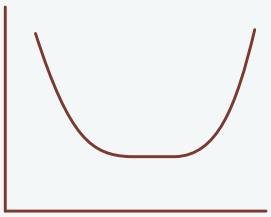
PROBABILISTIC GRAPH MODELING
FOR ENGINEERING DISCIPLINES · VOLUME II



Reliability Engineering by Derivation

Physical models, diagnosis, risk and acceptance

431 pages

Ricardo Marquez, Ph.D.

ricardo-marquez.com

Reliability Engineering by Derivation

Volume II

Physical models, diagnosis, risk and acceptance

A lesson sequence, written up as a book

Ricardo Marquez

October 31, 2026 (work in progress: one chapter per lesson, added as the lessons proceed)

How to read this book

This book teaches one skill: to take a reliability question stated in ordinary engineering language, write it as a probability or an expectation over the unknowns of a physical model, derive what that expression requires, compute it, check it, and act on it. Volume I holds Parts I to IV, Chapters 1 to 16: the method itself (a plant graph, a question as a probability over its state, and the five computations beneath every answer: the posterior, the trend, the evidence, the preposterior and the expected consequence), and its application to records of failures, lifetimes and repairs, to predictions of mission reliability, remaining life and future counts, and to the structure of systems (enumeration, cut sets, common cause, Markov chains). This second volume holds Parts V to VII, Chapters 17 to 28. It turns to the physical model itself: the state of a plant fitted to conservation relations and readings, the curvature that tells which entries the readings determine, the propagation of that uncertainty to engineering outputs; then to diagnosis (virtual sensors, bad data, fault isolation, model selection); and to risk, importance measures, acceptance tests and operating margins. Volume III holds Parts VIII and IX, Chapters 29 to 36: the value of information and the engineering decisions it supports, and four integrated case studies that run the whole method on one plant each. Chapter numbers run continuously across the three volumes; a reference to a chapter numbered 16 or below points to Volume I, whose results this volume builds on and does not repeat, and a chapter numbered 29 or above is named in words.

Every chapter follows the progression fixed in Volume I's preface and restated here in one line each: step 0, the plant graph; step 1, the question as a probability or expectation over the state; step 2, the graph extended for the question and the evidence attached as rows, with the assumptions they rest on; step 3, the posterior of the state, derived and computed once (the MAP state, the information and the covariance, with the exact posterior where one exists); step 4, the answers, each question read from that posterior down to its number (an entry or a gradient, a ratio of evidences, or an expected consequence); step 5, evaluation and checks against the full calculation; step 6, the report with its band: the probability, the Bayes factor, the expected consequence, or a crossing time with its band by inversion; step 7, the next measurement, ranked by the width it removes. The five kinds of question (estimation, prediction under a threshold, hypothesis comparison, diagnosis, design) and the five computations underneath them (the posterior, the trend, the evidence, the preposterior, the expected consequence) are those of Volume I. The notation chapter is reproduced here so that this volume can be read with Volume I closed, and the atlas of questions in the appendix now covers Volumes I and II.

The register is unchanged: every new quantity is introduced with its name, meaning, units and equation; every derivation proceeds one relation per line; every numerical example is worked with units; a *Jargon* box expands each shorthand term once. The plants of Volume I return: the equipment hall of Chapter 1, now two halls on one chilled-water loop with their sensors and meters; the pumps, valves and generators whose records were the evidence; and the redundant cooling plant, which grows into a data center in the case studies.

The framework the book teaches is physics-constrained probabilistic modeling: a plant written as a graph, its unknowns as a state vector, every piece of evidence as a row, and every question as a probability over that state. TERRA is the author's own implementation of this

framework; the book names the framework and the computations, not the software, and every number in it is reproducible from the derivations as written.

Each chapter closes with a statement of what to retain, a table of its terms with their meaning, units and the section that defines each, and exercises at three levels, formulation, derivation and application, each followed by its answer.

Contents

How to read this book **ii**

Notation and vocabulary **vii**

V Physical Models and Uncertainty Propagation **1**

17 Physics-constrained state estimation **2**
17.1 State estimation on a physical model 2
17.2 The posterior of the state (step 3) 13
17.3 The answers (step 4) 19
17.4 The checks (step 5) 26
17.5 The report (step 6) 29
17.6 The next measurement (step 7) 30
17.7 What to retain 33
17.8 Exercises 35

18 Maximum a posteriori estimation **37**
18.1 Fitting the state of a plant to its readings 37
18.2 The posterior of the state (step 3) 42
18.3 The answers (step 4) 47
18.4 The checks (step 5) 55
18.5 The report (step 6) 60
18.6 The next measurement (step 7) 61
18.7 What to retain 64
18.8 Exercises 66

19 The Hessian, covariance and identifiability **68**
19.1 The spread of a fitted state and what the readings determine 68
19.2 The posterior of the state (step 3) 73
19.3 The answers (step 4) 77
19.4 The checks (step 5) 90
19.5 The report (step 6) 96
19.6 The next measurement (step 7) 98
19.7 What to retain 101
19.8 Exercises 102

20 Propagating uncertainty to outputs **105**
20.1 Hall 1 at twice its present load 105
20.2 The posterior of the state (step 3) 110
20.3 The answers (step 4) 113

20.4 The checks (step 5) . . . 121
20.5 The report (step 6) . . . 126
20.6 The next measurement (step 7) . . . 128
20.7 What to retain . . . 132
20.8 Exercises . . . 134

VI Diagnosis and Competing Hypotheses 137

21 Virtual sensors and unobserved states 138
21.1 Quantities where no sensor is installed . . . 138
21.2 The posterior of the state (step 3) . . . 143
21.3 The answers (step 4) . . . 146
21.4 The checks (step 5) . . . 153
21.5 The report (step 6) . . . 159
21.6 The next measurement (step 7) . . . 160
21.7 What to retain . . . 164
21.8 Exercises . . . 165

22 Sensor bias and bad data 168
22.1 Readings that may be lying . . . 168
22.2 The posterior of the state under each drawing (step 3) . . . 174
22.3 The answers (step 4) . . . 178
22.4 The checks (step 5) . . . 194
22.5 The report (step 6) . . . 198
22.6 The next measurement (step 7) . . . 198
22.7 What to retain . . . 203
22.8 Exercises . . . 204

23 Fault detection and isolation 207
23.1 A plant that has changed . . . 207
23.2 The posterior of the state under each drawing (step 3) . . . 215
23.3 The answers (step 4) . . . 217
23.4 The checks (step 5) . . . 230
23.5 The report (step 6) . . . 234
23.6 The next measurement (step 7) . . . 235
23.7 What to retain . . . 239
23.8 Exercises . . . 241

24 Model selection and Bayes factors 244
24.1 Questions about the drawing . . . 244
24.2 The posterior of the state (step 3) . . . 252
24.3 The answers (step 4) . . . 259
24.4 The checks (step 5) . . . 269
24.5 The report (step 6) . . . 275
24.6 The next measurement (step 7) . . . 277
24.7 What to retain . . . 280
24.8 Exercises . . . 281

VII Risk, Importance and Acceptance 284

25 Risk and expected consequences 285
25.1 The expected loss of the hall's cooling . . . 285
25.2 The posterior of the state (step 3) . . . 291
25.3 The answers (step 4) . . . 292
25.4 The checks (step 5) . . . 309
25.5 The report (step 6) . . . 313
25.6 The next measurement (step 7) . . . 315
25.7 What to retain . . . 319
25.8 Exercises . . . 320

26 Reliability importance measures 322
26.1 The importance of each unit to the loss of cooling . . . 322
26.2 The posterior of the state (step 3) . . . 326
26.3 The answers (step 4) . . . 327
26.4 The checks (step 5) . . . 337
26.5 The report (step 6) . . . 340
26.6 The next measurement (step 7) . . . 342
26.7 What to retain . . . 346
26.8 Exercises . . . 347

27 Reliability demonstration and acceptance tests 349
27.1 Accepting a capacity read through a model . . . 349
27.2 The posterior of the state (step 3) . . . 354
27.3 The answers (step 4) . . . 356
27.4 The checks (step 5) . . . 367
27.5 The report (step 6) . . . 370
27.6 The next measurement (step 7) . . . 370
27.7 What to retain . . . 374
27.8 Exercises . . . 375

28 Operating margins and capacity risk 377
28.1 Margins on the path and on the chiller . . . 377
28.2 The posterior of the state (step 3) . . . 383
28.3 The answers (step 4) . . . 384
28.4 The checks (step 5) . . . 395
28.5 The report (step 6) . . . 400
28.6 The next measurement (step 7) . . . 400
28.7 What to retain . . . 404
28.8 Exercises . . . 405

A Atlas of probabilistic questions 408

Notation and vocabulary

Every symbol below is used with the same meaning in every chapter. The unknown always stands on the left of a conditioning bar and the evidence on the right. Vectors and matrices are bold.

Symbol	Unit	Meaning
The plant graph and its state		
$\mathbf{z}$, z_i	varies	State vector of the plant graph: every free quantity at once (temperatures, flows, conductances, log-rates); its i-th entry.
$\mathbf{y}$, y	varies	The evidence: all readings, counts and prior information treated as known; a single reading.
$r_k(\mathbf{z})$	varies	Row k of the model: a physical relation that must hold ($r_k = 0$), a reading compared with the state ($r_k = z_i - y$, in the units of the reading), or a prior ($r_k = z_i - z_i^0$).
w_k, σ_k	varies	Weight (precision) of row k and the standard deviation it corresponds to, $w_k = 1/\sigma_k^2$. A row may equivalently be written normalized, r_k/σ_k, with unit weight; both put the same term $\frac{1}{2}w_k r_k^2$ in the objective.
T, T_a	°C	Temperature of the equipment hall; ambient temperature (pinned).
Q, G	kW, kW/K	Heat load supplied to the hall; effective thermal conductance of the heat-removal path.
g, G_{ref}	–, kW/K	Log-conductance $\ln(G/G_{\text{ref}})$, the entry of the state that carries the conductance; its reference value.
y_T, σ_T	°C, K	A temperature reading of the hall and the standard deviation of its error.
$T_{\max}$, f^*	varies	A limit on a temperature; a requirement level on a state entry or on a derived quantity $f(\mathbf{z})$: the event $f \geq f^*$ is *the requirement holds*, and an upper limit is the complementary event (Chapter 2). From Part V a starred plant quantity (Q_1^*, T_1^*) is a scenario value: the quantity at a load not yet applied.
ω, ξ	–	Log-rate $\ln(\lambda/\lambda_{\text{ref}})$ and logit of a probability, $\ln(p/(1-p))$: the unbounded entries that carry a rate or a probability in the state.
u, v	–	Log-scale $\ln\eta$ and log-shape $\ln\beta$ of a Weibull lifetime, the two entries of the state in a lifetime fit.
Probability		
$\Pr(A)$, $\Pr(A \mid B)$	–	Probability of event A; conditional on event B.

Symbol	Unit	Meaning
$p(\cdot)$	varies	A probability density; $p(\mathbf{z} \mid \mathbf{y})$ is the posterior of the state, $p(\mathbf{y} \mid \mathbf{z})$ the likelihood, $p(\mathbf{z})$ the prior.
$\mathbb{E}$, Var, Cov	varies	Expectation, variance, covariance.
$\mathcal{N}(\mu, \sigma^2)$	–	Normal distribution with mean μ and variance σ^2; Φ its standard cumulative distribution function.
Beta(a, b), Gamma(a, b)	–	Beta distribution with shape parameters a, b; gamma distribution with shape a and rate b.
Binomial(n, p), Poisson(μ)	–	Number of failures in n trials with failure probability p; count with mean μ.
Inference objects		
$\mathcal{L}(\mathbf{z})$, $\mathcal{L}_k(\mathbf{z})$	–	The objective, minus the log posterior up to a constant, and the term row k puts in it: $\mathcal{L} = \sum_k \mathcal{L}_k$, with $\mathcal{L}_k = \frac{1}{2} w_k r_k(\mathbf{z})^2$ for a Gaussian row and $\mathcal{L}_k = -\ln p(y_k \mid \mathbf{z})$ for a count, lifetime or demand row (Section 4.2.7).
$\mathcal{L}'$, $\mathcal{L}''$, $\mathcal{L}'_k$, $\mathcal{L}''_k$	varies	For a state of one entry, say ω: the slope and the curvature of the objective, $\mathrm{d}\mathcal{L}/\mathrm{d}\omega$ and $\mathrm{d}^2\mathcal{L}/\mathrm{d}\omega^2$, and of row k's term; they add over the rows, and the Newton step is $\Delta\omega = -\mathcal{L}'/\mathcal{L}''$ (Chapter 5).
Working form	varies	The way a count row joins the Jacobian beside Gaussian rows (equation (4.43)): the line $\sqrt{\mu}\,\partial \ln \mu / \partial \mathbf{z}$ with the working residual $(\mu - k)/\sqrt{\mu}$ for a Poisson count k of mean μ, and $\sqrt{np(1-p)}\,\partial \xi / \partial \mathbf{z}$ with $(np - k)/\sqrt{np(1-p)}$ for a binomial count carried on the logit ξ. The pair reproduces the row's gradient and the $\mathbf{J}^\mathsf{T}\mathbf{J}$ part of its Hessian; its half square is not the row's term.
$\hat{\mathbf{z}}$	varies	The maximum a posteriori (MAP) state, the minimizer of $\mathcal{L}$.
$\mathbf{J}$	varies	Jacobian of the rows with respect to the state, evaluated at $\hat{\mathbf{z}}$: for a Gaussian row the line $\partial r_k / \partial z_i$, for a count row the line of its working form (Section 4.2.7).
$\mathbf{\Lambda}$, $\mathbf{H}$, $\mathbf{C}$	varies	Information matrix $\mathbf{J}^\mathsf{T}\mathbf{W}\mathbf{J}$ with $\mathbf{W} = \mathrm{diag}(w_k)$ ($\mathbf{J}^\mathsf{T}\mathbf{J}$ for normalized rows); the Hessian of $\mathcal{L}$, $\mathbf{H} = \mathbf{J}^\mathsf{T}\mathbf{J} + \mathbf{C}$, with $\mathbf{C}$ the curvature term the rows add beyond the information, zero for linear Gaussian rows and for count rows on their canonical parameter (equation (4.44)).
Λ	varies	For a state of one entry, the information, the curvature of the objective at the MAP, $\Lambda = \mathcal{L}''(\hat{\omega})$ (Chapter 5).
$\mathbf{\Sigma}$, Σ_{ij}	varies	Posterior covariance of the state under the Laplace approximation, $\mathbf{H}^{-1}$ ($\mathbf{\Lambda}^{-1}$ when $\mathbf{C} = \mathbf{0}$); its entries. The posterior spread of entry i is written $\sqrt{\Sigma_{ii}}$; σ_k always names a row's standard deviation and σ_p a physics row's tolerance.
σ^2_ω, σ_ω	varies	For a state of one entry, say ω or ξ: its Laplace variance $1/\Lambda$ and standard deviation (σ^2_ξ, σ_ξ for the logit). A derived quantity with derivative $g = \mathrm{d}f/\mathrm{d}\omega$ at the MAP has variance $g^2\sigma^2_\omega$.

Symbol	Unit	Meaning
ρ	K/kW	Thermal resistance of a path, $1/G$ (Chapter 4); not a correlation.
f_L, f_U	varies	The ends of the credible interval of a derived quantity $f(\mathbf{z})$ (Chapter 2).
y_{new}, $\tilde{y}_a$	varies	A future observation (the predictive) and the outcome of a candidate measurement a before it is taken (the preposterior).
M, M_j, B_{AB}	–	A drawing (model) of the plant; the members of a family; the Bayes factor $p(\mathbf{y} \mid M_A)/p(\mathbf{y} \mid M_B)$, written B_{10} for M_1 against M_0.
$f_t(\mathbf{z})$, $\mathbf{g}(t)$, $\bar{t}$	varies	A quantity read at time t from a state indexed by time, $f_t = \mathbf{g}(t)^{\mathsf{T}}\mathbf{z}$ with the gradient $\mathbf{g}(t) = (1, t - \bar{t})^{\mathsf{T}}$ about the center $\bar{t}$ of the record, the state holding a level entry α and a trend entry γ (the trend form, Chapter 2).
$U(a)$, u	varies	The expected value of a candidate action a over the outcomes it may produce, and the utility u of the updated posterior it is built from (the design form, Chapter 2).
$\mathbf{J}_a$, $\mathbf{\Lambda}_a$	varies	The Jacobian rows of a candidate measurement and its information $\mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a$ (expected counts at $\hat{\mathbf{z}}$ for a count row).
κ	–	Assurance level of an acceptance rule: the posterior probability a requirement must reach, $\Pr(R \geq R_{\min} \mid \mathbf{y}) \geq \kappa$ (Chapter 5).
π_0	–	The prior row's share of the information, $\mathcal{L}_0''/\mathcal{L}''$ for a state of one entry (Chapter 5); in Chapter 16 the stationary probability that no unit is up.
$f(\mathbf{z})$, $\mathbf{g}$	varies	A derived quantity of the state and its gradient $\partial f/\partial \mathbf{z}$ at $\hat{\mathbf{z}}$.
μ_f, σ_f^2	varies	Mean and variance of a derived quantity under the delta method: $f(\hat{\mathbf{z}})$ and $\mathbf{g}^{\mathsf{T}}\mathbf{\Sigma}\mathbf{g}$. From Part V the spread is written $\text{sd}(f \mid \mathbf{y}) = \sqrt{\mathbf{g}^{\mathsf{T}}\mathbf{\Sigma}\mathbf{g}}$.
$p(\mathbf{y} \mid M)$	–	Evidence (marginal likelihood) of model M.
B_{10}	–	Bayes factor of model M_1 against M_0, $p(\mathbf{y} \mid M_1)/p(\mathbf{y} \mid M_0)$; usually reported as a log in nats.
$\mathbf{\Upsilon}$, υ	varies	The gain of the correction form, $\mathbf{\Sigma}_0\mathbf{A}^{\mathsf{T}}(\mathbf{A}\mathbf{\Sigma}_0\mathbf{A}^{\mathsf{T}} + \mathbf{R})^{-1}$, which moves the prior mean toward the readings; its scalar value for one reading, the weight $\sigma_0^2/(\sigma_0^2 + \sigma_T^2)$ (Chapter 3).
$\mathbf{\Sigma}_a$	varies	Preposterior covariance after a candidate measurement a with information $\mathbf{\Lambda}_a$: $(\mathbf{\Sigma}^{-1} + \mathbf{\Lambda}_a)^{-1}$.
EIG	nat or bit	Expected information gain of a candidate measurement.
Reliability quantities (as in the companion book)		
T, $R(t)$, $F(t)$	h, –, –	Time to failure; reliability $\Pr(T > t)$; cumulative failure probability.
$f(t)$, $\lambda(t)$, $\Lambda(t)$	h^{-1}, h^{-1}, –	Failure density; hazard function; cumulative hazard.
λ, μ	h^{-1}	Constant failure rate; constant repair rate.
p, k, n	–	Failure probability per demand; number of failures observed; number of trials.

Symbol	Unit	Meaning
T_{tot}	h	Exposure (total operating time on test) behind a failure count.
β, η	–, h	Weibull shape and scale (characteristic life).
t_i, δ_i, c_i	h, –, h	Observed time of unit i; its event indicator (1 failed, 0 censored); its censoring time.
t_{50}, t_0, T_{new}	h	Median lifetime; an age already reached; the lifetime of a new unit.
ℓ, ℓ_p	–	Log-likelihood; profile log-likelihood (maximized over the other entries).
r, D	–, h	Number of repairs observed; their total duration.
U, A_{req}	–	Unavailability $1 - A$; a required availability.
α, γ	–, yr^{-1}	The log-rate case of the level and trend entries of the trend form: log-rate at the record's center and its trend per year in $\ln \lambda(t) = \alpha + \gamma\,(t - \bar{t})$; t^* the time the rate crosses a limit.
K_f, τ	–, h or yr	A future failure count; the future period or horizon it is counted over.
q_i, $\mathbf{s}$, $\phi(\mathbf{s})$	–	Unavailability of unit i; a configuration of units (up or down); the structure function, 1 when the configuration delivers the service.
R_{sys}, U_{sys}	–	Service availability and unavailability of several units under a drawing, by the 2^n enumeration.
β_{cc}	–	Common-cause share: the fraction of a unit's failures that fail all units sharing the cause (β is the Weibull shape).
A, MTTF, MTTR	–, h, h	Steady-state availability; mean time to failure; mean time to repair.
$L(a, \mathbf{z})$, $\mathcal{R}(a)$, C, C_e, c_a	R$	The loss of action a when the state is $\mathbf{z}$, its consequence; its expected value over the posterior, the risk $\mathcal{R}(a) = \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$; the consequence of an event, of event e; the fixed cost of the action (the expected-consequence form, Chapter 2).
Physical models and the state (Part V); a row naming a chapter is that chapter's own symbol		
$\mathbf{z}_{\text{full}}$, $\mathbf{r}_{\text{full}}$, $\mathbf{J}_{\text{full}}$	varies	The full state of a plant graph, every free quantity including the temperatures and meter quantities the balances fix; its residual vector and Jacobian (Chapter 17).
$h_i(\mathbf{z})$, σ_p, $\sigma_{p,P}$	varies, K, kW	The balance of node i as a function of the state; the tolerance of a hall's physics row; the tolerance of the header's (Chapters 17 and 18).
$\mathbf{J}_{\text{e}}$, $\mathbf{J}_{\text{p}}$, $\mathbf{J}_{\text{d}}$, $\mathbf{J}_\pi$	varies	The Jacobian of the evidence rows (readings and priors) and of the physics rows (Chapter 17); the design columns, the reading rows alone, and the prior rows, with information $\mathbf{\Lambda}_{\text{prior}} = \mathbf{J}_\pi^{\mathsf{T}} \mathbf{J}_\pi$ (Chapters 19 and 22).
$\mathbf{N}$, $\mathbf{N}_{\text{b}}$	varies	Null-space basis of the physics rows, $\mathbf{J}_{\text{p}}\mathbf{N} = \mathbf{0}$, whose columns are the free directions of the reduced state $\mathbf{z}$, with reduced Jacobian $\mathbf{J} = \mathbf{J}_{\text{e}}\mathbf{N}$; its block on the balanced entries (Chapter 17).

Symbol	Unit	Meaning
$\tilde{Q}_i$	kW	A load profiled out: the least-squares value with the other entries held at the prior point (Chapter 17).
ν, n_r, n_z	–	Redundancy, the number of rows less the number of entries, $\nu = n_r - n_z$: the degrees of freedom of $2\mathcal{L}(\hat{\mathbf{z}})$.
χ^2_ν, $\chi^2_\nu(c)$	–	Chi-square distribution with ν degrees of freedom; noncentral, with noncentrality c (Chapter 18).
$\mathbf{H}$, $\mathbf{S}$	varies	The exact Hessian of the objective, $\mathbf{H} = \mathbf{J}^\mathsf{T}\mathbf{J} + \mathbf{S}$, and its curvature term $\mathbf{S} = \sum_k r_k \nabla^2 r_k$, which Gauss–Newton drops; a subscript names the variable differentiated ($\mathbf{H_q}$, Chapter 25). $\mathbf{H}$ is a Hessian throughout the book.
$\Delta\mathbf{z}$, tol, ϖ	varies, –, –	The Gauss–Newton step, $\mathbf{J}^\mathsf{T}\mathbf{J}\,\Delta\mathbf{z} = -\mathbf{J}^\mathsf{T}\mathbf{r}$; the stopping tolerance on $\max_i \lvert\Delta z_i\rvert/(1+\lvert z_i\rvert)$; the Levenberg–Marquardt damping (Chapter 18). Chapter 23 writes the step $\mathbf{s}$, since there $\mathbf{\Delta}$ is a change between periods.
$\mathbf{j}_k$, $\mathbf{j}_a$	varies	The Jacobian row of reading k as a column, $\mathbf{j}_k = \mathbf{J}_k^\mathsf{T}$; of a candidate row a.
$\mathbf{K}$, K_{kk}, $1 - K_{kk}$	–	The hat matrix $\mathbf{J}(\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}\mathbf{J}^\mathsf{T}$; the leverage of row k, $\mathbf{j}_k^\mathsf{T}\mathbf{\Sigma}\mathbf{j}_k$; its redundancy number, the share of row k the other rows check, summing to ν (Chapters 18, 21 and 22).
r_k^N, r^*, p_fa	–	The normalized residual $r_k(\hat{\mathbf{z}})/\sqrt{1-K_{kk}}$; its test threshold, by Bonferroni $\Phi^{-1}(1 - p_\mathrm{fa}/2n_r)$; the false-alarm probability of the family of tests (Chapters 18 and 22).
s_j, $\mathbf{v}_j$, $\mathbf{V}$, $\mathbf{F}$	varies	The singular values of $\mathbf{J}$ in decreasing order and its right singular vectors, the columns of $\mathbf{V}$; the left factor, $\mathbf{J} = \mathbf{F}\,\mathrm{diag}(s_j)\mathbf{V}^\mathsf{T}$ ($\mathbf{W}$ is the weight matrix). The weakest direction $\mathbf{v}$, of spread $1/s_\mathrm{min}$ (Chapters 17 and 19).
$\mathrm{cond}(\cdot)$	–	Condition number, s_1/s_n for $\mathbf{J}$ and its square for $\mathbf{\Lambda}$ (Chapter 19); κ is the assurance level.
Λ_{ij}, ϱ_i^2, $\sigma_{0,i}$	varies	The entries of the information $\mathbf{\Lambda}$, $1/\Lambda_{ii}$ the conditional variance of entry i; its squared multiple correlation $1 - 1/(\Lambda_{ii}\Sigma_{ii})$; its prior spread (Chapter 19).
$\mathbf{\Lambda}_{-a}$, $\mathbf{\Sigma}_{-a}$, χ_p	varies	The information and covariance with the rows a removed; the radius of the credible ellipse of probability p (Chapter 19).
Propagation, virtual readings and diagnosis (Parts V and VI)		
$\mathrm{sd}(f \mid \mathbf{y})$	varies	The posterior standard deviation of a derived quantity, $\sqrt{\mathbf{g}^\mathsf{T}\mathbf{\Sigma}\mathbf{g}}$ by the delta method; it replaces σ_f from Part V on.
$[\mathbf{g}]_i$, V_i	varies	Entry i of the gradient; the diagonal term $[\mathbf{g}]_i^2\Sigma_{ii}$ of the variance budget (Chapter 20).
$\mathbf{d}$, $m(x)$, $v(x)$	varies	The deviation $\mathbf{z} - \hat{\mathbf{z}}$ (Chapter 20); the conditional mean and variance of the other entries given one entry x under the Laplace normal, $\Sigma_{5\mid1}$ the conditional variance of entry 5 given entry 1 (Chapters 20 and 21).

Symbol	Unit	Meaning
ζ_j, $\bar{\zeta}_j$, N_{eff}	–	The importance weight of draw j and its normalized value; the effective sample size $(\sum_j \bar{\zeta}_j^2)^{-1}$ (Chapter 20).
$\mathbf{y}_-$, $\mathcal{L}_-$, $\hat{\mathbf{z}}_-$, $\boldsymbol{\Sigma}_-$	varies	The evidence, objective, MAP state and covariance of the refit without one reading row; $\tilde{\boldsymbol{\Sigma}}_-$, $\tilde{\boldsymbol{\Lambda}}_-$ the rank-one downdate of $\boldsymbol{\Sigma}$ and $\boldsymbol{\Lambda}$ at $\hat{\mathbf{z}}$ (Chapter 21).
V_{-k}, V_a	varies	The variance of reading k predicted from the other rows, $\mathbf{g}_k^{\mathsf{T}}\boldsymbol{\Sigma}_-\mathbf{g}_k$; of a reading at a place a without one (Chapters 21 and 22).
$\text{sd}(y_k \mid \mathbf{y}_-)$, d_k	varies, –	The predictive spread of a held-out reading, $\sqrt{\sigma_k^2 + V_{-k}}$; the held-out departure $(y_k - f_k(\hat{\mathbf{z}}_-))/\text{sd}(y_k \mid \mathbf{y}_-)$ (Chapter 21).
T_r, $\dot{m}c_p$	°C, kW/K	The return temperature of the chilled-water loop; its heat-capacity flow (Chapters 21, 22 and 24).
b_k, σ_b, $\hat{b}_k$, b_k^*	K or kW	The bias entry of reading k, its prior spread and its estimate; the minimum detectable bias (Chapter 22).
B_{k0}, ψ_k	–	The Bayes factor of a bias on reading k against no bias; the ratio $\sigma_b^2(1-K_{kk})/\sigma_k^2$ of its closed form (Chapter 22).
$\mathcal{F}$, $\lvert\mathcal{F}\rvert$	–	A family of candidate drawings and its size; the posterior over the family is $\Pr(M_j \mid \mathbf{y})$ for $M_j \in \mathcal{F}$ (Chapters 22 and 23).
N, $\bar{y}_k$	d, varies	The number of days in a record and reading k averaged over them, with spread $\sigma_k/\sqrt{N}$ (Chapter 23; bold $\mathbf{N}$ is the null-space basis).
σ_{h}, σ_{w}	varies	The healthy model's prior spread of an entry, from the earlier period; the widened spread of a candidate drawing (Chapter 23).
$\boldsymbol{\Delta}$, Z_i, χ^2_Δ	varies, –, –	The change of the state between two periods, $\hat{\mathbf{z}}_2 - \hat{\mathbf{z}}_1$; its standardized entry; the joint statistic $\hat{\boldsymbol{\Delta}}^{\mathsf{T}}(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)^{-1}\hat{\boldsymbol{\Delta}}$ (Chapter 23).
$\Pr(M_d)$, $\ln B_{10}$	–	The prior probability of drawing M_d; the log Bayes factor in nats (Chapter 24).
$\text{KL}(A \,\|\, B)$, ς_d	nat, varies	The Kullback–Leibler divergence of the predictive under A from that under B; the predictive spread of a candidate reading under drawing d (Chapter 24).
c_{f}, Σ_{nn}, U_{c}	varies	The density of a flat prior; the posterior variance of an added entry; the unavailability of the cooling service (Chapter 24).
Risk, importance, acceptance and margins (Part VII)		
$\mathbf{q}$, $\boldsymbol{\Sigma}_{\mathbf{q}}$	–	The unavailabilities of the units as a vector and their posterior covariance (Chapter 25).
t_{yr}, $\mathcal{R}_{\text{yr}}$, L_{yr}	h, R$/yr, R$	The hours of a year, 8760 h; the annual risk $t_{\text{yr}}\mathcal{R}$, the expected loss of a year given $\mathbf{q}$; the loss of one particular year (Chapter 25).
$\mathcal{R}^*$, w_a	R$/yr, decades	A limit on the annual risk; the width of the central 95 % band of the risk after candidate a (Chapter 25).
ν_{loss}, MDT	yr^{-1}, h	The frequency of loss-of-service events; the mean down time of one (Chapter 25).

Symbol	Unit	Meaning
VaR_p, CVaR_p	R$	The p quantile of the predictive year's loss; the tail mean beyond it (Chapter 25).
U_i^+, U_i^-	–	The top event with unit i pinned down ($q_i = 1$) and pinned up ($q_i = 0$) (Chapter 26).
$I_{\mathrm{B},i}$, $\mathbf{I}_\mathrm{B}$	–	Birnbaum importance $U_i^+ - U_i^- = \partial U/\partial q_i$, the probability that unit i is critical; the vector of them (Chapter 26).
$I_{\mathrm{FV},i}$, $I_{\mathrm{CR},i}$	–	Fussell–Vesely importance $(U - U_i^-)/U$, the share a perfect unit removes; criticality importance $q_i I_{\mathrm{B},i}/U$, equal to it for units independent given $\mathbf{q}$ (Chapter 26).
RAW_i, RRW_i	–	Risk achievement worth U_i^+/U and risk reduction worth U/U_i^- (Chapter 26).
V_j, n_j	–, –	Uncertainty importance of shared entry j, its term $(g_j/U)^2\Sigma_{jj}$ in the variance of $\ln U$; the checks of its record (Chapter 26).
G^*, g^*, z_κ	kW/K, –, –	A required conductance and its log-conductance; the standard normal quantile $\Phi^{-1}(\kappa)$ (Chapters 27 and 28).
Q_test, ΔT^*, σ_c	kW, K, K	The held load of a test; the rise at the requirement, Q_test/G^* (the star marks the requirement level); the combined supply spread of sensor and set point (Chapter 27).
$\mathrm{OC}(\mathbf{z}; a)$, $U(a)$	–	The operating characteristic, the probability that test a accepts at a stated truth $\mathbf{z}$; averaged over the posterior, the probability the test ends accepted (Chapter 27).
ε_p, ε_c, π_g	–	The producer's and consumer's risks of a test; the prior's share Σ_{11}/σ_g^2 of the test's posterior (Chapter 27).
m_1^*, Q_cap, Q_κ	K, kW, kW	The margin $T_\mathrm{max} - T_1^*$ at a scenario load; the path's capacity $(T_\mathrm{max} - T_s)\,\mathrm{e}^{g_1}$; the largest load held with assurance κ (Chapter 28).
c_ch, P_pl, P^*, P_hr	–, kW, kW, kW	The chiller's capacity factor; its plate rating; the demand on it at a scenario; its headroom $P_\mathrm{pl}c_\mathrm{ch} - Q_1 - Q_2$ (Chapter 28).
ΔT_s, $(x)^+$, b_P	K, –, kW	The supply's rise when demand exceeds the delivered capacity; the positive part $\max(0, x)$; the bias of the chiller's heat meter (Chapters 23 and 28).

Vocabulary

Plant graph The representation of the physical system as nodes (places where a conserved quantity balances), edges (paths along which it flows) and sources, with the relations that must hold among them. Built before any question is asked.

State vector The collection of free quantities the inference estimates at once; outputs that are entries of the state need no further transformation.

Pinned, free A quantity treated as known (boundary condition) against one carried in the state.

Residual, row A relation written so that it equals zero when satisfied; physical rows constrain the state, reading rows compare it with evidence, prior rows carry handbook knowledge.

Prior, likelihood, posterior Knowledge before the evidence; the probability of the evidence given a state; knowledge after the evidence, by Bayes' theorem.

Marginalization Integrating a joint distribution over the unknowns not asked about; the law of total probability.

Aleatory, epistemic Variability in outcomes the model treats as random; incomplete knowledge of quantities that have one true value.

Maximum a posteriori (MAP) estimate The state at which the posterior density is highest; the minimizer of the objective.

Laplace approximation Replacing the posterior by the Gaussian whose covariance is the inverse curvature of the objective at the MAP state.

Information matrix The matrix $\mathbf{J}^\mathsf{T}\mathbf{W}\mathbf{J}$, the Hessian of the objective without the curvature term $\mathbf{C}$ of $\mathbf{H} = \mathbf{J}^\mathsf{T}\mathbf{J} + \mathbf{C}$ (equation (4.44)); precisions of independent rows add. For a state of one entry it is the number Λ.

Delta method Linearizing a derived quantity around the MAP state so that its variance is $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$.

Change of variables Carrying a density to a new coordinate $u = h(x)$:

$$p_u(u) = p_x(h^{-1}(u))\,|\mathrm{d}x/\mathrm{d}u|,$$

the Jacobian factor keeping the probability of each interval; for several entries the factor is $|\det \partial\mathbf{z}/\partial\mathbf{u}|$. It gives the shape of a derived quantity, where the delta method gives only its spread (Section 4.3.3).

Variance budget The expansion of $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$ into one term per state entry plus the cross terms from their covariance.

Threshold probability The probability that a state entry or derived quantity is below (or above) a limit; a normal tail under the Laplace approximation.

Posterior predictive The distribution of a future observation obtained by averaging its model over the posterior.

Credible interval An interval that contains the unknown with a stated posterior probability.

Evidence, Bayes factor The normalizing constant of one model's posterior; the ratio of two models' evidences, which compares the models rather than their parameters.

Preposterior The posterior that would result from a measurement not yet taken; the object that values a measurement before buying it.

Identifiability Whether the evidence determines the state entries separately or only in combination; read from the shape of the posterior, not from marginal spreads.

The five question categories Estimation, prediction under a threshold, hypothesis comparison, diagnosis and design, answered by a marginal, a tail of the posterior or of the trend, an evidence ratio, a ranking of evidences, and a preposterior or an expected consequence.

The five computations The objects that answer every category (Chapter 2): the *posterior* $p(\mathbf{z} \mid \mathbf{y})$, read for an entry or a derived quantity; the *trend*, the posterior over a state indexed by time or window with a trend entry, read as the probability that the trend is positive, as a crossing time with its band, or as a forecast; the *evidence* $p(\mathbf{y} \mid M)$ and its ratios; the *preposterior*, the expected posterior under a measurement not yet taken; and the *expected consequence* $\mathcal{R}(a) = \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$ of an action over the posterior, with the choice of the action that minimizes it (risk is probability times consequence).

Factor graph A graph of the state's entries and the factors (the rows) that bear on them; the plant graph with its rows is one, and the posterior is the normalized product of its factors (Chapters 1 and 4).

Variable elimination Integrating the entries a question does not name out of the joint distribution one at a time; for a Gaussian posterior it is the Schur complement, and its cost on a large sparse graph is set by the order of elimination (Chapter 3).

Belief propagation Computing the marginals of a graphical model by passing messages between neighboring factors and entries (the sum-product algorithm); exact on a graph without loops, where on a Gaussian graph it returns the Schur complements. Not used in this book, whose posteriors are factored jointly.

Kalman filter and smoother Variable elimination on a state indexed by time, taken in time order: the filter forward, the smoother back.

Importance sampling (reweighted draws) Drawing from an approximation, here the Laplace normal, and weighting each draw by the ratio of the exact posterior to it, so that the draws check the approximation (Chapter 20, in Volume II).

Markov chain Monte Carlo Drawing from a posterior by a Markov chain whose stationary distribution it is (Gibbs sampling, the Metropolis algorithm); the route when neither the Laplace normal nor reweighting is adequate, as for a hierarchical model (Chapter 33, in Volume III).

Explaining away Two causes independent before an effect is seen become dependent once it is seen: evidence for one lowers the probability of the other, as a biased sensor and a changed conductance compete for one disagreeing reading (Chapters 22 and 23, in Volume II).

Physics row, tolerance A balance of the plant graph written as a row with a small standard deviation; in the limit of zero tolerance the balance is exact and the state reduces to its free directions.

Reduced state The free directions left once the balances are exact, the columns of the null-space basis of the physics rows.

Profiling Holding some entries fixed and fitting the others by least squares, to see what the readings say about the fixed ones.

Predictive check Asking whether the readings are probable under the model fitted to them, by the misfit $2\mathcal{L}(\hat{\mathbf{z}})$ against its chi-square law or by a held-out reading against its predictive.

Singular value, weakest direction The scale at which the rows see a direction of the state; the combination of entries with the smallest singular value, which the evidence determines least.

Observability Whether the readings determine an entry at all: an unobservable direction is one the design columns do not see, and only a prior gives it a finite spread.

Condition number The ratio of the largest singular value to the smallest; it bounds the loss of digits in solving with the information matrix.

Virtual sensor The posterior of a quantity no instrument reads, pushed from the state through its function.

Leave-one-out (held-out) check Refitting without a reading and comparing the reading with the predictive of the remaining rows.

Redundancy number The share of a reading the other rows check, $1 - K_{kk}$; zero for a critical measurement, which no other row checks.

Minimum detectable bias The bias on a reading that the residual test flags with a stated probability.

Footprint The pattern a candidate fault leaves on the readings: its column of the Jacobian on the reading rows.

Fault detection and isolation, model-based diagnosis Deciding that the plant has departed from its healthy model, and which component did, by comparing the evidence of candidate drawings generated from the plant graph.

Occam factor The part of a model's evidence that penalizes the prior range its parameters were given and the readings did not use.

Savage–Dickey ratio The Bayes factor of a point value of one entry against the model that frees it, read as the ratio of the posterior to the prior density at that value.

Expected evidence gain The expected log Bayes factor a measurement not yet taken would add between two drawings, the Kullback–Leibler divergence of their predictives.

Value at risk, tail mean A quantile of the predictive loss of a year, and the mean of the loss beyond it.

Importance measures Birnbaum (the slope of the top event in a unit's unavailability), Fussell–Vesely (the share a perfect unit would remove), risk achievement and reduction worths (the factors of an outage and of a perfect unit); each a derived quantity of the state, with a band.

Pivotal decomposition The top event as a straight line in one unit's unavailability, between the unit pinned down and pinned up.

Uncertainty importance A shared entry's term in the variance of the top event: what its record leaves uncertain, as against what its unit carries.

Operating characteristic The probability that a test accepts, as a function of the true value; its complement at a good value is the producer's risk, its value at a bad one the consumer's risk.

Acceptance sampling by variables Accepting on a measured quantity read through a model, rather than on a count of failures (by attributes).

Operating margin, capacity The room between a predicted quantity and its limit at a load; the load at which the limit is reached. Their bands are posteriors, and the load held with an assurance is a tail inverted.

Derating A unit delivering less than its plate rating, estimated as a capacity factor against the plate as a point null.

Part V

Physical Models and Uncertainty Propagation

Chapter 17

Physics-Constrained State Estimation

17.1 State estimation on a physical model

Two data halls are cooled from one chilled-water loop, and on a steady afternoon its five instruments report: the hall sensors 17.654545 and 17.700°C, the supply header 12.10°C, the chiller's heat meter 52.5 kW and the electrical meter on hall 1's equipment 31.0 kW. No instrument reads a conductance, the quantity that falls first when a hall's coil fouls, and none reads hall 2's load. The first reading of the five numbers is a hand calculation. Hall 2's load is the heat meter less hall 1's, $52.5 - 31.0 = 21.5$ kW, and each conductance is its load over its hall's rise above the supply: $31.0/5.5545 = 5.58$ kW/K for hall 1 and $21.5/5.60 = 3.84$ kW/K for hall 2, against a handbook value of 5 kW/K for both. Every reading is reproduced exactly, and hall 2 looks about a quarter below the handbook.

The engineer responsible for the loop asks what the calculation leaves unsaid. It uses each reading once and has none to spare, so it would reproduce a faulty sensor just as exactly; it states no spread, so whether 3.84 differs from 5 by more than the readings' errors allow is not something it can say; and it sets aside the handbook value and the chiller's set point, which were known before any reading. Which entries the readings determine, how well, and whether the readings can be checked against each other are properties of the rows that tie each reading to the unknowns, and they can be read before the state is fitted. The question of this chapter is

how the state of a plant with several free quantities is written as one vector, how each reading is tied to that vector through a physical row, and what the residual vector, its weights and its Jacobian establish about the readings before the state is fitted.

It is answered with the rows of Chapter 4, physics rows with a tolerance beside reading and prior rows on one state vector. What changes from the chapters on records of events is the evidence: readings of temperatures and of heat on a plant whose balances tie them to the unknowns, and with it a state of eight entries of which the balances leave five directions free.

Jargon: state estimation, physics-constrained, measurement model, residual vector, normalized residual, physics row, model tolerance, Jacobian, null directions, design columns, redundancy, prior point, linearization, normal equations, Gauss–Newton step, predictive check, profiling, singular value

State estimation is the inference of the free quantities of a physical system from readings that are related to them through a model; in power systems and process plants it is usually a weighted least-squares fit of the state to the readings. It is *physics-constrained* when the model is the system's own conservation relations, so that the estimate satisfies them. The *measurement model* of a reading states the reading as a function of the state plus an error. The *residual vector* collects every row of the model, each written as predicted minus observed; a *normalized residual* is a row divided by the standard deviation of its error, a dimensionless number of standard deviations. A *physics row* is a balance of the plant written as a residual and divided by a *model tolerance* σ_p, the standard deviation by which the balance is allowed to fail; the exact balance is the limit $\sigma_p \to 0$. The *Jacobian* is the matrix of derivatives of the rows with respect to the entries of the state. The *null directions* of the physics rows are the moves of the state that leave their linearization at zero, the directions along which the balances keep holding; the *design columns* are the columns of the reading rows taken along those directions, which state what the readings alone can see of the states that satisfy the balances. *Redundancy* is the number of rows beyond the number of entries of the state, the degrees of freedom left for a check. The *prior point* is the state at which every entry with a prior sits at its prior mean. *Linearization* replaces each row by its tangent at a point; the *normal equations* are the linear equations whose solution minimizes the sum of squares of linearized rows, and the step they give is the *Gauss–Newton step*. A *predictive check* compares a reading with the distribution the model gives it before it is used. *Profiling* an entry replaces it by the value that minimizes the objective with the other entries held. The *singular values* of a matrix are the lengths by which it scales its orthogonal input directions, the right singular vectors; the smallest says how weakly the least-seen direction is seen.

Key idea

A reading is evidence about the state only through a physical row. In the full state of the plant a hall sensor reads one entry, the hall temperature; it sees the conductance, the load and the supply temperature only because the hall's physics row ties that temperature to all three. Projected onto the null directions of the physics rows, the readings and priors become rows over the free entries alone, with a Jacobian that is the projection of their own and a covariance that is the full state's in the limit of exact balances. The formulation, the rows with their weights and their Jacobian, decides which entries the readings determine, how well, and whether they can be checked, before any state is fitted; a formulation with no rows to spare fits any readings exactly and checks nothing.

17.1.1 From records of events to the state of a plant

In Parts II to IV of Volume I the state was small and the evidence was a count or a list of times. A failure rate was one entry of the state informed by a Poisson row; a lifetime model was two entries informed by a likelihood of observed and censored times; a common-cause share was one more entry informed by the count of simultaneous failures. The physical system stood behind those rows, but no row contained its physics; only the single hall of Part I, with one entry of interest, had a physical row. Part V reverses the emphasis. The plant is the equipment hall of Chapter 1, now two halls on one chilled-water loop, and the evidence is the set of readings that its sensors and meters report in one steady operating condition. The full state has eight entries, two conductances, two loads, three temperatures and the chiller's heat, of which the balances

leave five directions free, and every reading is connected to them by conservation of energy.

This is the setting the field calls state estimation. In electrical networks the state is the set of bus voltages and the readings are power flows and injections; in water networks the state is the set of heads and demands and the readings are pressures and flows; in a cooling plant the state is the set of temperatures, loads and conductances and the readings are the sensors and meters of the plant. The state-estimation chapter of the companion volume *Reliability Engineering* states the method and its vocabulary. Here its formulation is derived for the two halls from the probability rules of Volume I, so that the fit of Chapter 18, the spread and the observability of Chapter 19 and the propagation of Chapter 20 have one formulation to act on.

Part V asks five questions of the fitted plant, and the formulation of this chapter is what they act on. The first asks for the two halls' conductances and loads and the supply temperature, given the readings. By Bayes' theorem the posterior of the state $\mathbf{z}$ is

$$p(\mathbf{z} \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid \mathbf{z})\, p(\mathbf{z})}{p(\mathbf{y})}, \tag{17.1}$$

where $p(\mathbf{y} \mid \mathbf{z})$ is the probability of the readings given the state, built from their measurement models (Section 17.1.7), and $p(\mathbf{z})$ is the prior; every entry wanted is an entry of $\mathbf{z}$, so the answer is a marginal of this posterior, its mode and covariance at the fitted state (Chapters 18 and 19). The second asks whether the readings and the physics can be true together, a tail of the minimized objective against the chi-square of the redundant rows (Chapter 18). The third asks which directions of the state the readings resolve at the fitted state, a property of the posterior's shape rather than of any single marginal (Chapter 19). The fourth asks for hall 1's temperature at a future load of 60 kW and whether it stays below 25°C. It extends the graph with a scenario node, hall 1 at $Q_1^* = 60\,\mathrm{kW}$ with the same conductance and the same supply temperature, whose temperature is $T_1^* = T_s + Q_1^*\,\mathrm{e}^{-g_1}/G_{\mathrm{ref}}$, with T_s the supply temperature and g_1 hall 1's log-conductance (Section 17.1.4); the probability is the posterior mass of the states that satisfy the event,

$$\Pr(T_1^* < 25^{\circ}\mathrm{C} \mid \mathbf{y}) = \int \mathbf{1}\{T_s + 60\,\mathrm{e}^{-g_1}/G_{\mathrm{ref}} < 25\}\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}. \tag{17.2}$$

In Chapter 1 the boundary temperature was pinned and the same event was a threshold on the conductance alone; here it is $g_1 > \ln\!\big(60/(G_{\mathrm{ref}}(25 - T_s))\big)$, a threshold that moves with the free supply temperature, so the answer depends on two entries and on their correlation (Chapter 20). The fifth asks which measurement to take next for that temperature, the preposterior covariance of a candidate reading (Chapter 20). The chapters on bad data, on faults and on model selection, Chapters 22 to 24, ask of the same plant whether a reading is in error, which fault the readings point to and which drawing of the plant they support, from the residuals and from the evidence.

17.1.2 A reading seen through a physical row

Chapter 1 made the central observation on a single hall: the likelihood of the temperature reading does not depend on the conductance, and the reading informs the conductance only through the physics row that ties the temperature to it. With one entry of interest that observation could be made in one line. With eight entries and five readings it becomes the organizing fact of the formulation. Each reading is a reading of one entry: a hall sensor of its hall temperature, the heat meter of the chiller's heat, the load meter of one load. The balances are rows of their own, and a reading sees the other entries only along them. The hall balance moves the hall temperature with the conductance, the load and the supply temperature at once, so a hall sensor sees those three together and none of them alone. The header balance moves the chiller's heat with the sum of the two loads, so the heat meter sees the sum and neither load separately. Which combinations of the entries the readings determine, and which only the

priors determine, is fixed by these physics rows, and it can be read from their derivatives, a column and a null direction at a time, before the state is fitted.

Principle 17.1. *A reading constrains the state only through the physical rows that tie its entry to the others. The pattern of the rows' derivatives, the Jacobian, states which entries each row touches before any reading is processed, and an entry that no reading reaches, directly or along a physics row, is determined by its prior alone.*

17.1.3 The plant graph (step 0)

Two data halls are cooled from one chilled-water loop. The chiller supplies water at the supply temperature T_s to a header that feeds the cooling coils of both halls. Each hall's equipment dissipates a heat load Q_i, in kW, and the hall's heat-removal path, the coil with its air and water flows, carries that heat to the chilled water with an effective conductance G_i, in kW/K; the subscript $i = 1, 2$ names the hall. The warmed water returns to the chiller, which removes the heat from the loop. Figure 17.1 draws the plant graph. The two green boxes on the left and the right are the halls, free nodes whose temperatures T_1 and T_2 follow from the balances derived below; the white boxes above them are the heat loads, sources of unknown strength; the green box in the middle is the supply header, a free node carrying the supply temperature T_s; the white box at the bottom is the chiller, the source of chilled water and the sink of the heat. The solid green arrows are the physical paths of heat: from each load into its hall, from each hall to the header through the conductance G_i, and from the header to the chiller. The drawing routes the heat through the header node, whose temperature is the one the coils see; physically the heat travels back to the chiller in the return line, which the steady balance below does not need. The five dashed boxes are the readings, connected by dashed arrows to the node or source each one measures.

The graph is the system as it is, before any question is asked. Nothing in it is pinned except the reference value of the conductance; in Chapter 1 the ambient temperature was a pinned boundary, and here its place is taken by the supply temperature, which is free, because the chiller's control holds it near a set point but not exactly at it.

17.1.4 The balances of the halls and the header

Chapter 1 derived the steady energy balance of a hall with one heat-removal path, equation (1.1): the heat supplied equals the heat removed, and the heat removed is the conductance times the temperature difference across the path. For hall i the path ends at the chilled water at T_s instead of at the ambient air, and the same steps give

$$Q_i = G_i\,(T_i - T_s), \qquad i = 1, 2, \tag{17.3}$$

with Q_i in kW, G_i in kW/K and the temperatures in °C. Since $G_i > 0$, dividing by G_i and adding T_s,

$$T_i = T_s + \frac{Q_i}{G_i}, \qquad i = 1, 2 : \tag{17.4}$$

each hall sits above the supply temperature by its temperature rise Q_i/G_i. The supply header is a node too. In steady state the water of the loop stores no heat, and if the pump work and the heat gained by the pipes are neglected, the heat the chiller removes equals the heat the two halls deliver,

$$P = Q_1 + Q_2, \tag{17.5}$$

with P the chiller's heat duty in kW. Equations (17.4) and (17.5) are the physical rows of the plant, three relations that must hold among its eight free quantities.

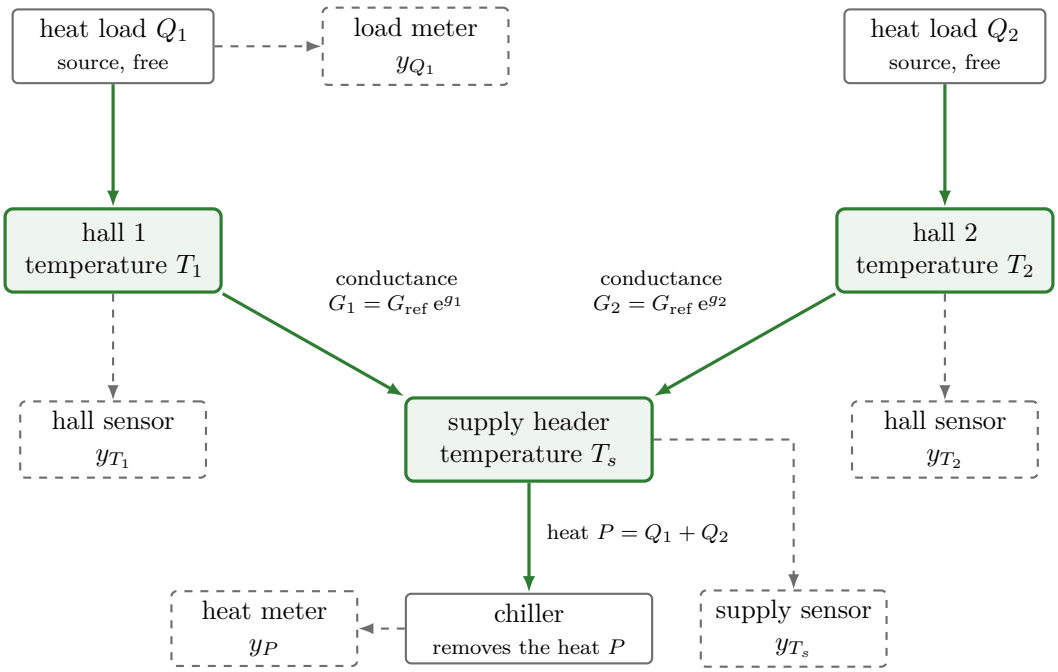


Figure 17.1: Plant graph of two data halls on one chilled-water loop in steady state. Free: the log-conductances g_1, g_2 of the two heat-removal paths, the heat loads Q_1, Q_2, the supply temperature T_s, the hall temperatures T_1, T_2 and the chiller's heat P, the eight entries of the full state; the three balances are physics rows among them. Pinned: the reference conductance $G_{\text{ref}} = 1\,\text{kW/K}$. Observed (dashed): T_1, T_2, T_s, P and Q_1. The three priors, on g_1, g_2 and T_s, are rows of the evidence and are not drawn.

The conductance is carried as the log-conductance of Section 1.1.4, $g_i = \ln(G_i/G_{\text{ref}})$, with $G_{\text{ref}} = 1\,\text{kW/K}$, so that $G_i = G_{\text{ref}}\, e^{g_i}$ is positive for every real g_i and a change of g_i is a relative change of G_i. With the reference at one, e^{g_i} is the conductance in kW/K and e^{-g_i} its reciprocal in K/kW; where a numerical expression below writes $Q_i\, e^{-g_i}$ without the reference, $G_{\text{ref}} = 1\,\text{kW/K}$ has been absorbed and the expression is in kelvin. Every free quantity of the graph is an entry of the full state, as in Chapter 4:

$$\mathbf{z}_{\text{full}} = \begin{bmatrix} g_1 & g_2 & Q_1 & Q_2 & T_s & T_1 & T_2 & P \end{bmatrix}^{\mathsf{T}}, \tag{17.6}$$

eight entries, with g_1, g_2 dimensionless, Q_1, Q_2, P in kW and T_s, T_1, T_2 in °C. The order puts the conductances, the loads and the supply temperature first and the three quantities on the left sides of equations (17.4) and (17.5) last. The balances enter the residual vector as physics rows with a tolerance (Section 17.1.7). When they hold exactly, the full state can move only in five directions, one per entry of

$$\mathbf{z} = \begin{bmatrix} g_1 & g_2 & Q_1 & Q_2 & T_s \end{bmatrix}^{\mathsf{T}}, \tag{17.7}$$

the reduced state. Section 17.2.1 derives it in matrix form, as the null directions of the physics rows, together with its rows and its Jacobian; the inference of the following chapters runs on it.

17.1.5 The readings and the priors

Table 17.1 lists the five readings: each row is one sensor or meter, with its symbol, the value it reports, the standard deviation of its error and the quantity of the plant it measures. The hall sensors and the supply sensor read temperatures; the chiller's heat meter reads the heat it

removes from the loop, which by equation (17.5) is the sum of the two loads; the electrical meter on hall 1's equipment reads its electrical power, which in steady state is the heat it dissipates. The readings are illustrative: they were generated from $G_1 = 5.5\,\mathrm{kW/K}$, $G_2 = 4.0\,\mathrm{kW/K}$, $Q_1 = 30\,\mathrm{kW}$, $Q_2 = 24\,\mathrm{kW}$ and $T_s = 12°\mathrm{C}$ with errors of $+0.2\,\mathrm{K}$, $-0.3\,\mathrm{K}$, $+0.1\,\mathrm{K}$, $-1.5\,\mathrm{kW}$ and $+1.0\,\mathrm{kW}$ added to the five readings in the order of the table.

Table 17.1: The readings of the two-hall plant. The hall-1 reading is printed to six decimals, 17.654545°C, because it is generated as $12 + 30/5.5 + 0.2$; the numbers of this chapter are computed with that value.

Sensor or meter	Symbol	Reading	Error sd	Measures
Hall 1 temperature sensor	y_{T_1}	17.654545°C	$\sigma_T = 0.5\,\mathrm{K}$	T_1, whose balance is $T_s + Q_1/G_1$
Hall 2 temperature sensor	y_{T_2}	17.700°C	$\sigma_T = 0.5\,\mathrm{K}$	T_2, whose balance is $T_s + Q_2/G_2$
Supply temperature sensor	y_{T_s}	12.10°C	$\sigma_s = 0.3\,\mathrm{K}$	T_s
Chiller heat meter	y_P	52.5 kW	$\sigma_P = 2\,\mathrm{kW}$	P, whose balance is $Q_1 + Q_2$
Hall 1 electrical meter	y_{Q_1}	31.0 kW	$\sigma_Q = 1.5\,\mathrm{kW}$	Q_1

Table 17.2 lists the priors: each row is one entry of the state, its prior distribution, and where the prior comes from. The handbook gives a conductance of $5\,\mathrm{kW/K}$ for a hall of this size and coil type, with a spread of 35 %: in the log-conductance this is $g_i \sim \mathcal{N}(g^0, \sigma_g^2)$ with $g^0 = \ln 5 = 1.6094$ and $\sigma_g = 0.3$, since $\mathrm{e}^{0.3} = 1.35$. The supply temperature has the chiller's set point $T_s^0 = 12°\mathrm{C}$ as its prior mean, with a standard deviation $\sigma_0 = 1\,\mathrm{K}$ for the control's offset. The two loads have no prior: nothing is known about them before the readings beyond their being positive, and they enter with a flat prior.

Table 17.2: The priors of the two-hall plant. The loads have none.

Entry	Prior	Source
$g_1 = \ln(G_1/G_{\mathrm{ref}})$	$\mathcal{N}(\ln 5,\ 0.3^2)$	handbook conductance 5 kW/K, spread 35 %
$g_2 = \ln(G_2/G_{\mathrm{ref}})$	$\mathcal{N}(\ln 5,\ 0.3^2)$	the same handbook value; independent of g_1
Q_1, Q_2	none (flat)	determined by the readings alone
T_s	$\mathcal{N}(12°\mathrm{C},\ 1\,\mathrm{K}^2)$	chiller set point and control offset

17.1.6 The questions over the state (step 1)

Table 17.3 lists the four questions of the opening, asked of the formulation before any fit. The first column is the question in engineering language; the second the general form of its category, in the notation of Chapter 2; the third its mathematical form over the reduced state of equation (17.7); the fourth its category; and the last where it is answered. Each is answered from the rows and their Jacobian at the prior point $\mathbf{z}^{(0)}$ (Section 17.2.2).

17.1.7 The evidence as rows (step 2)

Each reading has the measurement model of Chapter 1: the reading equals the quantity it measures plus an error, normal with zero mean and the stated standard deviation, independent of the other readings' errors. For the hall-1 sensor, for example,

$$y_{T_1} = T_1 + \varepsilon_{T_1}, \qquad \varepsilon_{T_1} \sim \mathcal{N}(0, \sigma_T^2),$$

Table 17.3: The questions asked of the two-hall plant before the fit, the general form of their category, their mathematical form over the reduced state $\mathbf{z} = (g_1, g_2, Q_1, Q_2, T_s)$, their category and where each is answered. $\mathbf{y}$ is the five readings with the three priors; $\mathbf{z}^{(0)}$ the prior point; $\tilde{\mathbf{z}}$ the handbook values of g_1, g_2, T_s with the loads that minimize the objective there; $\mathbf{J}$ the Jacobian of the reduced rows and $\mathbf{J}_y\mathbf{N}$ its reading rows; $\mathbf{v}$ the right singular vector of the smallest singular value $s_{\min}$; ν the number of rows beyond entries.

Engineering question	General form	Mathematical form	Category	Answered in
Do the readings agree with the handbook values, the set point and each other, and can that be checked?	a tail, eq. (2.12), of the predictive of a reading, eq. (2.13)	$r_k(\tilde{\mathbf{z}})$, its share $\frac{1}{2}r_k^2/\mathcal{L}(\tilde{\mathbf{z}})$ and $2\Phi(-\lvert r_k(\tilde{\mathbf{z}})\rvert)$; $\nu =$ rows $-$ entries	predictive check of the evidence	17.3.1
Which entries does each reading see, and do the readings determine all five?	$\mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y})$, $f = \mathbf{v}^\mathsf{T}\mathbf{z}$	the pattern of $\mathbf{J}_{\text{full}}$; $\mathrm{rank}(\mathbf{J}_y\mathbf{N}) = 5$; $\mathrm{sd}(\mathbf{v}^\mathsf{T}\mathbf{z} \mid \mathbf{y}) = 1/s_{\min}$	observability (estimation along the weakest direction)	17.3.2
How well would the readings fix each entry?	$\mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y})$	$\sqrt{\Sigma_{ii}}$, $\mathbf{\Sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}$ at $\mathbf{z}^{(0)}$, against the prior spreads	estimation (the spread)	17.3.3
Where does the first step of the fit take the state?	the location, eq. (2.10), of the posterior with its rows linearized at a point	$\mathbf{z}^{(1)} = \mathbf{z}^{(0)} + \Delta\mathbf{z}$, $(\mathbf{J}^\mathsf{T}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^\mathsf{T}\mathbf{r}$; $\mathcal{L}(\mathbf{z}^{(1)})$	estimation (the point)	17.3.3

with T_1 the sixth entry of the full state (17.6). The two meters measure loads in kW and the three sensors temperatures in °C, so the five errors have different units, and each row is divided by its own standard deviation. With the five readings and the three priors the evidence has eight rows; with the three physics rows of the balances the residual vector has eleven rows over the eight entries of the full state. Which row acts on which entry is a property of the graph, and it fixes where the Jacobian of the rows can be nonzero, marked $*$:

$$\mathbf{J}_{\text{full}} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & * & 0 & 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 & 0 & 0 \\ * & 0 & * & 0 & * & * & 0 & 0 \\ 0 & * & 0 & * & * & 0 & * & 0 \\ 0 & 0 & * & * & 0 & 0 & 0 & * \end{bmatrix} \begin{matrix} \text{reading, hall 1} \\ \text{reading, hall 2} \\ \text{reading, supply} \\ \text{reading, heat meter} \\ \text{reading, load meter} \\ \text{prior, } g_1 \\ \text{prior, } g_2 \\ \text{prior, } T_s \\ \text{physics, hall 1} \\ \text{physics, hall 2} \\ \text{physics, header} \end{matrix}, \tag{17.8}$$

with the columns $g_1, g_2, Q_1, Q_2, T_s, T_1, T_2, P$ of equation (17.6). Every reading row and every prior row touches exactly one entry. The columns of g_1, g_2 and Q_2 contain no reading row at all: no sensor reads a conductance, and no meter reads the second load. Those three entries are reached only through the physics rows, the conductance g_i through the balance of its hall and Q_2 through the balance of hall 2 and the header balance. The supply temperature's column has four marks, its own sensor, its prior and both hall balances, and it is the entry that couples the two halls. The header row marks the two loads alike. All eleven rows are Gaussian: the reading and prior rows with the standard deviations of Tables 17.1 and 17.2, the physics rows

with a model tolerance σ_p, as written below.

The modeling choices. Seven modeling choices are made for the whole model, and each buys something and can fail. The reading errors are taken as normal, independent and unbiased, each with the standard deviation of Table 17.1: this is what makes each normalized row a number of standard deviations and lets the squares of the rows add, and it fails for a miscalibrated meter, whose bias a single snapshot cannot separate from the state (Section 17.6.5). The handbook conductance and the chiller's set point are carried as prior rows, and the loads are given none: the priors are what supplies rows beyond entries, and a hall unlike the handbook's would turn its prior row from a check into a source of disagreement, which Section 17.4.4 measures. Every free quantity of the graph is carried as an entry, the hall temperatures and the chiller's heat included, and the balances as physics rows: this is what makes each reading row a comparison of one entry with its reading, so that the pattern of the Jacobian is the graph's own, and it carries three entries the readings could do without, which the projection of Section 17.2.1 removes, valid as far as the limit of Section 17.4.1 holds. The halls and the loop are taken as steady, with one lumped heat-removal path per hall: this is what gives each balance its three terms and fixes the null directions, and a storage term or a second path would add entries and change the columns (Section 17.6.5). The rows are linearized at the prior point, the handbook conductances, an even split of a design load and the set point: this gives a covariance and a step before any fit, and it misleads where the rows curve over the length of the step, which Section 17.4.3 measures. The exact balances are taken as the limit $\sigma_p \to 0$ and the computation runs on the reduced state: this removes the three carried entries from the normal equations without a substitution made by hand, and it holds only if the free-entry block of the full covariance converges to the reduced one, which Table 17.7 checks. The conductance is carried as its logarithm: this keeps every conductance positive and puts the handbook's relative spread in a normal prior row, and on the conductance itself the Jacobian's entries would change as $1/G_i^2$ across the range of conductances (Section 17.1.7).

The rows of the readings, the priors and the balances. The measurement model of the hall-1 sensor states that $y_{T_1} - T_1$ is a normal error of standard deviation σ_T. The hall temperature is the sixth entry of the full state, so the density of the reading given the state,

$$p(y_{T_1} \mid \mathbf{z}) = \frac{1}{\sqrt{2\pi}\,\sigma_T} \exp\left[-\frac{(T_1 - y_{T_1})^2}{2\sigma_T^2}\right],$$

is a function of one entry. The balance of hall 1, equation (17.4), is taken to hold up to a model tolerance σ_p, the device of Chapter 4 (equation (4.16)): the hall temperature is normal about the temperature the balance predicts,

$$p(T_1 \mid g_1, Q_1, T_s) = \mathcal{N}\big(T_1;\; T_s + Q_1\,\mathrm{e}^{-g_1}/G_{\mathrm{ref}},\; \sigma_p^2\big),$$

with σ_p in kelvin, the amount by which the steady balance may fail. Each density is the exponential of minus half the square of a normalized row, and the two rows are

$$r_1(\mathbf{z}) = \frac{T_1 - y_{T_1}}{\sigma_T}, \qquad r_9(\mathbf{z}) = \frac{T_1 - T_s - Q_1\,\mathrm{e}^{-g_1}/G_{\mathrm{ref}}}{\sigma_p}, \tag{17.9}$$

the reading row and the physics row of hall 1, numbered as in the residual vector below. The reading row compares the reading with an entry; the physics row compares that entry with the temperature the balance predicts from (g_1, Q_1, T_s). A state with a large conductance and a small load predicts a hall barely above the supply: if its T_1 stays at the reading, the physics row grows positive, and if T_1 follows the balance down, the reading row grows negative, so the

evidence argues against that state through one row or the other. The rows of hall 2 are the same with the subscript 2, rows 2 and 10.

The heat meter and the header balance pair the same way. The meter's measurement model is $y_P = P + \varepsilon_P$, and the header balance (17.5) holds up to a tolerance $\sigma_{p,P}$ in kW:

$$r_4(\mathbf{z}) = \frac{P - y_P}{\sigma_P}, \qquad r_{11}(\mathbf{z}) = \frac{P - Q_1 - Q_2}{\sigma_{p,P}}.$$

The supply sensor and the load meter read T_s and Q_1, and their rows are $(T_s - y_{T_s})/\sigma_s$ and $(Q_1 - y_{Q_1})/\sigma_Q$. The tolerances are set at $\sigma_p = 0.01\,\mathrm{K}$ and $\sigma_{p,P} = 0.01\,\mathrm{kW}$: tight enough that the answers no longer move as they shrink further (Table 17.7), and within about two decades of the readings' spreads, which keeps the normal equations conditioned well enough for double precision.

Common mistake

Writing rows on incomparable scales. A balance row kept in watts with a standard deviation of a few milliwatts beside meter rows of a few kilowatts makes that row of the Jacobian many decades larger than the others, since a row's derivatives are divided by its standard deviation; the information matrix is then ill-conditioned and an iteration that should converge in a few steps crawls or stalls. Normalizing every row by its standard deviation, and giving a physics row a tolerance that is small against the physics yet within about two decades of the meters' spreads, as $\sigma_p = 0.01\,\mathrm{K}$ and $\sigma_{p,P} = 0.01\,\mathrm{kW}$ are here, keeps the rows on one scale.

Every quantity as an entry. Carrying the hall temperatures and the chiller's heat as entries makes every reading row a comparison of one entry with its reading and puts all of the physics into three rows of its own. The reading rows are then linear in the state, and the whole nonlinearity of the problem sits in the two hall physics rows, which hold the products $Q_i\,\mathrm{e}^{-g_i}$. The same choice turns three extensions into a change of one row. A storage term in the balance, the transient of Section 1.6.5, makes the hall temperature at one time depend on its value at the previous time, and the physics row becomes the discretized balance between two consecutive entries. A model discrepancy (Section 1.1.4) is the tolerance σ_p itself, set to a physical value instead of taken to its limit. And several rows that refer to the same temperature, a second sensor in the same hall or a question about the temperature itself, each touch the one entry, whose mean and variance are then read off the fit as the full-load temperature was in Section 4.3.6.

Common mistake

Substituting a reading for the state. Setting $T_s = y_{T_s} = 12.10\,^{\circ}\mathrm{C}$ and fitting the other four entries treats the supply sensor as exact. The supply temperature then cannot move toward what the two hall readings and the set point say, the hall rows absorb the sensor's error in the conductances, and the fit reports spreads that ignore the supply sensor's 0.3 K. A reading is a row with an error; the quantity it measures stays a free entry of the state.

Three kinds of row and their normalization. Chapter 1 distinguished three kinds of row (Section 1.1.7): physical rows, which state relations that must hold; reading rows, which compare the state with a measured value; and prior rows, which carry knowledge held before the readings. The two-hall plant has all three. Its physical rows are the two hall balances and the header balance, each divided by its tolerance. Its reading rows are the five rows of Table 17.1, each the comparison of one entry with its reading, as in equation (17.9). Its prior rows are the three rows of Table 17.2. A prior row has the form of a reading row, the entry minus its prior mean divided by the prior's standard deviation, because a normal prior $z_i \sim \mathcal{N}(z_i^0, \sigma^2)$ has the

density $\exp[-(z_i - z_i^0)^2/(2\sigma^2)]$ up to a constant: a handbook value is a reading of the entry taken before the test.

Every row is written normalized, divided by the standard deviation of its error: a reading or prior row by the spread of the sensor or the handbook, a physics row by its tolerance. The normalized row is dimensionless, and its weight is one: the term it puts into the objective, $\frac{1}{2}r_k^2$, equals the term $\frac{1}{2}w_k d_k^2$ of the unnormalized difference d_k, in the units of the reading, with weight $w_k = 1/\sigma_k^2$. Normalization therefore changes nothing in the objective. It changes what a row's value means: a normalized residual of 2 says that the state and the reading disagree by twice what the sensor's error allows, whatever the sensor measures, so that a temperature row in kelvin and a meter row in kilowatts can be compared on one scale. With the normalized rows stacked in the residual vector $\mathbf{r}$ and their Jacobian $\mathbf{J}$, the information matrix is $\mathbf{J}^\mathsf{T}\mathbf{J}$; written with the unnormalized differences, whose Jacobian is $\tilde{\mathbf{J}}$, and the weight matrix $\mathbf{W} = \mathrm{diag}(w_k)$, it is the same matrix $\tilde{\mathbf{J}}^\mathsf{T}\mathbf{W}\tilde{\mathbf{J}}$, equation (4.26).

The residual vector. Collecting the eleven rows,

$$\mathbf{r}_{\text{full}}(\mathbf{z}_{\text{full}}) = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \\ r_6 \\ r_7 \\ r_8 \\ r_9 \\ r_{10} \\ r_{11} \end{bmatrix} = \begin{bmatrix} (T_1 - y_{T_1})/\sigma_T \\ (T_2 - y_{T_2})/\sigma_T \\ (T_s - y_{T_s})/\sigma_s \\ (P - y_P)/\sigma_P \\ (Q_1 - y_{Q_1})/\sigma_Q \\ (g_1 - g^0)/\sigma_g \\ (g_2 - g^0)/\sigma_g \\ (T_s - T_s^0)/\sigma_0 \\ (T_1 - T_s - Q_1\,\mathrm{e}^{-g_1}/G_{\text{ref}})/\sigma_p \\ (T_2 - T_s - Q_2\,\mathrm{e}^{-g_2}/G_{\text{ref}})/\sigma_p \\ (P - Q_1 - Q_2)/\sigma_{p,P} \end{bmatrix} \begin{matrix} \text{reading, hall 1} \\ \text{reading, hall 2} \\ \text{reading, supply} \\ \text{reading, heat meter} \\ \text{reading, load meter} \\ \text{prior, } g_1 \\ \text{prior, } g_2 \\ \text{prior, } T_s \\ \text{physics, hall 1} \\ \text{physics, hall 2} \\ \text{physics, header} \end{matrix} \tag{17.10}$$

a vector of eleven dimensionless numbers for each candidate full state, in three blocks: the five reading rows $\mathbf{r}_\mathrm{y}$, the three prior rows and the three physics rows $\mathbf{r}_\mathrm{p}$. The reading and prior rows together are written $\mathbf{r}_\mathrm{e}$, the rows of the evidence, and their Jacobian $\mathbf{J}_\mathrm{e}$, whose first five rows, the reading rows, are $\mathbf{J}_\mathrm{y}$; the physics rows have the Jacobian $\mathbf{J}_\mathrm{p}$. Rows 1 to 8 and 11 are linear in the full state; rows 9 and 10 are not, through the products $Q_i\,\mathrm{e}^{-g_i}$, and those two physics rows are the whole nonlinearity of the problem.

The objective. For rows with independent normal errors the minus logarithm of the posterior is, up to a constant, half the sum of the squared normalized rows: Chapter 4 derived this for the three-entry state of one hall with its physics rows, equation (4.24), and Chapter 18 derives the multi-row form for this plant. The objective of the full formulation is therefore

$$\mathcal{L}_{\text{full}}(\mathbf{z}_{\text{full}}) = \tfrac{1}{2}\sum_{k=1}^{11} r_k^2 = \tfrac{1}{2}\,\mathbf{r}_{\text{full}}^\mathsf{T}\mathbf{r}_{\text{full}}, \qquad p(\mathbf{z}_{\text{full}} \mid \mathbf{y}) \propto \exp\bigl[-\mathcal{L}_{\text{full}}\bigr], \tag{17.11}$$

with unit weights because the rows are normalized; the flat priors of the loads contribute no row. It is a single number for each candidate state, the sum of the squared disagreements of that state with every reading, every prior and every balance, each in units of its own standard deviation.

The Jacobian and the null directions of the balances. The Jacobian of the full residual vector is the 11×8 matrix

$$\mathbf{J}_{\text{full}} = \frac{\partial \mathbf{r}_{\text{full}}}{\partial \mathbf{z}_{\text{full}}^\mathsf{T}}, \qquad J_{ki} = \frac{\partial r_k}{\partial z_i}, \tag{17.12}$$

with the rows k of equation (17.10) down and the entries i of equation (17.6) across; its nonzero cells are the marks of the pattern (17.8). The hall physics rows carry the physics and their derivatives carry it too. Row $8+i$, for $i = 1, 2$, is $r_{8+i} = (T_i - T_s - Q_i\,\mathrm{e}^{-g_i}/G_{\mathrm{ref}})/\sigma_p$, and the derivative of the exponential is $\partial\,\mathrm{e}^{-g_i}/\partial g_i = -\mathrm{e}^{-g_i}$, so

$$\frac{\partial r_{8+i}}{\partial g_i} = \frac{Q_i\,\mathrm{e}^{-g_i}}{G_{\mathrm{ref}}\,\sigma_p}, \qquad \frac{\partial r_{8+i}}{\partial Q_i} = -\frac{\mathrm{e}^{-g_i}}{G_{\mathrm{ref}}\,\sigma_p} = -\frac{1}{G_i\,\sigma_p}, \qquad \frac{\partial r_{8+i}}{\partial T_s} = -\frac{1}{\sigma_p}, \qquad \frac{\partial r_{8+i}}{\partial T_i} = \frac{1}{\sigma_p}, \tag{17.13}$$

and the derivatives of the row with respect to the other hall's entries and to P are zero. On the balance, $Q_i\,\mathrm{e}^{-g_i}/G_{\mathrm{ref}} = T_i - T_s$, the first entry is the temperature rise of the hall in units of the tolerance: a relative increase of 1 % in the conductance, $\delta g_i = 0.01$, lowers the temperature the balance predicts by 1 % of the rise. The second is the reciprocal conductance: one more kilowatt raises the predicted hall by $1/G_i$ kelvin. The third is the same for both halls: the whole hall moves with the water that cools it. The fourth says that the row compares the entry T_i with that prediction.

The derivatives of the other rows. The other nine rows are linear in the state and their derivatives are constants. Each reading row has a single entry, the reciprocal of its sensor's standard deviation in the column of the entry it reads: $\partial r_1/\partial T_1 = \partial r_2/\partial T_2 = 1/\sigma_T$, $\partial r_3/\partial T_s = 1/\sigma_s$, $\partial r_4/\partial P = 1/\sigma_P$ and $\partial r_5/\partial Q_1 = 1/\sigma_Q$. Each prior row has one entry too, the reciprocal of its prior's standard deviation: $1/\sigma_g$ for g_1 and g_2, and $1/\sigma_0$ for T_s. The header physics row has three, $\partial r_{11}/\partial P = 1/\sigma_{p,P}$ and $\partial r_{11}/\partial Q_1 = \partial r_{11}/\partial Q_2 = -1/\sigma_{p,P}$: the two loads enter the header balance with the same weight, so the balance, and the heat meter that sees the loads through it, cannot tell them apart. A Jacobian entry of a normalized row has the units of the reciprocal of its entry of the state: none for a log-conductance, kW^{-1} for a load or the chiller's heat, K^{-1} for a temperature.

The null directions of the physics rows. Split the full state into its first five entries, the reduced state $\mathbf{z} = (g_1, g_2, Q_1, Q_2, T_s)$ of equation (17.7), and its last three, $\mathbf{z}_{\mathrm{b}} = (T_1, T_2, P)$, and the physics block of the Jacobian into the matching columns, $\mathbf{J}_{\mathrm{p}} = [\,\mathbf{J}_{\mathrm{pa}} \;\; \mathbf{J}_{\mathrm{pb}}\,]$, a 3×5 and a 3×3 block. A move $\Delta\mathbf{z}_{\mathrm{full}}$ keeps the linearized balances when $\mathbf{J}_{\mathrm{p}}\Delta\mathbf{z}_{\mathrm{full}} = \mathbf{0}$, that is $\mathbf{J}_{\mathrm{pa}}\Delta\mathbf{z} + \mathbf{J}_{\mathrm{pb}}\Delta\mathbf{z}_{\mathrm{b}} = \mathbf{0}$. The block $\mathbf{J}_{\mathrm{pb}} = \mathrm{diag}(1/\sigma_p,\ 1/\sigma_p,\ 1/\sigma_{p,P})$ is invertible, so

$$\Delta\mathbf{z}_{\mathrm{b}} = \mathbf{N}_{\mathrm{b}}\,\Delta\mathbf{z}, \qquad \mathbf{N}_{\mathrm{b}} = -\mathbf{J}_{\mathrm{pb}}^{-1}\mathbf{J}_{\mathrm{pa}}, \qquad \mathbf{N} = \begin{bmatrix} \mathbf{I}_5 \\ \mathbf{N}_{\mathrm{b}} \end{bmatrix}, \qquad \mathbf{J}_{\mathrm{p}}\mathbf{N} = \mathbf{0}, \tag{17.14}$$

and the five columns of the 8×5 matrix $\mathbf{N}$ are the null directions of the physics rows, the tangent of the set on which the balances hold, one column per free direction: each moves one entry of the reduced state by one unit and the three last entries as the linearized balances require, by its lower block. The tolerances cancel in that block $\mathbf{N}_{\mathrm{b}}$, which by equation (17.13) and the header row is

$$\mathbf{N}_{\mathrm{b}}(\mathbf{z}) = \begin{bmatrix} -Q_1\mathrm{e}^{-g_1}/G_{\mathrm{ref}} & 0 & 1/G_1 & 0 & 1 \\ 0 & -Q_2\mathrm{e}^{-g_2}/G_{\mathrm{ref}} & 0 & 1/G_2 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix},$$

with columns $(g_1, g_2, Q_1, Q_2, T_s)$ and rows (T_1, T_2, P) in °C, °C and kW. Its columns read as the physics: on the balance the null direction of g_i lowers T_i by the hall's rise $T_i - T_s$ per unit of g_i; the direction of Q_i raises T_i by $1/G_i$ kelvin and P by one kilowatt per kilowatt; the direction of T_s raises both halls with the supply. The physics block has rank three, so these five directions are all of its null space.

The log-conductance as the entry. The conductance is carried as $g_i = \ln(G_i/G_{\mathrm{ref}})$ rather than as G_i, and the Jacobian shows three reasons. The first is positivity, as in Chapter 1: every real value of g_i is a positive conductance, so the fit can move the entry freely. The second is the prior. A handbook spread is a relative spread, 35 % of the value, and a normal prior on g_i is a prior of that kind, a factor $\mathrm{e}^{\pm 0.3}$ at one standard deviation; written on G_i the same knowledge is a log-normal density, which no row linear in G_i can carry. The third is the form of the Jacobian entries. In the log-conductance the hall physics row has the entry and the null direction has the component

$$\frac{\partial r_{8+i}}{\partial g_i} = \frac{T_i - T_s}{\sigma_p}, \qquad (\mathbf{N}_{\mathrm{b}})_{i,g_i} = -(T_i - T_s),$$

the temperature rise of the hall, which the hall reading sees as $-(T_i - T_s)/\sigma_T$: it has a physical meaning, it is the same for every conductance that produces the same rise, and it scales with the quantity the sensor sees. In the conductance itself they would be

$$\frac{\partial r_{8+i}}{\partial G_i} = \frac{Q_i}{G_i^2\,\sigma_p}, \qquad (\mathbf{N}_{\mathrm{b}})_{i,G_i} = -\frac{Q_i}{G_i^2},$$

an entry that depends on the units of G_i and that changes as $1/G_i^2$ across the range of conductances, so that the linearization would hold over a narrower range.

The exact balances as a limit. As $\sigma_p \to 0$ the physics rows become the exact balances. A state that misses a hall balance by δ kelvin pays $\frac{1}{2}(\delta/\sigma_p)^2$ in the objective, a term that grows without bound as the tolerance shrinks, so the posterior concentrates on the set of states where the three balances hold. On that set the full state can still move in five directions, the null directions of the physics rows derived in Section 17.1.7, and the readings and priors decide where on it the state lies; the reduced state of Section 17.2.1 is the set described by those five directions. The limit changes no count: the three physics rows bring three rows, and the three directions they remove are the three entries T_1, T_2, P they name, so the redundancy is eleven rows less eight entries, the same as eight readings and priors less five free directions, equation (17.27). The limit is taken at the end of a computation and checked numerically, never used to remove an entry by hand; Section 17.4.1 checks it on the covariance and on the first step of the fit. A tolerance of physical size, a balance that fails by a few tenths of a kelvin, is a different model, one that adds the discrepancy's variance to every quantity the balance determines (equation (4.34)).

17.2 The posterior of the state (step 3)

In the book's notation $p(\mathbf{z} \mid \mathbf{y})$ is the posterior density of the state and $f(\mathbf{z})$ a quantity derived from it. Four objects recur in the specialized forms. The *reduced state* $\mathbf{z}$ of equation (17.7) is the vector of the five free entries, on which the rows act through the full state, whose physics rows hold in the limit $\sigma_p \to 0$. The *MAP state* $\hat{\mathbf{z}}$ is the minimum of the objective that the eight rows of the evidence put on the reduced state; it is not reached here, and every object is evaluated at the prior point $\mathbf{z}^{(0)}$, the linearization point of the first step. The *information matrix* $\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$, with $\mathbf{J}$ the 8×5 Jacobian of the normalized rows on the reduced state, is the Gauss–Newton curvature of the objective, and $\mathbf{\Sigma} = \mathbf{\Lambda}^{-1}$ is the covariance the fit would report if the point were its minimum. No closed form exists, since two of the rows multiply a load by an exponential of a log-conductance.

All four questions are read from one posterior: the posterior of the state that the eight rows of the evidence put on the reduced state, its balances exact. This section derives it, sets out the linearized step that computes it, and evaluates both for the plant at the prior point; Section 17.3 reads every answer from it.

17.2.1 The reduced state, its rows and its covariance

With exact balances the full state lies on the set where the three physics rows vanish, and the columns of $\mathbf{N}$ are the tangent of that set. The reduced state $\mathbf{z}$ names a point of the set by its first five entries; the point itself is

$$\mathbf{z}_{\text{full}}(\mathbf{z}) = \big(\mathbf{z},\ T_s + Q_1\,\mathrm{e}^{-g_1}/G_{\text{ref}},\ T_s + Q_2\,\mathrm{e}^{-g_2}/G_{\text{ref}},\ Q_1 + Q_2\big),$$

the state of the set with those free entries, at which rows 9 to 11 are zero. Its derivative with respect to $\mathbf{z}$ is $\mathbf{N}$, since $\mathbf{N}_{\text{b}}$ of equation (17.14) is the derivative of the last three entries along the set. The rows of the reduced state are the eight rows of the evidence at that point, $\mathbf{r}(\mathbf{z}) = \mathbf{r}_{\text{e}}\big(\mathbf{z}_{\text{full}}(\mathbf{z})\big)$:

$$\mathbf{r}(\mathbf{z}) = \begin{bmatrix} (T_s + Q_1\,\mathrm{e}^{-g_1}/G_{\text{ref}} - y_{T_1})/\sigma_T \\ (T_s + Q_2\,\mathrm{e}^{-g_2}/G_{\text{ref}} - y_{T_2})/\sigma_T \\ (T_s - y_{T_s})/\sigma_s \\ (Q_1 + Q_2 - y_P)/\sigma_P \\ (Q_1 - y_{Q_1})/\sigma_Q \\ (g_1 - g^0)/\sigma_g \\ (g_2 - g^0)/\sigma_g \\ (T_s - T_s^0)/\sigma_0 \end{bmatrix}, \tag{17.15}$$

numbered as rows 1 to 8 of the full vector. Its first row is the hall-1 reading row on the reduced state,

$$r_1(\mathbf{z}) = \frac{T_s + Q_1\,\mathrm{e}^{-g_1}/G_{\text{ref}} - y_{T_1}}{\sigma_T}, \tag{17.16}$$

the difference, in standard deviations of the sensor, between the temperature at the point of the set and the temperature the sensor reports. The objective on the reduced state is

$$\mathcal{L}(\mathbf{z}) = \tfrac{1}{2}\,\mathbf{r}(\mathbf{z})^{\mathsf{T}}\mathbf{r}(\mathbf{z}), \qquad p(\mathbf{z} \mid \mathbf{y}) \propto \exp\big[-\mathcal{L}(\mathbf{z})\big], \tag{17.17}$$

the full objective (17.11) on the set, where its physics rows contribute nothing.

The reduced Jacobian. A move $\Delta\mathbf{z}$ of the reduced state moves the point of the set by $\mathbf{N}\,\Delta\mathbf{z}$, so the chain rule gives the Jacobian of the reduced rows as a projection of the rows of the evidence onto the null directions,

$$\mathbf{J}(\mathbf{z}) = \frac{\partial \mathbf{r}}{\partial \mathbf{z}^{\mathsf{T}}} = \mathbf{J}_{\text{e}}\,\mathbf{N}, \tag{17.18}$$

an 8×5 matrix, the reading and prior rows only. Each reading row of $\mathbf{J}_{\text{e}}$ has its single entry $1/\sigma_k$, so its row of $\mathbf{J}$ is $1/\sigma_k$ times the row of $\mathbf{N}$ of the entry it reads:

$$\mathbf{J}(\mathbf{z}) = \begin{bmatrix} -Q_1\mathrm{e}^{-g_1}/(G_{\text{ref}}\sigma_T) & 0 & 1/(G_1\sigma_T) & 0 & 1/\sigma_T \\ 0 & -Q_2\mathrm{e}^{-g_2}/(G_{\text{ref}}\sigma_T) & 0 & 1/(G_2\sigma_T) & 1/\sigma_T \\ 0 & 0 & 0 & 0 & 1/\sigma_s \\ 0 & 0 & 1/\sigma_P & 1/\sigma_P & 0 \\ 0 & 0 & 1/\sigma_Q & 0 & 0 \\ 1/\sigma_g & 0 & 0 & 0 & 0 \\ 0 & 1/\sigma_g & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/\sigma_0 \end{bmatrix}.$$

The hall-1 sensor, whose row in the full state touches only T_1, sees its own conductance and load and the shared supply temperature on the reduced state, three entries at once and no entry alone; the heat meter, whose row touches only P, sees Q_1 and Q_2 with the same weight. The first five rows are the design block $\mathbf{J}_{\text{y}}\mathbf{N}$ of equation (17.28). The two hall rows are the only ones that change with the state; they are the reduced state's nonlinear rows, the hall balances carried into the readings.

The reduced covariance as the limit. The covariance of the reduced state is $\mathbf{\Sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}$ with the reduced Jacobian (17.18), and it is the block of the full-state covariance $\mathbf{\Sigma}_{\text{full}} = (\mathbf{J}_{\text{full}}^\mathsf{T}\mathbf{J}_{\text{full}})^{-1}$ in the rows and columns of the free entries, in the limit of exact balances. The derivation changes coordinates. Write a move of the full state as $\Delta\mathbf{z}_{\text{full}} = \mathbf{N}\,\Delta\mathbf{u} + \mathbf{B}\,\Delta\mathbf{v}$, with $\mathbf{B} = [\,\mathbf{0}\;\;\mathbf{I}_3\,]^\mathsf{T}$ the 8×3 matrix that moves only T_1, T_2 and P. The square matrix $[\,\mathbf{N}\;\;\mathbf{B}\,]$ is invertible, and its first five rows are $[\,\mathbf{I}_5\;\;\mathbf{0}\,]$, so $\Delta\mathbf{u}$ is the move of the free entries. Since $\mathbf{J}_\text{p}\mathbf{N} = \mathbf{0}$ and $\mathbf{J}_\text{p}\mathbf{B} = \mathbf{J}_\text{pb}$, the Jacobian in the new coordinates and its information matrix are

$$\mathbf{J}_{\text{full}}\,[\,\mathbf{N}\;\;\mathbf{B}\,] = \begin{bmatrix} \mathbf{J}_\text{e}\mathbf{N} & \mathbf{J}_\text{e}\mathbf{B} \\ \mathbf{0} & \mathbf{J}_\text{pb} \end{bmatrix}, \qquad \begin{bmatrix} \mathbf{J}^\mathsf{T}\mathbf{J} & \mathbf{C} \\ \mathbf{C}^\mathsf{T} & \mathbf{D} + \mathbf{J}_\text{pb}^\mathsf{T}\mathbf{J}_\text{pb} \end{bmatrix},$$

with $\mathbf{C} = \mathbf{J}^\mathsf{T}\mathbf{J}_\text{e}\mathbf{B}$ and $\mathbf{D} = \mathbf{B}^\mathsf{T}\mathbf{J}_\text{e}^\mathsf{T}\mathbf{J}_\text{e}\mathbf{B}$. The covariance of $\mathbf{u}$, which is the free-entry block of $\mathbf{\Sigma}_{\text{full}}$, is the inverse of the Schur complement of the lower block, by the block-inversion formula of Section 3.5.3, and the lower block grows without bound as the tolerances shrink, $\mathbf{J}_\text{pb}^\mathsf{T}\mathbf{J}_\text{pb} = \operatorname{diag}(\sigma_p^{-2},\ \sigma_p^{-2},\ \sigma_{p,P}^{-2})$:

$$(\mathbf{\Sigma}_{\text{full}})_{\mathbf{zz}} = \Big[\mathbf{J}^\mathsf{T}\mathbf{J} - \mathbf{C}\big(\mathbf{D} + \mathbf{J}_\text{pb}^\mathsf{T}\mathbf{J}_\text{pb}\big)^{-1}\mathbf{C}^\mathsf{T}\Big]^{-1} \;\longrightarrow\; (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1} = \mathbf{\Sigma} \qquad (\sigma_p,\ \sigma_{p,P} \to 0). \quad (17.19)$$

The correction inside the brackets is of order σ_p^2. The same coordinates give the covariance of the carried entries in the limit, $\mathbf{N}_\text{b}\mathbf{\Sigma}\mathbf{N}_\text{b}^\mathsf{T}$: the hall temperatures and the chiller's heat are known through the free entries, each by its row of $\mathbf{N}_\text{b}$. The equality holds at every point of the set, and Chapters 18 to 24 therefore run on the reduced state: its eight rows (17.15), its Jacobian $\mathbf{J} = \mathbf{J}_\text{e}\mathbf{N}$ and its covariance $(\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}$. The exact balance is this limit, never a substitution made by hand.

> **Principle 17.2.** *The design columns of a plant are fixed by its graph and its sensors, not by the values the sensors report: the zeros of its Jacobian say which entries a row cannot touch at any state, the null directions of its physics rows say how the balances carry a reading to the entries it does not read, and the rank of the reduced Jacobian* $\mathbf{J}_\text{e}\mathbf{N}$ *on the reading rows says whether the readings, without the priors, determine the state at all.*

17.2.2 The linearized step

Every object of the formulation is evaluated at one point before the fit, the prior point: the state that puts every entry with a prior at its prior mean, the two loads, which have none, at a starting value of 28 kW each, the design load of a hall (illustrative), and the three remaining entries on their balances, so that the physics rows start at zero:

$$\begin{aligned} \mathbf{z}_{\text{full}}^{(0)} &= (\ln 5,\ \ln 5,\ 28\,\text{kW},\ 28\,\text{kW},\ 12^\circ\text{C},\ 17.6^\circ\text{C},\ 17.6^\circ\text{C},\ 56\,\text{kW}), \\ \mathbf{z}^{(0)} &= (\ln 5,\ \ln 5,\ 28\,\text{kW},\ 28\,\text{kW},\ 12^\circ\text{C}), \end{aligned} \quad (17.20)$$

the second line being the first five entries of the first, the prior point of the reduced state; the hall temperatures are $12 + 28/5$ and the chiller's heat the sum of the loads. The superscript in parentheses numbers the points of an iteration, $\mathbf{z}^{(0)}$ the start and $\mathbf{z}^{(1)}$ the point after one step; a superscript without parentheses, as in g^0 and T_s^0, marks a prior mean.

The gradient of the objective says how it would change for a small move of each entry. Differentiating equation (17.17) by the chain rule, with the reduced Jacobian (17.18),

$$\nabla\mathcal{L}(\mathbf{z}) = \sum_k r_k(\mathbf{z})\,\frac{\partial r_k}{\partial \mathbf{z}} = \mathbf{J}(\mathbf{z})^\mathsf{T}\,\mathbf{r}(\mathbf{z}) = \mathbf{N}^\mathsf{T}(\mathbf{J}_\text{e}^\mathsf{T}\mathbf{r}_\text{e}), \quad (17.21)$$

where the inner product is the gradient of the full objective in the eight entries of the full state, at a point where the physics rows vanish, and $\mathbf{N}^\mathsf{T}$ collects it along each null direction.

Minimizing the objective means solving $\nabla\mathcal{L} = \mathbf{J}^\mathsf{T}\mathbf{r} = 0$, five nonlinear equations in five unknowns. The linearized problem replaces the rows by their tangents at the current point. For a step $\Delta\mathbf{z}$ from $\mathbf{z}^{(0)}$, the first-order Taylor expansion of each row is

$$\mathbf{r}(\mathbf{z}^{(0)} + \Delta\mathbf{z}) \approx \mathbf{r}(\mathbf{z}^{(0)}) + \mathbf{J}(\mathbf{z}^{(0)})\,\Delta\mathbf{z}, \tag{17.22}$$

exact for rows 3 to 8, which are linear, and approximate for the two hall rows. Inserting the expansion in the objective gives the linearized objective

$$\mathcal{L}_{\text{lin}}(\Delta\mathbf{z}) = \tfrac{1}{2}\left\|\mathbf{r}^{(0)} + \mathbf{J}^{(0)}\Delta\mathbf{z}\right\|^2 = \tfrac{1}{2}\,\mathbf{r}^{(0)\mathsf{T}}\mathbf{r}^{(0)} + \Delta\mathbf{z}^\mathsf{T}\mathbf{J}^{(0)\mathsf{T}}\mathbf{r}^{(0)} + \tfrac{1}{2}\,\Delta\mathbf{z}^\mathsf{T}\mathbf{J}^{(0)\mathsf{T}}\mathbf{J}^{(0)}\,\Delta\mathbf{z},$$

with $\mathbf{r}^{(0)} = \mathbf{r}(\mathbf{z}^{(0)})$ and $\mathbf{J}^{(0)} = \mathbf{J}(\mathbf{z}^{(0)})$, a quadratic function of the step. Its gradient with respect to $\Delta\mathbf{z}$ is $\mathbf{J}^{(0)\mathsf{T}}\mathbf{r}^{(0)} + \mathbf{J}^{(0)\mathsf{T}}\mathbf{J}^{(0)}\,\Delta\mathbf{z}$, and setting it to zero gives the normal equations

$$\left(\mathbf{J}^{(0)\mathsf{T}}\mathbf{J}^{(0)}\right)\Delta\mathbf{z} = -\mathbf{J}^{(0)\mathsf{T}}\mathbf{r}^{(0)}, \qquad \Delta\mathbf{z} = -\left(\mathbf{J}^{(0)\mathsf{T}}\mathbf{J}^{(0)}\right)^{-1}\mathbf{J}^{(0)\mathsf{T}}\mathbf{r}^{(0)}, \tag{17.23}$$

whose solution exists and is unique when $\mathbf{J}^{(0)}$ has full column rank. The matrix $\mathbf{J}^\mathsf{T}\mathbf{J}$ is the information matrix $\mathbf{\Lambda}$ of equation (4.26) for normalized rows, the Gauss–Newton approximation to the curvature of the objective, and the right-hand side is minus the gradient (17.21). The step is the least-squares solution of $\mathbf{J}^{(0)}\Delta\mathbf{z} = -\mathbf{r}^{(0)}$, eight equations in five unknowns. With $\mathbf{J} = \mathbf{J}_\text{e}\mathbf{N}$ the information matrix is $\mathbf{N}^\mathsf{T}(\mathbf{J}_\text{e}^\mathsf{T}\mathbf{J}_\text{e})\mathbf{N}$, and because every row of the evidence touches one entry of the full state, the inner matrix is diagonal, the precision each reading or prior puts on its own entry:

$$\mathbf{J}_\text{e}^\mathsf{T}\mathbf{J}_\text{e} = \operatorname{diag}(\sigma_g^{-2},\ \sigma_g^{-2},\ \sigma_Q^{-2},\ 0,\ \sigma_s^{-2} + \sigma_0^{-2},\ \sigma_T^{-2},\ \sigma_T^{-2},\ \sigma_P^{-2}),$$

in the order of equation (17.6): the priors on g_1 and g_2, the load meter on Q_1, nothing on Q_2, the supply sensor and the prior on T_s, the hall sensors on T_1 and T_2, the heat meter on P. All of the coupling between entries comes from the null directions: each entry of $\mathbf{J}^\mathsf{T}\mathbf{J}$ is a product of two columns of $\mathbf{N}$ through this diagonal.

The second-order term. The tangent step leaves the hall balances unmet by a term of second order. Write $h_i(\mathbf{z}) = T_s + Q_i\,\mathrm{e}^{-g_i}/G_\text{ref}$ for the temperature the balance of hall i assigns, whose gradient is row i of $\mathbf{N}_\text{b}$. A step along the tangent moves T_i by $\nabla h_i^\mathsf{T}\Delta\mathbf{z}$, while the point of the set moves it by $h_i(\mathbf{z}^{(0)} + \Delta\mathbf{z}) - h_i(\mathbf{z}^{(0)})$, and the miss is the second-order term of h_i along the step,

$$\sigma_p\, r_{8+i}(\mathbf{z}_\text{full}^{(0)} + \mathbf{N}\Delta\mathbf{z}) \approx -\tfrac{1}{2}\,\Delta\mathbf{z}^\mathsf{T}\,\nabla^2 h_i\,\Delta\mathbf{z}, \qquad \nabla^2 h_i = \begin{bmatrix} Q_i\,\mathrm{e}^{-g_i} & -\mathrm{e}^{-g_i} \\ -\mathrm{e}^{-g_i} & 0 \end{bmatrix}, \tag{17.24}$$

with G_ref absorbed, written in the rows and columns of (g_i, Q_i), the only entries in which h_i is not linear. The linearization (17.22) assumes that this term is small over the length of the step. The same matrix divided by σ_T is the second derivative of the reduced hall row, $\partial^2 r_i/\partial\mathbf{z}\,\partial\mathbf{z}^\mathsf{T}$, so it is also the matrix that enters the curvature term $\sum_k r_k\,\partial^2 r_k/\partial\mathbf{z}\,\partial\mathbf{z}^\mathsf{T}$, which the Gauss–Newton information $\mathbf{J}^\mathsf{T}\mathbf{J}$ drops: the curvature of the balance set as the readings see it. In the full state the same term is carried by the physics rows, and it does not vanish in the limit $\sigma_p \to 0$, because r_{8+i}/σ_p tends to a finite value while $\sigma_p r_{8+i}$ tends to zero.

17.2.3 The posterior of the plant at the prior point

The plant's inputs are the five readings and their standard deviations of Table 17.1, the three priors of Table 17.2, the reference $G_\text{ref} = 1\,\text{kW/K}$, the tolerances $\sigma_p = 0.01\,\text{K}$ and $\sigma_{p,P} = 0.01\,\text{kW}$, and the prior point (17.20). The hall-1 reading enters as printed, 17.654545°C.

Table 17.4 lists the rows with their values at the prior point. Each row of the table is one row of the vector (17.10): its index, what it compares, the unnormalized difference with its unit, and the standard deviation that normalizes it. The last two columns are its values at $\mathbf{z}_{\text{full}}^{(0)}$, unnormalized and normalized.

Table 17.4: The full residual vector (17.10), row by row, with its values at the prior point $\mathbf{z}_{\text{full}}^{(0)}$ of equation (17.20), at which the hall temperatures and the chiller's heat sit on their balances. $G_{\text{ref}} = 1\,\text{kW/K}$; the tolerances σ_p and $\sigma_{p,P}$ of rows 9 to 11 are 0.01 K and 0.01 kW.

k	Row	Difference (unit)	sd	At $\mathbf{z}_{\text{full}}^{(0)}$	r_k
1	hall 1 reading	$T_1 - y_{T_1}$ (K)	0.5 K	−0.0545 K	−0.109
2	hall 2 reading	$T_2 - y_{T_2}$ (K)	0.5 K	−0.100 K	−0.200
3	supply reading	$T_s - y_{T_s}$ (K)	0.3 K	−0.100 K	−0.333
4	chiller heat meter	$P - y_P$ (kW)	2 kW	3.50 kW	1.750
5	hall-1 load meter	$Q_1 - y_{Q_1}$ (kW)	1.5 kW	−3.00 kW	−2.000
6	prior on g_1	$g_1 - \ln 5$ (–)	0.3	0	0
7	prior on g_2	$g_2 - \ln 5$ (–)	0.3	0	0
8	prior on T_s	$T_s - 12$ (K)	1 K	0	0
9	hall 1 balance	$T_1 - T_s - Q_1 \mathrm{e}^{-g_1}$ (K)	0.01 K	0	0
10	hall 2 balance	$T_2 - T_s - Q_2 \mathrm{e}^{-g_2}$ (K)	0.01 K	0	0
11	header balance	$P - Q_1 - Q_2$ (kW)	0.01 kW	0	0

One row of the table in full, the hall-1 reading at the prior point: the sixth entry is $T_1 = 17.600°\text{C}$, which sits on the balance because $17.6 - 12 - 28\,\mathrm{e}^{-\ln 5} = 0$, so that row 9 is zero; the sensor reports 17.654545°C, the difference is −0.054545 K, and divided by 0.5 K it is −0.109: the prior point puts the hall about a tenth of a standard deviation too cool.

At the prior point, $G_i = 5\,\text{kW/K}$, $Q_i = 28\,\text{kW}$ and $T_i - T_s = 5.6\,\text{K}$, and with $\sigma_p = 0.01\,\text{K}$, the four nonzero entries of each hall physics row, equation (17.13), are

$$\frac{5.6}{0.01} = 560, \qquad -\frac{1}{5 \times 0.01} = -20\,\text{kW}^{-1}, \qquad -\frac{1}{0.01} = -100\,\text{K}^{-1}, \qquad \frac{1}{0.01} = 100\,\text{K}^{-1}.$$

The reading rows have $1/\sigma_T = 2\,\text{K}^{-1}$ on T_1 and T_2, $1/\sigma_s = 3.333\,\text{K}^{-1}$ on T_s, $1/\sigma_P = 0.5\,\text{kW}^{-1}$ on P and $1/\sigma_Q = 0.667\,\text{kW}^{-1}$ on Q_1; the prior rows $1/\sigma_g = 3.333$ on g_1 and g_2 and $1/\sigma_0 = 1\,\text{K}^{-1}$ on T_s; the header row $\pm 100\,\text{kW}^{-1}$. Assembled, the Jacobian is

$$\mathbf{J}_{\text{full}}(\mathbf{z}_{\text{full}}^{(0)}) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 3.333 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0.667 & 0 & 0 & 0 & 0 & 0 \\ 3.333 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3.333 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 560 & 0 & -20 & 0 & -100 & 100 & 0 & 0 \\ 0 & 560 & 0 & -20 & -100 & 0 & 100 & 0 \\ 0 & 0 & -100 & -100 & 0 & 0 & 0 & 100 \end{bmatrix} = \begin{bmatrix} \mathbf{J}_{\text{e}} \\ \mathbf{J}_{\text{p}} \end{bmatrix}, \tag{17.25}$$

eleven rows by eight columns, in the order $(g_1, g_2, Q_1, Q_2, T_s, T_1, T_2, P)$ across, with the eight rows of the evidence above and the three physics rows below. Figure 17.2 draws the same matrix as a grid. Rows run down, one per row of the residual vector, and entries of the full state run across; a filled cell is a nonzero entry, with its value at the prior point, and an empty cell a structural zero, an entry the row does not contain at any state. The upper frame marks

the five reading rows and the lower frame the three physics rows, whose cells show σ_p times the entry so that the values do not depend on the tolerance. The filled cells are the marks of the pattern (17.8), and no others.

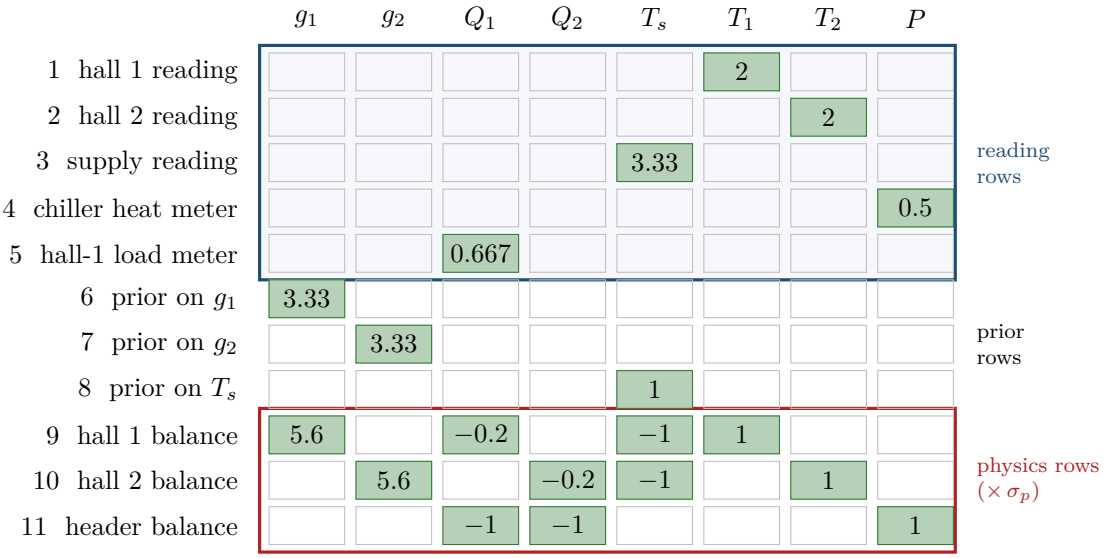


Figure 17.2: The pattern of the full Jacobian $\mathbf{J}_{\text{full}}(\mathbf{z}^{(0)}_{\text{full}})$ of equation (17.25). Rows: the eleven rows of the residual vector; columns: the eight entries of the full state. Green cells: nonzero entries with their values at the prior point, the physics rows 9 to 11 multiplied by their tolerance (σ_p in kelvin for the halls, $\sigma_{p,P}$ in kilowatts for the header); empty cells: structural zeros. Upper frame: the five reading rows; lower frame: the three physics rows.

The tolerances cancel in $\mathbf{N}_{\text{b}}$ of equation (17.14), which at the prior point is

$$\mathbf{N}_{\text{b}}(\mathbf{z}^{(0)}) = \begin{bmatrix} -5.6 & 0 & 0.2 & 0 & 1 \\ 0 & -5.6 & 0 & 0.2 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix} :$$

the null direction of g_1 lowers T_1 by 5.6 K per unit of g_1, the rise of the hall; the direction of Q_1 raises T_1 by 0.2 K and P by 1 kW per kilowatt; the direction of T_s raises both halls with the supply. Table 17.5 lists the reduced Jacobian $\mathbf{J}_{\text{e}}\mathbf{N}$ at the prior point. Each row of the table is a row of the evidence and each column a null direction, named by the entry of the reduced state it moves by one unit.

The precision each row of the evidence puts on its own entry is

$$\mathbf{J}_{\text{e}}^{\mathsf{T}}\mathbf{J}_{\text{e}} = \text{diag}(11.11,\ 11.11,\ 0.444,\ 0,\ 12.11,\ 4,\ 4,\ 0.25),$$

and through the null directions the information matrix at the prior point is

$$\mathbf{J}^{(0)\mathsf{T}}\mathbf{J}^{(0)} = \begin{bmatrix} 136.55 & 0 & -4.48 & 0 & -22.40 \\ 0 & 136.55 & 0 & -4.48 & -22.40 \\ -4.48 & 0 & 0.854 & 0.250 & 0.800 \\ 0 & -4.48 & 0.250 & 0.410 & 0.800 \\ -22.40 & -22.40 & 0.800 & 0.800 & 20.11 \end{bmatrix}.$$

Its first entry is $11.11 + 4 \times 5.6^2 = 136.55$, the prior on g_1 plus the hall-1 sensor's precision on T_1, which the null direction of g_1 moves by -5.6 K; the entry $-22.4 = 4 \times (-5.6) \times 1$ couples g_1 and T_s because both null directions move T_1, the entry the hall-1 sensor reads; and the Q_2 entry $0.410 = 4 \times 0.2^2 + 0.25$ comes entirely through T_2 and P, the zero of the diagonal on Q_2 itself.

Table 17.5: The reduced Jacobian $\mathbf{J}(\mathbf{z}^{(0)}) = \mathbf{J}_{\mathrm{e}}\mathbf{N}$ at the prior point, from equations (17.25), (17.14) and (17.18). Columns: the null directions of g_1, g_2, Q_1, Q_2 and T_s, in units of none, none, kW^{-1}, kW^{-1} and K^{-1}. Rows 1 to 5 are the reading rows; their columns are the design columns.

k	Row	g_1	g_2	Q_1	Q_2	T_s
1	hall 1 reading	−11.200	0	0.400	0	2.000
2	hall 2 reading	0	−11.200	0	0.400	2.000
3	supply reading	0	0	0	0	3.333
4	chiller heat meter	0	0	0.500	0.500	0
5	hall-1 load meter	0	0	0.667	0	0
6	prior on g_1	3.333	0	0	0	0
7	prior on g_2	0	3.333	0	0	0
8	prior on T_s	0	0	0	0	1.000

The gradient (17.21) of the full objective at the prior point, where the physics rows vanish, is

$$\mathbf{J}_{\mathrm{e}}^{\mathsf{T}}\mathbf{r}_{\mathrm{e}} = (0,\ 0,\ -1.333,\ 0,\ -1.111,\ -0.218,\ -0.400,\ 0.875),$$

one component per entry that a reading or a prior touches, in the units of the reciprocal entries. Multiplied by $\mathbf{N}^{\mathsf{T}}$, it gives the gradient on the reduced state,

$$\nabla\mathcal{L}(\mathbf{z}^{(0)}) = (1.222,\ 2.240,\ -0.502,\ 0.795,\ -1.729),$$

with norm 3.22. Each component collects the full gradient along one null direction: the direction of T_s moves T_s, T_1 and T_2 by one kelvin each, so its component is $-1.111 - 0.218 - 0.400 = -1.729\,\mathrm{K}^{-1}$, and it asks for a warmer supply because all three temperature rows predict too cool; the direction of g_1 moves T_1 by $-5.6\,\mathrm{K}$, so its component is $-5.6 \times (-0.218) = 1.222$, and it asks for a smaller conductance, which would warm hall 1.

The reduced rows, their Jacobian, the information matrix $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ with its inverse $\mathbf{\Sigma}$, and the gradient, all at the prior point, are what step 4 uses: every answer of the chapter is computed from them. Before the fit no exact posterior exists; the full state at a finite tolerance, equation (17.19), gives the value set beside the spreads and the step.

17.3 The answers (step 4)

The four questions of Table 17.3 fall into three categories: a predictive check of the evidence (the first), observability (the second) and estimation, of the spread and of the point (the third and fourth). Each is answered from the posterior of Section 17.2 at the prior point: its engineering question, the general form of its category, the form it takes for the plant, the quantities it uses, and the plant's numbers. Every answer is computed from the reduced state of step 3, its rows (17.15), its Jacobian and its covariance $\mathbf{\Sigma}$ of equation (17.19), and that is the value the chapter reports. Before the fit there is no exact posterior to set beside it; the full state at a finite tolerance, which reaches the same values to the order of σ_p^2, gives the value set beside the spreads and the step.

17.3.1 The disagreement of the readings and the redundancy

Engineering question. Do the readings agree with the handbook values, the set point and each other, and can that agreement be checked? The engineer wants to know which readings the handbook state contradicts, and whether the plant has readings to spare against which a faulty one would show.

General form. A reading is checked against the distribution the model gives it, the predictive (2.13) with y_{new} the reading, and its surprise is the tail of that predictive beyond the observed value, the form (2.12) with f the distance of the reading from the predictive's center,

$$\Pr(|y_{\text{new}} - \mathbb{E}[y_{\text{new}}]| \geq |y_k - \mathbb{E}[y_{\text{new}}]|).$$

Whether a check is possible at all is a count of rows against entries; at the fitted state the check becomes a tail of the minimized objective, which Chapter 18 derives.

Specialized form. The question is equations (2.13) and (2.12) with y_{new} each of the five readings y_k in turn, f its distance $|y_k - \mathbb{E}[y_{\text{new}}]|$ from the predictive's center, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of the reduced state of step 3 taken at one point $\tilde{\mathbf{z}}$, so that the predictive of y_k is its reading row at $\tilde{\mathbf{z}}$. One choice enters here. The predictive of each reading is taken at one state $\tilde{\mathbf{z}}$: the entries with a prior, the conductances and the supply temperature, at their prior means, and the loads, which have none, at the values that minimize the objective there. Profiling the loads out in this way makes the check one of the handbook knowledge rather than of an arbitrary starting value of the loads; it spends the meters on placing the loads, so their own residuals are smaller than an independent check of them would give, and taking the predictive at a point rather than averaged over the prior's spread overstates the surprise of a reading whose entry the prior leaves loose, which the fitted check of Chapter 18 does not.

Every row is normalized, $r_k = d_k/\sigma_k$ with d_k the difference in the units of the reading (Section 17.1.7). At $\tilde{\mathbf{z}}$ the predictive of reading k is normal about the value the state predicts, with the sensor's standard deviation, so $r_k(\tilde{\mathbf{z}})$ is the reading's distance from the predictive's center in standard deviations, whatever the row measures; its tail is $2\Phi(-|r_k|)$ and its share of the objective $\mathcal{L} = \frac{1}{2}\sum_k r_k^2$ is $\frac{1}{2}r_k^2/\mathcal{L}$. The physics rows are zero by construction and the priors' rows are zero at their means, so the whole objective at $\tilde{\mathbf{z}}$ is the readings' disagreement with the handbook knowledge. With g_1, g_2 and T_s fixed the reduced rows (Section 17.2.1) are linear in the loads, and the profiled loads are the least-squares solution in two entries, one step of the normal equations restricted to the two load columns $\mathbf{J}_Q$ of the reduced Jacobian from the prior point,

$$(\tilde{Q}_1,\ \tilde{Q}_2)^{\mathsf{T}} = (Q_1^{(0)},\ Q_2^{(0)})^{\mathsf{T}} - (\mathbf{J}_Q^{\mathsf{T}}\mathbf{J}_Q)^{-1}\mathbf{J}_Q^{\mathsf{T}}\,\mathbf{r}(\mathbf{z}^{(0)}), \tag{17.26}$$

exact because the rows are linear in the step.

Whether a check is possible is a count. The full state has eleven rows over eight entries, and the reduced state eight rows over five. The difference is the same,

$$\nu = \text{rows} - \text{entries} = 11 - 8 = 8 - 5 = 3, \tag{17.27}$$

because each physics row brings one row and removes one direction; it is the number of degrees of freedom of the fit, the rows of redundancy. Five rows, suitably chosen, are enough to fix the five entries; the other three can only agree or disagree with them, and the size of their disagreement at the fitted state is the consistency check of Chapter 18. At the minimum of the objective of a linear model with normal errors, twice the objective is a chi-square variable with ν degrees of freedom, so a value far in its upper tail says that the readings and the physics cannot all be true.

Without redundancy the check is empty. Remove the three prior rows and the five reading rows of the reduced state remain, five equations in five unknowns. If their Jacobian has full rank they have an exact solution, the state at which every row is zero, and in the full state the five reading rows put T_1, T_2, T_s, P and Q_1 at their readings while the three physics rows fix Q_2, g_1 and g_2, eight rows over eight entries with a square Jacobian. The objective is zero for any readings, and nothing in the fit would show a reading in error. Observability is a separate property: whether the readings determine the entries at all is a question of rank, the next question's, and a square system of full rank is observable and still uncheckable. The priors are

rows like the readings and they count. In this plant they are the redundancy: they compare the fitted conductances and supply temperature with handbook knowledge, so that a reading that drives a conductance far from its handbook value shows up as a large prior residual. A second sensor in one hall would add a fourth degree of freedom, one more row on the entry T_1, and the loss of the load meter would leave two.

Principle 17.3. *Redundancy is the number of rows beyond the number of entries of the state. It is what allows the readings to be checked against the physics and against each other; a square system fits any readings exactly and checks nothing.*

Building blocks. The eleven rows and their normalization (Section 17.1.7), the eight rows of the reduced state and their Jacobian (Section 17.2.1, with their values at the prior point in Section 17.2.3), the prior point (Section 17.2.2); the profiled loads (17.26) and the normal tail Φ; the count (17.27); the square system of the five readings.

The disagreement with the handbook and the redundancy. The objective (17.17) at the prior point, equal there to the full objective (17.11), is

$$\begin{aligned}\mathcal{L}(\mathbf{z}^{(0)}) &= \tfrac{1}{2}(0.1091^2 + 0.2000^2 + 0.3333^2 + 1.7500^2 + 2.0000^2 + 0 + 0 + 0)\\ &= \tfrac{1}{2}\,(0.0119 + 0.0400 + 0.1111 + 3.0625 + 4.0000) = 3.6128,\end{aligned}$$

98 % of it on the two meters. That share belongs to the starting loads of 28 kW each, which have no prior, and not to the handbook. The profiled loads (17.26), from the load columns of Table 17.5 and the residuals of Table 17.4, are

$$\tilde{Q}_1 = 29.406\,\text{kW}, \qquad \tilde{Q}_2 = 25.204\,\text{kW},$$

and at $\tilde{\mathbf{z}} = (\ln 5,\ \ln 5,\ 29.406\,\text{kW},\ 25.204\,\text{kW},\ 12°\text{C})$ the hall temperatures on their balances are 17.881 and 17.041°C and the chiller's heat 54.609 kW. The five reading rows there, their shares of $\mathcal{L}(\tilde{\mathbf{z}}) = 2.1485$ and their two-sided predictive tails are:

Reading	$r_k(\tilde{\mathbf{z}})$	share	$2\Phi(-\lvert r_k\rvert)$
hall 1 sensor	+0.453	4.8 %	0.65
hall 2 sensor	−1.318	40.5 %	0.19
supply sensor	−0.333	2.6 %	0.74
chiller heat meter	+1.055	25.9 %	0.29
hall-1 load meter	−1.063	26.3 %	0.29

Each line is one reading: its normalized row at $\tilde{\mathbf{z}}$, its share of the objective, and the probability of a reading at least as far from its predictive's center. The largest disagreement is the hall-2 sensor: it reads 17.700°C against the 17.041°C that the handbook conductance gives hall 2 with the load the meters leave it, a hall 0.66 K warm, which is the direction of a low hall-2 conductance. No reading is surprising alone; the smallest tail is 0.19.

The count (17.27) gives $\nu = 3$, and all three rows of redundancy are the priors. Without them the five reading rows of the reduced state have a design block of full rank (Section 17.3.2), and their exact solution, the state at which every row is zero, is

$$\mathbf{z} = (1.7194,\ 1.3453,\ 31.0\,\text{kW},\ 21.5\,\text{kW},\ 12.10°\text{C}), \qquad G_1 = 5.581\,\text{kW/K}, \quad G_2 = 3.839\,\text{kW/K},$$

the hand calculation of the opening. The square system reproduces any readings exactly: had the hall-2 sensor reported 22.0°C, a gross error of more than four degrees, the system would have returned $G_2 = 21.5/9.9 = 2.172\,\text{kW/K}$ with every residual zero.

17.3.2 The reach of each reading and the weakest direction

Engineering question. Which entries does each reading see, and do the readings determine all five free entries? The hand calculation of the opening used every reading to fix one quantity; the engineer asks which combinations of the conductances, the loads and the supply temperature each reading actually constrains, and whether any direction of the state is left to the priors alone.

General form. Whether the readings determine the state is estimation along the direction they see least: the posterior variance (2.10) of a linear combination $f(\mathbf{z}) = \mathbf{v}^{\mathsf{T}}\mathbf{z}$, with $\mathbf{v}$ the unit vector along which that variance is largest. Which entries each row touches is the pattern of the Jacobian, a property of the graph and of its sensors that holds under every posterior.

Specialized form. The question is equation (2.10) with the state $\mathbf{z}$ the five free entries $(g_1, g_2, Q_1, Q_2, T_s)$ of the reduced state, the derived quantity $f(\mathbf{z}) = \mathbf{v}^{\mathsf{T}}\mathbf{z}$ along the weakest direction, and $p(\mathbf{z} \mid \mathbf{y})$ the Gaussian of the rows taken at the prior point of step 3, $\mathcal{N}(\mathbf{z}^{(0)}, \boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma} = (\mathbf{N}^{\mathsf{T}}\mathbf{J}_{\mathrm{y}}^{\mathsf{T}}\mathbf{J}_{\mathrm{y}}\mathbf{N})^{-1}$ for the readings alone and $\boldsymbol{\Sigma}$ of equation (17.19) with the priors; the variance is then $\mathbf{v}^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{v}$. The pattern (17.8) says which entry each row touches. A reading sees the entries it does not touch only along the null directions $\mathbf{N}$ of the physics rows, the moves of the full state that keep the linearized balances (Section 17.1.7), and what it sees of the five free directions is its row of the reduced Jacobian $\mathbf{J} = \mathbf{J}_{\mathrm{e}}\mathbf{N}$, the rows of the evidence projected onto those directions (Section 17.2.1). The reading rows of that Jacobian, $\mathbf{J}_{\mathrm{y}}\mathbf{N}$ with $\mathbf{J}_{\mathrm{y}}$ the five reading rows of $\mathbf{J}_{\mathrm{e}}$, are the design columns read row by row: a zero says that the reading cannot see that direction at any state, and a column with no nonzero entry is a direction that only a prior reaches. The readings alone determine all five free entries when the design block has full column rank,

$$\operatorname{rank}(\mathbf{J}_{\mathrm{y}}\mathbf{N}) = 5, \tag{17.28}$$

and its smallest singular value $s_{\min}$ says how weakly the worst-seen direction is seen. With the singular value decomposition $\mathbf{J}_{\mathrm{y}}\mathbf{N} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathsf{T}}$, the covariance of the readings alone is $(\mathbf{N}^{\mathsf{T}}\mathbf{J}_{\mathrm{y}}^{\mathsf{T}}\mathbf{J}_{\mathrm{y}}\mathbf{N})^{-1} = \mathbf{V}\mathbf{S}^{-2}\mathbf{V}^{\mathsf{T}}$, so the right singular vector $\mathbf{v}$ of $s_{\min}$ is the direction of largest variance, and the quantity $f = \mathbf{v}^{\mathsf{T}}\mathbf{z}$, whose gradient is $\mathbf{v}$, has

$$\operatorname{sd}(\mathbf{v}^{\mathsf{T}}\mathbf{z} \mid \mathbf{y}) = \sqrt{\mathbf{v}^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{v}} = \frac{1}{s_{\min}}; \tag{17.29}$$

with $\boldsymbol{\Sigma}$ the covariance of the rows taken, here the readings alone; with the priors the same holds for the singular values of the whole reduced Jacobian $\mathbf{J}$ and $\boldsymbol{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$. The structural zeros hold at every state; the values in the nonzero cells depend on the point, so the rank and the singular values are read at the prior point here and at the fitted state in Chapter 19.

Building blocks. The pattern (17.8); the derivatives of the rows (Section 17.1.7); the null directions $\mathbf{N}$ and the reduced Jacobian (Section 17.2.1, with their values at the prior point in Section 17.2.3); the singular values and the weakest right singular vector of the design block and of $\mathbf{J}$.

The design block and its weakest direction. A line of Table 17.5 is a line of $\mathbf{J}_{\mathrm{e}}$ times $\mathbf{N}$: the hall-1 reading row is 2 times the first row of $\mathbf{N}_{\mathrm{b}}$, $(-11.2,\ 0,\ 0.4,\ 0,\ 2)$, and the heat meter sees Q_1 and Q_2 with the same weight 0.5. The load Q_1 is seen by three reading rows, among them the one direct reading of the plant, and Q_2 by two, both in combination with another entry; no column of the design block is zero. At the prior point the design block has rank five, the condition (17.28), with singular values 11.59, 11.21, 3.22, 0.898 and 0.371, and

the whole reduced Jacobian, priors included, 12.03, 11.69, 3.47, 0.905 and 0.388. The weakest right singular vector of the design block is

$$\mathbf{v} = (0.015,\ -0.033,\ 0.409,\ -0.912,\ 0.000),$$

in the order $(g_1, g_2, Q_1, Q_2, T_s)$, a combination of the loads that is mostly hall 2's, and by equation (17.29)

$$\mathrm{sd}(\mathbf{v}^{\mathsf{T}}\mathbf{z} \mid \mathbf{y}) = 1/0.371 = 2.70\,\mathrm{kW}\ \text{(readings alone)}, \qquad 1/0.388 = 2.58\,\mathrm{kW}\ \text{(with the priors)}:$$

the readings alone determine all five free entries, and the priors strengthen the weakest direction a little, mostly through the prior on g_2, which the hall-2 reading ties to Q_2.

17.3.3 The spread of the state and the first step of the fit

Engineering question. How well would the readings fix each entry? Where does the first step of the fit take the state? The hand calculation gave five numbers and no spread; the engineer asks for the spread each entry would carry, the hall-2 conductance first, and for the point toward which the evidence first moves the handbook state.

General form. Both are estimation, the posterior distribution of a quantity, equation (2.9), summarized by a location and a spread, equation (2.10). For an entry, $f(\mathbf{z}) = z_i$, the spread is the posterior variance, and the location of the posterior is its mode $\hat{\mathbf{z}}$,

$$\mathrm{Var}(z_i \mid \mathbf{y}) = \int \big(z_i - \mathbb{E}[z_i \mid \mathbf{y}]\big)^2\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \qquad \hat{\mathbf{z}} = \arg\max_{\mathbf{z}}\ p(\mathbf{z} \mid \mathbf{y}).$$

The first step of the fit is not that mode. It is the mode of the posterior whose rows are replaced by their tangents at the starting point, $\mathbf{z}^{(1)} = \arg\max_{\mathbf{z}} \exp[-\mathcal{L}_{\mathrm{lin}}]$, with $\mathcal{L}_{\mathrm{lin}}$ the objective of the linearized rows; it coincides with $\hat{\mathbf{z}}$ only when the rows are linear or the start is already the mode.

Specialized form. The two questions are equations (2.9) and (2.10) with the state $\mathbf{z}$ the five free entries $(g_1, g_2, Q_1, Q_2, T_s)$ of the reduced state, $f(\mathbf{z}) = z_i$ for each entry (and for a carried entry its row of $\mathbf{N}_{\mathrm{b}}$ times $\mathbf{z}$), and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3 in the Laplace form, with the covariance $\boldsymbol{\Sigma}$ of equation (17.19) taken at the prior point. On the reduced state the Laplace approximation of Chapter 4 replaces the posterior by $\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$, an entry read off with location $\hat{z}_i$ and spread $\sqrt{\Sigma_{ii}}$. Before the fit the same objects are taken at the prior point: the information matrix $\boldsymbol{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$ of the eight reduced rows at $\mathbf{z}^{(0)}$, and its inverse $\boldsymbol{\Sigma}$, whose diagonal gives the spread $\sqrt{\Sigma_{ii}}$ of each free entry, set beside its prior's, σ_g for g_i, σ_0 for T_s and none for the loads. The correlation of the two loads, $\mathrm{corr}(Q_1, Q_2) = \Sigma_{34}/\sqrt{\Sigma_{33}\Sigma_{44}}$, says how the heat meter's sum ties them, and the carried entries, the hall temperatures and the chiller's heat, have the covariance $\mathbf{N}_{\mathrm{b}}\boldsymbol{\Sigma}\mathbf{N}_{\mathrm{b}}^{\mathsf{T}}$, each through its row of the matrix $\mathbf{N}_{\mathrm{b}}$ that the balances define (Section 17.2.1). For the location, $\hat{\mathbf{z}}$ is the minimum of the objective, and the first Gauss–Newton step $\Delta\mathbf{z}$ toward it solves the normal equations with the same matrix $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ and the gradient $\mathbf{J}^{\mathsf{T}}\mathbf{r}$ at $\mathbf{z}^{(0)}$ (Section 17.2.2). The point after the step, $\mathbf{z}^{(1)} = \mathbf{z}^{(0)} + \Delta\mathbf{z}$, is judged by its objective, evaluated with the nonlinear rows, against the value the tangent rows predict: the gap between the two is what the linearization left out. The value beside each comes from the full state at a finite tolerance: the free-entry block of $\boldsymbol{\Sigma}_{\mathrm{full}}$, equation (17.19), and the normal equations with the full Jacobian (17.25), which give the same spreads and the same step to the order of σ_p^2.

Building blocks. The reduced rows and Jacobian at $\mathbf{z}^{(0)}$, the matrix $\mathbf{N}_\mathrm{b}$, the information matrix and its inverse (Section 17.2.1); the gradient, the normal equations and the linearized objective (Section 17.2.2), with their values at the prior point in Section 17.2.3; and, for the values beside them, the full Jacobian (17.25) at $\sigma_p = 0.01\,\mathrm{K}$.

The spread of each entry at the prior point. The inverse of the information matrix of Section 17.2.3 gives the spreads of the five free entries at the prior point:

$$\sqrt{\Sigma_{11}} = 0.1083, \qquad \sqrt{\Sigma_{22}} = 0.1241,$$
$$\sqrt{\Sigma_{33}} = 1.459\,\mathrm{kW}, \qquad \sqrt{\Sigma_{44}} = 2.394\,\mathrm{kW}, \qquad \sqrt{\Sigma_{55}} = 0.2801\,\mathrm{K},$$

in the order $(g_1, g_2, Q_1, Q_2, T_s)$. The log-conductances are known to factors of $\mathrm{e}^{0.108}$ and $\mathrm{e}^{0.124}$, 11 and 13 % of the conductance, against the handbook's 35 %; the supply temperature to 0.28 K, against the set point's 1 K and the sensor's 0.3 K; hall 1's load to 1.46 kW, against its meter's 1.5 kW; and hall 2's load, which no meter reads, to 2.39 kW, two thirds again as wide as hall 1's. The correlation of the two loads is $\mathrm{corr}(Q_1, Q_2) = -0.5793$: through the heat meter's sum, a kilowatt more in one hall is a kilowatt less in the other. The hall-1 temperature, a carried entry, has $\sqrt{(\mathbf{N}_\mathrm{b}\boldsymbol{\Sigma}\mathbf{N}_\mathrm{b}^\mathsf{T})_{11}} = 0.4804\,\mathrm{K}$, slightly below its sensor's 0.5 K. The values beside them come from the full state at $\sigma_p = 0.01\,\mathrm{K}$: its free-entry block gives the same five spreads, the same correlation and the same 0.4804 K to every printed digit, and differs from $\boldsymbol{\Sigma}$ by at most 1.0×10^{-4} in any element.

The first step of the fit. Solving the normal equations (17.23) at the prior point gives the step of Table 17.6. Each row of the table is one entry of the state, with its value at $\mathbf{z}^{(0)}$, the step, and its value after the step, $\mathbf{z}^{(1)} = \mathbf{z}^{(0)} + \Delta\mathbf{z}$; the lower rows give the derived quantities and the objective at both points, and the value of the linearized objective at the step.

Table 17.6: One linearized step from the prior point, by the normal equations (17.23). The objective at $\mathbf{z}^{(1)}$ is evaluated with the nonlinear rows (17.15); the linearized objective is the prediction of the tangent rows (17.22). T_1, T_2 and P are the last three entries of the point $\mathbf{z}_\mathrm{full}(\mathbf{z})$ of the balance set.

Quantity	At $\mathbf{z}^{(0)}$	Step $\Delta\mathbf{z}$	At $\mathbf{z}^{(1)}$
g_1	1.6094	0.0985	1.7080
g_2	1.6094	−0.1934	1.4160
Q_1 (kW)	28.000	2.739	30.739
Q_2 (kW)	28.000	−5.932	22.068
T_s (°C)	12.000	0.107	12.107
$G_1 = \mathrm{e}^{g_1}$ (kW/K)	5.000		5.518
$G_2 = \mathrm{e}^{g_2}$ (kW/K)	5.000		4.121
T_1, T_2 (°C)	17.600, 17.600		17.678, 17.463
P (kW)	56.000		52.807
Objective $\mathcal{L}$	3.6128		0.4085
Linearized objective $\mathcal{L}_\mathrm{lin}$			0.3179
Gradient norm $\lVert\mathbf{J}^\mathsf{T}\mathbf{r}\rVert$	3.223		3.065

The step does what the residuals of the prior point asked for. It moves the split of the load, about 2.7 kW into hall 1 and 5.9 kW out of hall 2, so that the total falls from 56 to 52.8 kW, close to the heat meter, and hall 1 rises toward its meter. It then adjusts the conductances so that each hall keeps its observed rise: hall 1 has more load, so its conductance rises by 10 % (g_1 by 0.0985); hall 2 has less, so its conductance falls by 18 % (g_2 by −0.193). The supply

temperature moves 0.107 K toward its sensor's 12.10°C, a tenth of its prior's standard deviation of 1 K and a third of the sensor's 0.3 K. The objective falls from 3.6128 to 0.4085, against the tangent's prediction of 0.318. The value beside the step comes from the full state: the normal equations with the full Jacobian (17.25) at $\sigma_p = 0.01$ K give a step whose first five entries agree with $\Delta\mathbf{z}$ to 1×10^{-5}.

Figure 17.3 draws the eight normalized residuals at both points. The horizontal axis lists the rows of the residual vector, the vertical axis is the residual in standard deviations, the dark bars are the prior point and the light bars the point after the step, and the dashed lines mark one standard deviation either way. At the prior point two bars stand out, the two meters near ±2, set there by the even split of the starting loads; after the step every row lies within one standard deviation, and the largest residuals are now on the hall-2 reading, −0.47, and on the prior of g_2, −0.64: the step has moved the disagreement from the meters onto the conductance of hall 2, whose estimate now lies 0.64 prior standard deviations below the handbook value. That prior row is the redundancy at work: the low conductance the readings ask for shows as a prior residual, where the square system would have shown nothing.

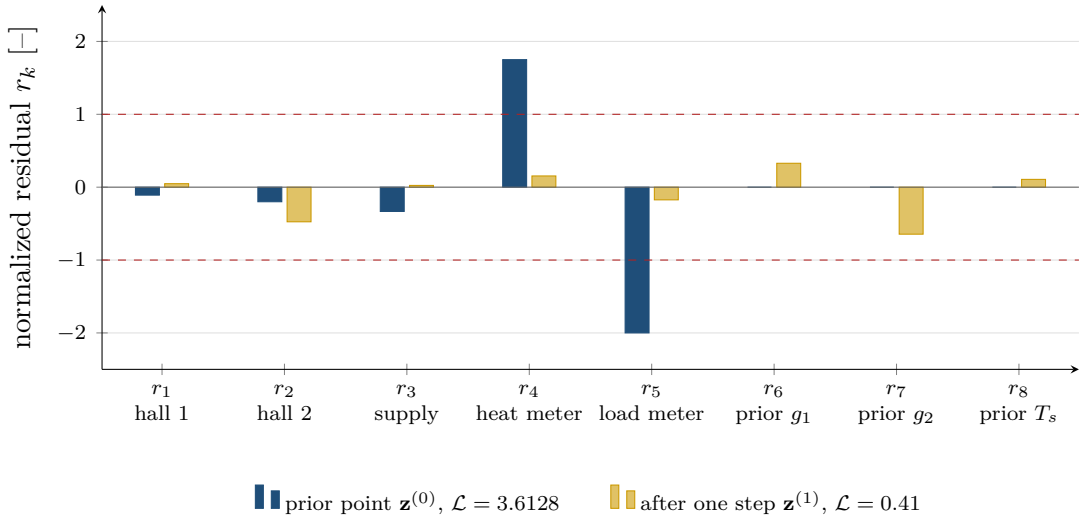


Figure 17.3: The eight normalized residuals of equation (17.15) at the prior point $\mathbf{z}^{(0)}$ (dark) and after one linearized step, $\mathbf{z}^{(1)}$ (light), from Tables 17.4 and 17.6. Dashed lines: ±1 standard deviation.

Figure 17.4 shows the step on the objective itself, in two slices through the five-dimensional state. In each panel the horizontal axis is the log-conductance g_1 and the vertical axis the load Q_1; the closed curves are contours of the objective (17.17) with the other three entries held fixed, at the levels written beside them. In panel (a) the other entries are held at the prior point, $g_2 = \ln 5$, $Q_2 = 28$ kW, $T_s = 12$°C; in panel (b) at their values after the step. The dot is (g_1, Q_1) at $\mathbf{z}^{(0)}$, the square at $\mathbf{z}^{(1)}$, and the arrow is the step's component in this plane. The contours are tilted ellipses: a higher load in hall 1 is consistent with its reading only with a higher conductance, so the low-objective region runs from lower left to upper right, along the curve of constant temperature rise. In panel (a) the step lands on a higher contour, near 5.0, than the one it left: with Q_2 held at 28 kW the heat meter forbids the extra load in hall 1, and the slice's own minimum, 3.46, lies close to the prior point. In panel (b) the same point lies within a hundredth of the slice's minimum, 0.407 at $g_1 = 1.704$, $Q_1 = 30.75$ kW. The step is a move in all five entries at once, and a slice through the objective shows its descent only when the slice is taken where the other entries have moved too.

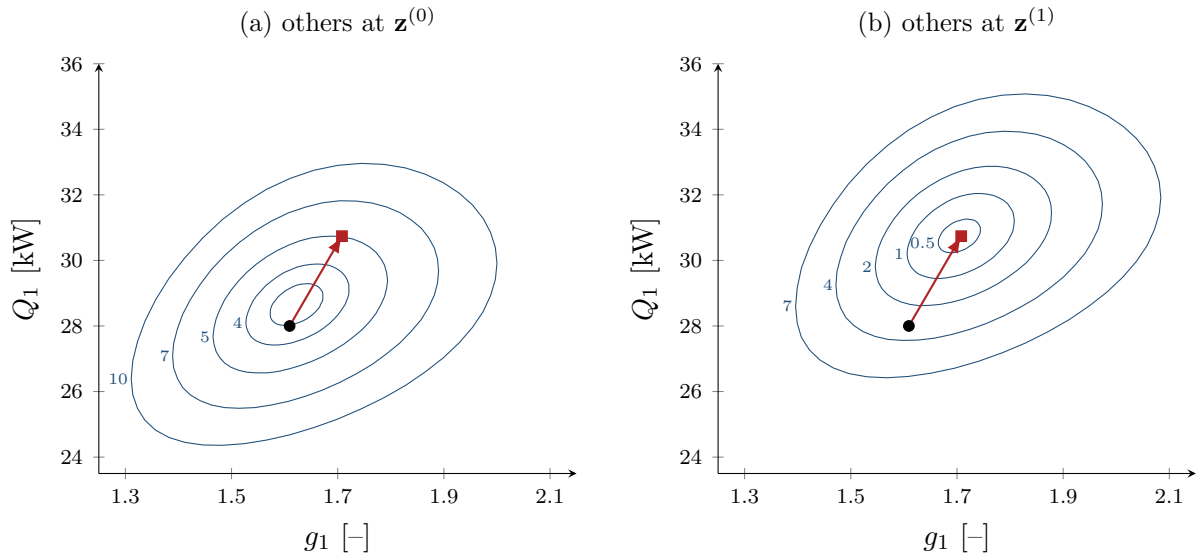


Figure 17.4: Contours of the objective (17.17) over (g_1, Q_1) with (g_2, Q_2, T_s) held at the prior point (a), $(\ln 5,\ 28\,\text{kW},\ 12°\text{C})$, and after the step (b), $(1.416,\ 22.07\,\text{kW},\ 12.107°\text{C})$; levels as labeled, with the unlabeled innermost contour of panel (a) at 3.6. Dot: (g_1, Q_1) at $\mathbf{z}^{(0)}$; square: at $\mathbf{z}^{(1)}$; arrow: the step's component in the plane. Slice minima: (a) 3.46 at $(1.622,\ 28.65\,\text{kW})$, (b) 0.407 at $(1.704,\ 30.75\,\text{kW})$.

17.4 The checks (step 5)

The plant has three physics rows with a model tolerance, so the limit $\sigma_p \to 0$ is checked on the two objects the chapter reports, the covariance and the step, against the full state at finite tolerances; the analytic Jacobian is checked against differences of the rows, the tangent against the rows themselves, and the spreads against the priors. No draws are taken: before the fit the only posterior formed is the Gaussian at the prior point, and Chapter 19 checks the Gaussian at the fitted state.

17.4.1 The exact balances as the limit of a vanishing tolerance

Table 17.7 checks the limit (17.19) at the prior point. Each column is one pair of tolerances, σ_p in kelvin and $\sigma_{p,P}$ in kilowatts with the same number, and the last column is the reduced covariance; the rows are the posterior spreads of the five free entries, one correlation, the spread of T_1 and the largest difference between the free-entry block and $\mathbf{\Sigma}$.

By $\sigma_p = 0.01\,\text{K}$ every printed spread and the correlation have stopped moving, and the largest difference falls a hundredfold per decade of the tolerance, as the order σ_p^2 of the correction says. The equality holds at every point of the set, and at the fitted state of Chapter 18 the reduced covariance is the one Chapter 19 computes.

The same step in the full state. With the full Jacobian (17.25) at $\sigma_p = 0.01\,\text{K}$, the normal equations $(\mathbf{J}_{\text{full}}^{\mathsf{T}}\mathbf{J}_{\text{full}})\,\Delta\mathbf{z}_{\text{full}} = -\mathbf{J}_{\text{full}}^{\mathsf{T}}\mathbf{r}_{\text{full}}$, eight equations in eight unknowns, give a step whose first five entries agree with $\Delta\mathbf{z}$ of Table 17.6 to 1×10^{-5}; at $\sigma_p = 0.1\,\text{K}$ they still differ in the fourth decimal ($\Delta g_1 = 0.0982$, $\Delta g_2 = -0.1928$). Its last three entries are $\mathbf{N}_{\text{b}}\Delta\mathbf{z} = (0.103\,\text{K},\ 0.004\,\text{K},\ -3.193\,\text{kW})$, the moves of T_1, T_2 and P along the tangent of the balances. In the limit the full step is $\mathbf{N}\,\Delta\mathbf{z}$: the linearized physics rows confine $\Delta\mathbf{z}_{\text{full}}$ to the null directions, and along them the rows of the evidence are the reduced rows. Along the tangent T_2 moves by only 0.004 K, while at the point of the set with the new free entries it sits at 17.463°C;

Table 17.7: The free-entry block of $\boldsymbol{\Sigma}_{\text{full}} = (\mathbf{J}_{\text{full}}^{\mathsf{T}}\mathbf{J}_{\text{full}})^{-1}$ at the prior point, for three tolerances, against the reduced covariance $\boldsymbol{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$ of equation (17.19). The spread of T_1 in the last column is $\sqrt{(\mathbf{N}_{\text{b}}\boldsymbol{\Sigma}\mathbf{N}_{\text{b}}^{\mathsf{T}})_{11}}$.

Quantity	$\sigma_p = 0.1$	0.01	0.001	$\boldsymbol{\Sigma}$, reduced
$\sqrt{\Sigma_{g_1 g_1}}$ [–]	0.1094	0.1083	0.1083	0.1083
$\sqrt{\Sigma_{g_2 g_2}}$ [–]	0.1250	0.1241	0.1241	0.1241
$\sqrt{\Sigma_{Q_1 Q_1}}$ [kW]	1.459	1.459	1.459	1.459
$\sqrt{\Sigma_{Q_2 Q_2}}$ [kW]	2.396	2.394	2.394	2.394
$\sqrt{\Sigma_{T_s T_s}}$ [K]	0.2801	0.2801	0.2801	0.2801
correlation of Q_1, Q_2 [–]	−0.5789	−0.5793	−0.5793	−0.5793
$\sqrt{\Sigma_{T_1 T_1}}$ [K]	0.4804	0.4804	0.4804	0.4804
largest $\lvert(\boldsymbol{\Sigma}_{\text{full}})_{\mathbf{zz}} - \boldsymbol{\Sigma}\rvert$ [mixed]	1.0×10^{-2}	1.0×10^{-4}	1.0×10^{-6}	0

the difference, 0.141 K, is what the linearization leaves in the hall-2 balance, the subject of Section 17.4.3.

17.4.2 The Jacobian against finite differences

An analytic Jacobian is checked against a numerical one before it is used. The central difference of row k in entry i is

$$J_{ki} \approx \frac{r_k(\mathbf{z} + h\,\mathbf{e}_i) - r_k(\mathbf{z} - h\,\mathbf{e}_i)}{2h},$$

with $\mathbf{e}_i$ the i-th unit vector and a step h; its error is of order h^2 times the third derivative of the row. With $h = 10^{-6}$ at the prior point and $\sigma_p = 0.01$ K, the eighty-eight entries of the central-difference Jacobian agree with those of equation (17.25), computed from equations (17.13) and the constant entries of Section 17.1.7, to 2.5×10^{-7} in the largest difference, which is the round-off of the difference quotient on physics rows whose terms reach $1/\sigma_p$ times a temperature, a relative error of 5×10^{-10} on the largest entry. Every structural zero is zero in the numerical Jacobian as well, which confirms the pattern (17.8) and Figure 17.2.

17.4.3 The linearization against the rows

The linearized objective predicted 0.318 at the step, and the objective itself is 0.4085. In the full state the difference has one place. The rows of the evidence are linear in the full state, so along the full step $\mathbf{N}\Delta\mathbf{z}$ they take exactly their tangent values, $r_1 = 0.098$ and $r_2 = -0.192$ for the hall readings and 0.318 for the objective; the whole departure is in the two hall physics rows, which the tangent step leaves 0.025 K and 0.141 K from their balances. The point of the set with the new free entries removes the miss by moving T_1 by −0.025 K and T_2 by −0.141 K, and the hall reading rows become 0.048 and −0.475. The miss is the second-order term (17.24), whose matrix at the prior point is

$$\nabla^2 h_i(\mathbf{z}^{(0)}) = \begin{bmatrix} 5.6 & -0.2 \\ -0.2 & 0 \end{bmatrix}.$$

Figure 17.5 shows the departure for hall 1 along one null direction. The horizontal axis is the log-conductance g_1, with the other free entries at the prior point; the vertical axis is the hall-1 reading row on the reduced state. The solid curve is the row (17.16), $r_1 = (12 + 28\,\mathrm{e}^{-g_1} - 17.654545)/0.5$, the reading compared with the hall on its balance; the dashed line is its tangent at $g_1 = \ln 5$, of slope −11.2, the entry of Table 17.5, the reading compared with the hall moved along the null direction of g_1; the shaded band is $|r_1| \le 1$. The curve is

convex, so the tangent lies below it on both sides, and the gap between them is the miss of equation (17.24) in standard deviations of the sensor, $\frac{1}{2} \times (5.6/0.5) \times \delta g_1^2 = \frac{1}{2} \times 11.2 \times \delta g_1^2$: it is 0.05 at the step of g_1, $\delta g_1 = 0.0985$, and 0.5, half a standard deviation of the sensor, at $\delta g_1 = \pm 0.3$, one standard deviation of the prior. The tangent is therefore a good description of the row over the small moves of the last iterations of a fit, and a poor one over the width of the prior.

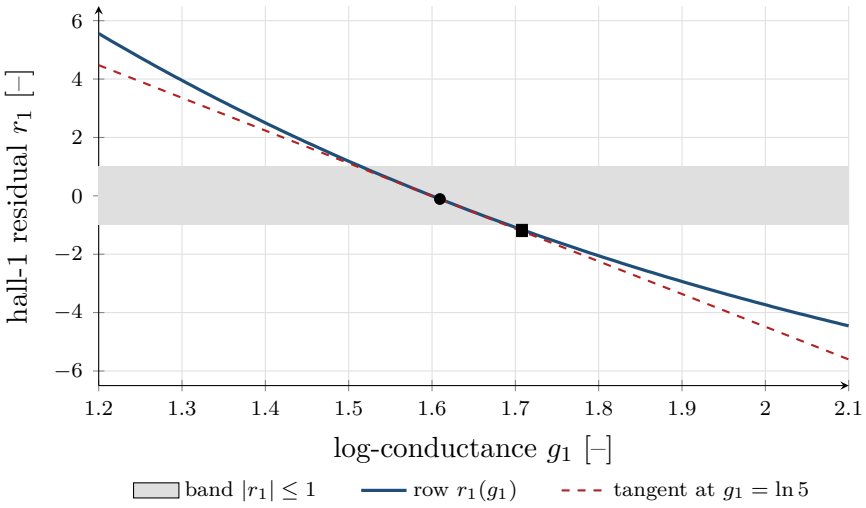


Figure 17.5: The hall-1 reading row against the log-conductance, with $Q_1 = 28\,\text{kW}$ and $T_s = 12°\text{C}$: $r_1 = (12 + 28\,\text{e}^{-g_1} - 17.654545)/0.5$ (solid), its tangent at the prior point $g_1 = \ln 5$, $-0.109 - 11.2\,(g_1 - \ln 5)$ (dashed), and the band of one standard deviation (shaded). Dot: the prior point. Squares: row and tangent at $g_1 = 1.708$, the step's value, -1.160 and -1.213.

For hall 2 the step was larger, and both entries of the product moved at once, $(\Delta g_2, \Delta Q_2) = (-0.1934,\ -5.932\,\text{kW})$. Along the tangent the rise $T_2 - T_s$ changes by the (g_2, Q_2) part of the second row of $\mathbf{N}_\text{b}$ times the step, $(-5.6,\ 0.2) \cdot (-0.1934,\ -5.932) = 1.083 - 1.186$, from 5.6 to 5.497 K. The quadratic form (17.24) splits the correction into the diagonal entry, $\frac{1}{2} \times 5.6 \times 0.1934^2 = +0.105\,\text{K}$, and the off-diagonal pair, $2 \times \frac{1}{2} \times (-0.2) \times (-0.1934) \times (-5.932) = -0.229\,\text{K}$; the rise at the point of the set is in fact $22.068/4.121 = 5.355\,\text{K}$. The off-diagonal entry dominates. It is $-\text{e}^{-g_2}$, the change of the null direction's g_2 component per kilowatt of load: the tangent credits the fall of the conductance with a rise of 1.083 K computed at the old load of 28 kW, but the conductance's fall acts on the smaller load of 22.07 kW. With the diagonal's +0.10 K and the higher orders, the tangent overstates the rise by 0.14 K, the miss of the hall-2 balance. That is 0.28 standard deviations of the sensor, the difference between -0.19 and -0.47 in row 2.

A single linearized step is therefore a preview, not a fit. At $\mathbf{z}^{(1)}$ the gradient of the objective is still 3.06 in norm, almost as large as at the prior point, because the hall-2 row and the prior on g_2 now pull against each other; the step has to be repeated from $\mathbf{z}^{(1)}$ with the Jacobian evaluated there, whose hall-2 row has changed from $(0,\ -11.2,\ 0,\ 0.4,\ 2)$ to $(0,\ -10.71,\ 0,\ 0.485,\ 2)$, twice the second row of $\mathbf{N}_\text{b}$ at the new point. Chapter 18 derives that repetition, the Gauss–Newton iteration, and carries it to convergence from this starting point. Had the conductance itself been the entry, the null direction would move the hall by $-Q_i/G_i^2 = -28/25 = -1.12\,\text{K}$ per kW/K at the prior point, which the hall reading sees as $-1.12/0.5 = -2.24$ per kW/K, an entry that would change as $1/G_i^2$ over the same step.

Common mistake

Reading the linearized prediction as the result. The linearized objective at the step, 0.318, is a prediction of the tangent rows, not a property of the plant; the objective there is 0.4085, and only the iteration of Chapter 18 reaches the minimum. The same caution applies to a slice of the objective such as Figure 17.4: a slice with the other entries held fixed can place the minimum far from where a joint move goes.

17.4.4 Sensitivity to the priors

The three priors are the plant's redundancy, and their weight on the spreads is measured by changing them. Table 17.8 repeats the spreads of Section 17.3.3 at the prior point under five sets of prior rows. Each row of the table is one set; the columns are the spreads $\sqrt{\Sigma_{ii}}$ of the five free entries and the correlation of the two loads. The first row is the priors of Table 17.2; the second and third double and halve the handbook's spread σ_g; the fourth doubles the set point's σ_0; the last removes all three prior rows, which leaves the five readings as a square system on the reduced state.

Table 17.8: Spreads of the free entries at the prior point under five sets of prior rows, from $\boldsymbol{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$ on the reduced state. The last row has no prior rows: the five readings alone, the square system of Section 17.3.1.

Prior rows	$\sqrt{\Sigma_{11}}$ [–]	$\sqrt{\Sigma_{22}}$ [–]	$\sqrt{\Sigma_{33}}$ [kW]	$\sqrt{\Sigma_{44}}$ [kW]	$\sqrt{\Sigma_{55}}$ [K]	$\mathrm{corr}(Q_1, Q_2)$ [–]
as given, $\sigma_g = 0.3$, $\sigma_0 = 1\,\mathrm{K}$	0.1083	0.1241	1.459	2.394	0.2801	−0.579
handbook spread doubled, $\sigma_g = 0.6$	0.1140	0.1329	1.489	2.470	0.2853	−0.594
handbook spread halved, $\sigma_g = 0.15$	0.0918	0.1009	1.384	2.205	0.2668	−0.537
set-point spread doubled, $\sigma_0 = 2\,\mathrm{K}$	0.1089	0.1246	1.459	2.394	0.2887	−0.579
no prior rows, readings only	0.1171	0.1372	1.500	2.500	0.3000	−0.600

The readings carry almost all of the precision. With every prior row removed the spreads grow by eleven percent at most: the conductances' from 0.108 and 0.124 to 0.117 and 0.137, hall 2's load from 2.39 to 2.50 kW, the supply temperature's from 0.280 to 0.300 K, the last two being exactly the heat meter less the load meter, $\sqrt{2^2 + 1.5^2} = 2.5\,\mathrm{kW}$, and the supply sensor's own spread. Doubling the handbook's spread moves every entry by less than the removal does, and doubling the set point's moves essentially only the supply temperature, by 0.009 K. Only a handbook twice as sure as this one changes the conductances' spreads materially, by 15 and 19 %. The priors' role on this plant is the one Section 17.3.1 found: they add little precision, and they are the only rows against which the readings can be checked.

17.5 The report (step 6)

The hand calculation that opened the chapter is the square system of the formulation, and it checks nothing: it reads 5.58 and 3.84 kW/K off the five readings with every residual zero, and a hall-2 sensor four kelvin in error would have given 2.17 kW/K with every residual zero as well. Set against the handbook conductances and the set point, with the loads profiled out, the reading that disagrees most is the hall-2 sensor, at −1.32 standard deviations and 40 % of the objective: hall 2 runs warmer than the handbook conductance allows, the direction of the hand calculation's low conductance, but within the readings' errors. The formulation keeps

three rows of redundancy, and all three are the handbook and set-point priors; they add little precision, but they are what lets a wrong reading show.

17.5.1 The answers and their bands

Table 17.9 gives the answers to the four questions of Table 17.3, in its order. Each row is a question; the second column gives the answer, from the reduced state, and the third its band: for the spreads the range over the priors of Table 17.8 that do not claim more than the handbook, and for the others the check of Section 17.4 or the count that bounds them.

Table 17.9: The report of the formulation of the two-hall plant before the fit, on the reduced state: the predictive check at the handbook values with the loads profiled out, $\tilde{\mathbf{z}}$, and the other answers at the prior point $\mathbf{z}^{(0)}$.

Question	Answer	Band
Do the readings agree with the handbook values, the set point and each other, and can that be checked?	$\mathcal{L}(\tilde{\mathbf{z}}) = 2.15$; largest row the hall-2 sensor, -1.32 standard deviations and 40 % of it; the meters at $+1.05$ and -1.06; smallest two-sided tail 0.19	$\nu = 3$ redundant rows, all three the priors; without them a square system that fits any readings
Which entries does each reading see, and do the readings determine all five?	rank five: the readings alone determine all five; the conductances and Q_2 only through the balances, each in combination with another entry; weakest direction mostly hall 2's load, sd = 2.70 kW	2.58 kW with the priors; smallest singular value 0.371 of the reading block, 0.388 with the priors
How well would the readings fix each entry?	0.108 and 0.124 in g_1, g_2; 1.46 kW and 2.39 kW in Q_1, Q_2; 0.280 K in T_s; $\mathrm{corr}(Q_1, Q_2) = -0.58$	Q_2 from 2.39 to 2.50 kW, g_2 from 0.124 to 0.137, over the priors as given, doubled and removed
Where does the first step of the fit take the state?	$G_1 = 5.52$, $G_2 = 4.12$ kW/K, $Q_1 = 30.74$, $Q_2 = 22.07$ kW, $T_s = 12.107$°C; $\mathcal{L}$ from 3.6128 to 0.4085	tangent's prediction 0.318; the same step in the full state to 10^{-5}

17.5.2 The formulation handed to the fit

None of the four answers is the state of the plant. The first step stops where the hall-2 row and the prior on g_2 pull against each other, the spreads are taken at a point the fit will leave, and the threshold and the next measurement for hall 1 at 60 kW wait for the fitted state. What the formulation fixes does not wait: the pattern of the Jacobian and the count of redundant rows hold at every state, and the full rank of the design block at any point of the set near this one. Chapter 18 derives the objective from Bayes' theorem for this plant, iterates the linearized step to the maximum a posteriori state, and checks the fit by its residuals, its chi-square and the physical admissibility of the result.

17.6 The next measurement (step 7)

The report leaves hall 2 the weakly known hall: its load, which no meter reads, is the heat meter's total less hall 1's meter, known to 2.39 kW, and its conductance, the quantity the opening's engineer worried about, to a log-spread of 0.124. The candidates are the three measurements that Chapter 20 weighs at the fitted state for hall 1's scenario temperature, and the hall-2 twin of the first; here they are weighed at the prior point for the spreads of the five free entries.

17.6.1 The candidate rows

Each candidate is one more row, and on the reduced state its Jacobian row is its row of the full state times $\mathbf{N}$. At the prior point the four are:

- *A second hall-1 temperature sensor* of $\sigma = 0.5\,\mathrm{K}$, reading T_1 as the first one does, with the reduced row $\mathbf{J}_a = (-11.2,\ 0,\ 0.4,\ 0,\ 2)$, the first row of Table 17.5;
- *a second hall-2 temperature sensor* of $\sigma = 0.5\,\mathrm{K}$, with $\mathbf{J}_a = (0,\ -11.2,\ 0,\ 0.4,\ 2)$, the second row of Table 17.5;
- *an added supply-temperature sensor* of $\sigma = 0.1\,\mathrm{K}$, beside the present one, with $\mathbf{J}_a = (0,\ 0,\ 0,\ 0,\ 10)\,\mathrm{K}^{-1}$;
- *an electrical meter on hall 2's load* of $\sigma = 1.5\,\mathrm{kW}$, the twin of hall 1's, with $\mathbf{J}_a = (0,\ 0,\ 0,\ 0.667,\ 0)\,\mathrm{kW}^{-1}$, the only candidate whose row touches Q_2 directly.

Each adds one row and no entry, so each raises the redundancy from three to four, and each extra row is checked against readings rather than against a prior: two sensors on one entry check each other, and a third meter on two loads and their sum checks the other two.

17.6.2 The preposterior of the spreads

A candidate row is Gaussian and linear in the reduced state, so the covariance it would leave does not depend on its unseen reading, and the preposterior covariance of equation (2.17) is, for one row,

$$\boldsymbol{\Sigma}_a = \left(\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a\right)^{-1} = \boldsymbol{\Sigma} - \frac{\boldsymbol{\Sigma}\mathbf{J}_a^{\mathsf{T}}\,\mathbf{J}_a\boldsymbol{\Sigma}}{1 + \mathbf{J}_a\boldsymbol{\Sigma}\mathbf{J}_a^{\mathsf{T}}}, \tag{17.30}$$

the second form being the rank-one update of the inverse. The width a candidate removes from an entry is $1 - \sqrt{\Sigma_{a,ii}}/\sqrt{\Sigma_{ii}}$, and its expected information gain over all five entries is equation (2.18). For the hall-2 load meter, $(\boldsymbol{\Sigma}\mathbf{J}_a^{\mathsf{T}})_4 = 3.8196$ and $\mathbf{J}_a\boldsymbol{\Sigma}\mathbf{J}_a^{\mathsf{T}} = 2.5464$, so

$$\Sigma_{a,44} = 5.7295 - \frac{3.8196^2}{1 + 2.5464} = 5.7295 - \frac{14.590}{3.5464} = 5.7295 - 4.1139 = 1.6156\,\mathrm{kW}^2,$$

a spread of 1.271 kW for hall 2's load, a width removed of 46.9 %.

Table 17.10 lists the four candidates. Each row is the present formulation or one candidate; the first five columns are the spreads of the free entries after the candidate, with the width removed in parentheses where it exceeds one percent, and the last is the expected information gain in nats.

Table 17.10: The preposterior of four candidate rows at the prior point, equation (17.30): the spread of each free entry after the candidate, with the width removed where it exceeds 1 %, and the expected information gain, equation (2.18).

Candidate	$\sqrt{\Sigma_{a,11}}$ [–]	$\sqrt{\Sigma_{a,22}}$ [–]	$\sqrt{\Sigma_{a,33}}$ [kW]	$\sqrt{\Sigma_{a,44}}$ [kW]	$\sqrt{\Sigma_{a,55}}$ [K]	EIG [nat]
present	0.1083	0.1241	1.459	2.394	0.2801	–
second hall-1 sensor, 0.5 K	0.0927 (14.4 %)	0.1241	1.458	2.393	0.2800	0.33
second hall-2 sensor, 0.5 K	0.1083	0.1120 (9.7 %)	1.458	2.390	0.2800	0.33
added supply sensor, 0.1 K	0.0993 (8.3 %)	0.1171 (5.7 %)	1.459	2.393	0.0942 (66.4 %)	1.09
hall-2 load meter, 1.5 kW	0.1054 (2.6 %)	0.1054 (15.0 %)	1.271 (12.9 %)	1.271 (46.9 %)	0.2800	0.63

Each candidate narrows what its row reads. The second hall-1 sensor narrows hall 1's conductance by a seventh and nothing else; it also takes the spread of T_1 from 0.480 to 0.346 K.

Its hall-2 twin gains the same information, 0.33 nats, but narrows hall 2's conductance by only 9.7 %: that conductance is tied more strongly to its loosely known load, $\mathrm{corr}(g_2, Q_2) = 0.62$ against $\mathrm{corr}(g_1, Q_1) = 0.44$ in hall 1, and a second reading of the rise leaves the load's share of its spread in place. The supply sensor gains the most information, 1.09 nats, but it spends it on the supply temperature, already known to 0.28 K, whose width it cuts by two thirds, and it narrows the conductances by only 8 and 6 %, through the hall rows that share T_s. The hall-2 load meter is the only candidate that reaches hall 2's load: it nearly halves the spread of hall 2's load, narrows hall 2's conductance by 15 % and hall 1's load by 13 % through the heat meter's sum, and loosens the tie between the loads from $\mathrm{corr}(Q_1, Q_2) = -0.58$ to -0.35. After it the two halls are known equally well, each conductance to 0.105 and each load to 1.27 kW.

17.6.3 The recommendation

The next measurement for the questions of this chapter is the electrical meter on hall 2's load. It is the one candidate that turns hall 2's load from a difference of two meters into a reading, and with it the one that sharpens hall 2's conductance, the quantity the hand calculation put a quarter below the handbook, by 15 % against the second hall-2 sensor's 9.7 %. Its row of redundancy is the one that checks the split of the load: three meters on two loads and their sum, where at present the split is checked only through the handbook's prior on g_2. A better supply sensor would add more information in total, almost all of it on the supply temperature. For hall 1's temperature at 60 kW, Chapter 20 ranks three of these candidates at the fitted state and puts the second hall-1 sensor first: the best next measurement depends on the quantity asked about.

17.6.4 What the analysis showed

The hand calculation read hall 2's conductance as 3.84 kW/K, about a quarter below the handbook, and reproduced every reading; the formulation shows that it would have read a faulty hall-2 sensor the same way and had no reading to spare. Set against the handbook with the loads profiled out, the reading that disagrees most is the hall-2 sensor, at -1.32 standard deviations and 40 % of the objective: hall 2 runs warm for its share of the load, the sign of a low conductance, but not by more than the readings' errors allow. Whether that is a fouled coil or a load larger than the meters say is a question about the split of the load between the halls. The heat meter sees only the total, hall 1's meter takes its share, and hall 2 is left with the remainder and both meters' errors, which makes its load the less known of the two and its conductance, read through that load and its own rise, less sure than hall 1's. The handbook and the chiller's set point add almost nothing to what the readings fix, but they are the only rows the readings can be checked against. What would improve the picture is not a better temperature sensor but a meter on hall 2's own load.

17.6.5 Limits of the model

Five assumptions carry the formulation, and each is a choice made when the graph was built.

Steady state. The halls and the loop store no heat, so the balances have no storage term and one set of readings describes one operating condition. A record taken through a change of load needs the transient of Section 1.6.5, with one hall temperature per time as an entry.

One lumped path per hall. Each hall's heat leaves by one path with one effective conductance; heat lost through the envelope, or a conductance that changes with the fan speed, would need further paths or entries.

Exact balances. The balances hold in the limit $\sigma_p \to 0$: the pump work, the pipe gains and any heat that bypasses the coils are neglected, and if they matter the tolerance takes a physical value and the balance becomes a model discrepancy.

Normal, independent, unbiased reading errors. A miscalibrated meter violates the last of these, and at one snapshot its bias is indistinguishable from a change of the state it reads; its detection is a diagnosis question of the chapter on bad data, Chapter 22.

Electrical power as heat. The electrical power of hall 1's equipment is taken to be its heat load, which holds in steady state when every watt delivered is dissipated in the hall.

The linearization is not an assumption of the formulation but of one step of its computation. The rows themselves are exact under the assumptions above; the tangent of Section 17.4.3 is used only to choose the next point, and its errors disappear as the iteration converges. The covariance at the prior point is an assumption of the same kind: it is the curvature at a point the fit will leave, and the Laplace approximation of Chapter 4, which uses the curvature at the minimum, makes its own assumption about the posterior, whose failure Chapter 19 states.

17.7 What to retain

Definition 17.1 (Physics-constrained state estimation)**.** Every free quantity of a plant graph is an entry of the full state $\mathbf{z}_{\text{full}}$. Each balance is a physics row, the balance's residual divided by a model tolerance σ_p, and the exact balance is the limit $\sigma_p \to 0$; each reading is a normalized row $(z_i - y_k)/\sigma_k$ on the entry it reads, and each prior a row $(z_i - z_i^0)/\sigma$. The null directions $\mathbf{N}$ of the physics rows, $\mathbf{J}_{\text{p}}\mathbf{N} = \mathbf{0}$, span the reduced state $\mathbf{z}$ of free entries; its rows are the reading and prior rows on the balance set, its Jacobian is the projection $\mathbf{J} = \mathbf{J}_{\text{e}}\mathbf{N}$, and its covariance $(\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$ is the free-entry block of the full-state covariance in the limit. The residual vector defines the objective $\mathcal{L} = \frac{1}{2}\mathbf{r}^{\mathsf{T}}\mathbf{r}$ and the posterior $p(\mathbf{z} \mid \mathbf{y}) \propto \exp(-\mathcal{L})$; the Jacobian states what each row touches, the rank of the reduced reading rows whether the readings alone determine the state, and the excess of rows over entries the redundancy available for a check. One linearized step solves $\mathbf{J}^{\mathsf{T}}\mathbf{J}\,\Delta\mathbf{z} = -\mathbf{J}^{\mathsf{T}}\mathbf{r}$; the information matrix of normalized rows is $\mathbf{J}^{\mathsf{T}}\mathbf{J}$, and of unnormalized rows $\tilde{\mathbf{J}}^{\mathsf{T}}\mathbf{W}\tilde{\mathbf{J}}$.

Four conclusions carry beyond the two halls. A reading is evidence about the state only through a physical row: in the full state no reading row touches a conductance, and the reduced Jacobian $\mathbf{J}_{\text{e}}\mathbf{N}$ is what turns the hall readings into rows over the conductances, the loads and the supply temperature. The exact balance is a limit, checked numerically on the covariance and on the step, and the reduced state of the following chapters is obtained from it in matrix form, never by a substitution made by hand. The Jacobian's pattern of zeros and the columns of $\mathbf{N}_{\text{b}}$, read before any number is processed, already say which entries each reading sees and that the heat meter cannot separate the two loads. And a formulation is checkable only through its rows to spare: here they are the three priors, which add little precision and all of the redundancy, while a single linearized step, which takes the objective from 3.6128 to 0.4085 against a prediction of 0.318, is a preview of the fit and not the fit. Table 17.11 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 17.11: The terms of physics-constrained state estimation and where each is defined.

Term	Expression	Units	Meaning	Defined in
Hall balance	$Q_i = G_i\,(T_i - T_s)$	kW	heat in equals heat removed	17.1.4
Header balance	$P = Q_1 + Q_2$	kW	chiller removes the halls' heat	17.1.4
Full state	$\mathbf{z}_{\text{full}} = (g_1, g_2, Q_1, Q_2, T_s, T_1, T_2, P)$	–, kW, °C	every free quantity of the graph	17.1.4
Physics row	$(T_i - T_s - Q_i \mathrm{e}^{-g_i}/G_{\text{ref}})/\sigma_p$	–	a balance with its tolerance	17.1.7
Reading row	$(T_i - y_{T_i})/\sigma_T$	–	a reading compared with its entry	17.1.7
Normalized residual	$r_k = d_k/\sigma_k$, weight one	–	disagreement in standard deviations	17.1.7
Full residual vector	$\mathbf{r}_{\text{full}} \in \mathbb{R}^{11}$, eq. (17.10)	–	all rows at once	17.1.7
Full Jacobian	$\mathbf{J}_{\text{full}}$, 11×8, eq. (17.25)	$1/z_i$	response of each row to each entry	17.1.7
Physics-row entries	$(T_i - T_s)/\sigma_p$, $-1/(G_i\sigma_p)$, $-1/\sigma_p$, $1/\sigma_p$	varies	rise, reciprocal conductance, supply, hall	17.1.7
Null directions	$\mathbf{N} = [\mathbf{I}_5;\ \mathbf{N}_{\text{b}}]$, $\mathbf{J}_{\text{p}}\mathbf{N} = \mathbf{0}$	varies	moves that keep the balances	17.1.7
Reduced state	$\mathbf{z} = (g_1, g_2, Q_1, Q_2, T_s)$	–, kW, °C	the free directions	17.2.1
Reduced rows	$\mathbf{r}(\mathbf{z}) \in \mathbb{R}^8$, eq. (17.15)	–	evidence on the balance set	17.2.1
Reduced Jacobian	$\mathbf{J} = \mathbf{J}_{\text{e}}\mathbf{N}$, 8×5	$1/z_i$	what each row sees of the free directions	17.2.1
Design columns	rows of $\mathbf{J}$ on the readings	–	what the readings alone see	17.2.1
Weakest direction	$\mathbf{v}$, $\mathrm{sd}(\mathbf{v}^{\mathsf{T}}\mathbf{z} \mid \mathbf{y}) = 1/s_{\min}$	varies	the combination the readings see least	17.3.2
Reduced covariance	$(\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1} = \lim(\boldsymbol{\Sigma}_{\text{full}})_{\mathbf{zz}}$	varies	the exact balance as a limit	17.2.1
Objective	$\mathcal{L} = \frac{1}{2}\,\mathbf{r}^{\mathsf{T}}\mathbf{r}$	–	minus log posterior	17.2.1
Prior point	$\mathbf{z}^{(0)} = (\ln 5, \ln 5, 28, 28, 12)$	–	priors at their means	17.2.2
Gradient	$\nabla\mathcal{L} = \mathbf{J}^{\mathsf{T}}\mathbf{r} = \mathbf{N}^{\mathsf{T}}\mathbf{J}_{\text{e}}^{\mathsf{T}}\mathbf{r}_{\text{e}}$	$1/z_i$	direction of steepest rise	17.2.2
Normal equations	$\mathbf{J}^{\mathsf{T}}\mathbf{J}\,\Delta\mathbf{z} = -\mathbf{J}^{\mathsf{T}}\mathbf{r}$	–	one linearized step	17.2.2
Linearized objective	$\frac{1}{2}\lVert\mathbf{r}^{(0)} + \mathbf{J}^{(0)}\Delta\mathbf{z}\rVert^2$	–	the tangent's prediction	17.2.2
Balance Hessian	$\nabla^2 h_i$, eq. (17.24)	varies	the miss a tangent step leaves	17.2.2
Profiled loads	$\tilde{Q}_1$, $\tilde{Q}_2$, eq. (17.26)	kW	loads that best fit the handbook state	17.3.1
Redundancy	$\nu = 11 - 8 = 8 - 5 = 3$	–	rows available for a check	17.3.1
Square system	rows = entries	–	fits any readings, checks nothing	17.3.1
Preposterior	$\boldsymbol{\Sigma}_a = (\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a)^{-1}$	varies	a candidate row's covariance	17.6.2

17.8 Exercises

Exercise 17.1. *Formulation.* One hall is cooled from a supply held exactly at $T_s = 12°\mathrm{C}$ and carries a known load $Q = 30\,\mathrm{kW}$; its conductance is unknown, with the handbook prior $g \sim \mathcal{N}(\ln 5,\ 0.3^2)$, and one sensor reads $y_T = 17.655°\mathrm{C}$ with $\sigma_T = 0.5\,\mathrm{K}$. (a) Write the full state, the physics row and the two rows of the evidence. (b) Count the degrees of freedom. (c) Evaluate the rows and the objective at the prior mean of g with T on its balance. ((a) $\mathbf{z}_{\text{full}} = (g,\ T)^{\mathsf{T}}$; rows $r_1 = (T - 17.655)/0.5$, $r_2 = (g - \ln 5)/0.3$, $r_3 = (T - 12 - 30\,\mathrm{e}^{-g})/\sigma_p$. (b) Three rows, two entries: one degree of freedom, as two rows over the one free direction. (c) At $(\ln 5,\ 18°\mathrm{C})$, $\mathbf{r} = (0.69,\ 0,\ 0)$ and $\mathcal{L} = \frac{1}{2} \times 0.69^2 = 0.238$.)

Exercise 17.2. *Formulation.* For the two-hall plant: (a) count rows, entries and degrees of freedom of the full and the reduced state with all readings and priors, without the hall-1 load meter, without the three priors, and with a second sensor in hall 1; (b) list, from Figure 17.2, the entries each reading row touches, and, from Table 17.5, the free directions each reading sees; (c) state what a storage term in the hall balance would change in the state vector and the rows. ((a) Full: $11, 8, 3$; $10, 8, 2$; $8, 8, 0$, a square system; $12, 8, 4$. Reduced: $8, 5, 3$; $7, 5, 2$; $5, 5, 0$; $9, 5, 4$. (b) Each reading row touches one entry (T_1, T_2, T_s, P, Q_1); along the null directions the hall readings see their own g_i and Q_i and T_s, the supply sensor T_s, the heat meter Q_1 and Q_2 with equal weight, the load meter Q_1. (c) With $C_{\text{th}}\,\mathrm{d}T_i/\mathrm{d}t = Q_i - G_i(T_i - T_s)$, where C_{th} is the hall's thermal capacitance in kJ/K (Section 1.6.5), the hall temperature at every time of the record becomes an entry, and the physics row becomes the discretized balance linking two consecutive times.)

Exercise 17.3. *Derivation.* (a) Derive the four nonzero Jacobian entries of the hall-i physics row and the null direction's components $(\mathbf{N}_{\text{b}})_{i\cdot}$. (b) Show that in the conductance itself the null direction moves T_i by $-Q_i/G_i^2$ per kW/K, evaluate what the hall reading sees of it at the prior point, and show that the handbook prior cannot be written as a row linear in G_i. ((a) With $\sigma_p r_{8+i} = T_i - T_s - Q_i\mathrm{e}^{-g_i}/G_{\text{ref}}$: $(Q_i\mathrm{e}^{-g_i}/G_{\text{ref}},\ -1/G_i,\ -1,\ 1)/\sigma_p$ in (g_i, Q_i, T_s, T_i), and $(\mathbf{N}_{\text{b}})_{i\cdot} = -\mathbf{J}_{\text{pb}}^{-1}\mathbf{J}_{\text{pa}}$ has $(-(T_i - T_s),\ 1/G_i,\ 1)$ in (g_i, Q_i, T_s). (b) $\partial(Q_i/G_i)/\partial G_i = -Q_i/G_i^2 = -1.12\,\mathrm{K}$ per kW/K, which the reading sees as $-1.12/0.5 = -2.24$ per kW/K. With $g_i \sim \mathcal{N}(\ln 5, 0.3^2)$, G_i has the log-normal density $\exp[-(\ln G_i - \ln 5)^2/(2 \times 0.09)]/(G_i\, 0.3\sqrt{2\pi})$, whose minus logarithm $\ln G_i + (\ln G_i - \ln 5)^2/0.18$ is not the square of a linear function of G_i; a normal prior $\mathcal{N}(5, 1.5^2)$ in its place, the log-normal's spread to first order ($0.3 \times 5\,\mathrm{kW/K}$, a 30 % spread against the 35 % factor $\mathrm{e}^{0.3}$), would put probability $\Phi(-3.33) = 4.3 \times 10^{-4}$ on negative conductances.)

Exercise 17.4. *Derivation.* For the hall of the first exercise: (a) write the 3×2 Jacobian at $(\ln 5,\ 18°\mathrm{C})$ with $\sigma_p = 0.01\,\mathrm{K}$, the null direction $\mathbf{n}$ of the physics row, and the reduced Jacobian $\mathbf{J}_{\text{e}}\mathbf{n}$; (b) take one Gauss–Newton step along $\mathbf{n}$, and compare the reading and prior rows after the step, the objective at the point of the balance with the same g, and the exact minimum. ((a) Rows (reading, prior, physics), columns (g, T): $\mathbf{J}_{\text{full}} = [\,0,\ 2;\ 3.333,\ 0;\ 600,\ 100\,]$; $\mathbf{n} = (1,\ -6)^{\mathsf{T}}$, since the balance moves T by $-30\,\mathrm{e}^{-g} = -6\,\mathrm{K}$ per unit of g; $\mathbf{J}_{\text{e}}\mathbf{n} = (-12,\ 3.333)^{\mathsf{T}}$. (b) $\Delta g = -(\mathbf{J}_{\text{e}}\mathbf{n})^{\mathsf{T}}\mathbf{r}_{\text{e}}/\|\mathbf{J}_{\text{e}}\mathbf{n}\|^2 = 12 \times 0.69/(144 + 11.11) = 0.0534$, so $\Delta\mathbf{z}_{\text{full}} = (0.0534,\ -0.320\,\mathrm{K})$, the same step the full normal equations give at $\sigma_p = 0.01\,\mathrm{K}$; the reading and prior rows then give 0.0171, the linearized prediction, and the hall misses its balance by 0.008 K; on the balance at $g = 1.6628$, $G = 5.274\,\mathrm{kW/K}$, the objective is 0.0180, and the exact minimum is 0.0179 at $g = 1.6640$, $G = 5.280\,\mathrm{kW/K}$: one step comes within 1 % of the minimum because the balance is nearly linear over a step of 0.05.)

Exercise 17.5. *Application.* A third hall with its own sensor ($\sigma_T = 0.5\,\mathrm{K}$) and handbook prior is added to the loop, and the loop's return temperature is read by a sensor of $\sigma_{T_r} = 0.3\,\mathrm{K}$, with the loop flow metered at $\dot{m} = 3.0\,\mathrm{kg/s}$ (taken as exact) and $c_p = 4.19\,\mathrm{kJ/(kg\,K)}$; the three halls together carry about 80 kW. (a) Write the full state, the return-temperature reading row and

the loop balance $Q_1 + Q_2 + Q_3 = \dot{m}c_p(T_r - T_s)$ as a physics row. (b) Count rows, entries and degrees of freedom. (c) Give the return rows' Jacobian entries and what the return reading sees along the null directions. ((a) $\mathbf{z}_{\text{full}} = (g_1, g_2, g_3, Q_1, Q_2, Q_3, T_s, T_1, T_2, T_3, P, T_r)$; reading row $(T_r - y_{T_r})/\sigma_{T_r}$; physics row $(T_r - T_s - (Q_1 + Q_2 + Q_3)/(\dot{m}c_p))/\sigma_p$ with $\dot{m}c_p = 12.57\,\text{kW/K}$, a rise of 6.4 K at 80 kW. (b) Seven reading rows, four priors (g_1, g_2, g_3, T_s) and five physics rows (three halls, the header, the loop): sixteen rows over twelve entries, four degrees of freedom; on the reduced state of seven free entries, eleven rows. (c) Reading row: $1/\sigma_{T_r} = 3.33\,\text{K}^{-1}$ on T_r only; physics row: $1/\sigma_p$ on T_r, $-1/\sigma_p$ on T_s, $-1/(\dot{m}c_p\sigma_p)$ on each load. The null directions move T_r by $1/(\dot{m}c_p) = 0.0796\,\text{K}$ per kilowatt of each load and by 1 K per kelvin of T_s, so the return reading sees $0.265\,\text{kW}^{-1}$ on each load, $3.33\,\text{K}^{-1}$ on T_s and nothing of the conductances: a second meter of the total load, of equivalent standard deviation $\dot{m}c_p\sigma_{T_r} = 3.8\,\text{kW}$, coupled to the supply temperature.)

Exercise 17.6. *Application.* The building management system logs each hall's temperature rise $y_{\text{rise},i} = y_{T_i} - y_{T_s}$, computed from the hall sensor and the shared supply sensor, and the analyst attaches the two rises and the supply reading as three independent rows. (a) Write the rise rows on the full state and their rows of the reduced Jacobian. (b) Compute the covariance of the three errors and say what treating them as independent does. (c) State two correct formulations. ((a) $(T_i - T_s - y_{\text{rise},i})/\sigma_{\text{rise}}$ with $\sigma^2_{\text{rise}} = \sigma^2_T + \sigma^2_s = 0.34\,\text{K}^2$, touching T_i with $+1/\sigma_{\text{rise}}$ and T_s with $-1/\sigma_{\text{rise}}$; the null direction of T_s moves T_i and T_s together, so along it the two entries cancel, and the reduced rows are $(-(T_i - T_s),\ 1/G_i,\ 0)/\sigma_{\text{rise}}$ in (g_i, Q_i, T_s): the rises see the conductances and loads but not the supply. (b) $\text{Var}(y_{\text{rise},i}) = 0.34$, $\text{Cov}(y_{\text{rise},1}, y_{\text{rise},2}) = \sigma^2_s = 0.09$, $\text{Cov}(y_{\text{rise},i}, y_{T_s}) = -0.09\,\text{K}^2$, a correlation of 0.26 between the rises: the supply sensor's error enters three rows, and independent weights count it three times and overstate the information. (c) Attach the raw readings y_{T_1}, y_{T_2}, y_{T_s} as in equation (17.10); or keep the rises with the full weight matrix $\mathbf{W} = \mathbf{C}^{-1}$ of their covariance $\mathbf{C}$, equivalently rows whitened by the Cholesky factor of $\mathbf{C}$, whose Jacobian is that factor's inverse times the raw one.)

Chapter 18

Maximum a Posteriori Estimation

18.1 Fitting the state of a plant to its readings

The two data halls of Chapter 17 have reported five readings at one steady operating condition: 17.655 and 17.700°C in the halls, 12.10°C at the supply header, 52.5 kW on the chiller's heat meter and 31.0 kW on the electrical meter of hall 1. The facilities engineer, looking for a faulty sensor, compares each reading with what the handbook says it should be. With the handbook's conductance of 5 kW/K, hall 1, carrying its meter's 31.0 kW, should stand $31.0/5 = 6.2$ K above the supply sensor's 12.10°C, at 18.30°C; hall 2, carrying the 21.5 kW that the chiller removes beyond hall 1's meter, should stand at $12.10 + 21.5/5 = 16.40$°C. The hall-1 sensor reads 17.655°C, 1.3 of its standard deviations of 0.5 K low; the hall-2 sensor reads 17.700°C, 2.6 standard deviations high. The first reading of the record is that the hall-2 sensor is suspect and should be recalibrated.

The engineer who must sign the work order asks what that comparison rests on. It takes the handbook's conductance as exact, although the handbook gives it only within 35 %; it takes the loads as known, although they come from meters with errors of 1.5 and 2 kW; and it judges each reading alone, although the five readings and the three priors bear on the same five unknowns through the balances, so that a hall reading out of line with the handbook may say more about the hall's heat-removal path than about its sensor. The question of this chapter is what the reading-by-reading comparison does not settle,

what the most probable state of the plant is, given its balances, its readings and its prior information, and whether the readings and the physics agree, judged by rows that could have disagreed.

Jargon: maximum a posteriori (MAP) state, Gauss–Newton iteration, normal equations, exact Hessian, standardized and normalized residual, redundancy, degrees of freedom of a fit, hat matrix, redundancy number, chi-square distribution, noncentral chi-square distribution, square system, admissibility, stopping rule, Levenberg–Marquardt damping

The *maximum a posteriori (MAP) state* is the value of the state at which the posterior density is highest, the minimizer of the objective. The *Gauss–Newton iteration* minimizes a sum of squared rows by replacing the rows, at each iterate, by their first-order expansion and solving the resulting linear least-squares problem, whose *normal equations* are $\mathbf{J}^\mathsf{T}\mathbf{J}\,\Delta\mathbf{z} = -\mathbf{J}^\mathsf{T}\mathbf{r}$. The *exact Hessian* is the full matrix of second derivatives of the objective; the Gauss–Newton iteration uses only its part $\mathbf{J}^\mathsf{T}\mathbf{J}$. A *standardized residual* is a row's misfit at the fit divided by the row's standard deviation. The *redundancy* of a fit is the number of rows beyond the number of state entries, $\nu = n_r - n_z$, also called its *degrees of freedom*. The *hat matrix* $\mathbf{K}$ maps the rows' data onto their fitted values; its complement $1 - K_{kk}$ on the diagonal is the *redundancy number* of row k, the variance of its standardized residual at the fit, and the standardized residual divided by the root of its redundancy number is the *normalized residual* of state-estimation practice, the statistic of bad-data detection. The *chi-square distribution* with ν degrees of freedom, χ^2_ν, is the distribution of a sum of ν squared independent standard normal variables; the *noncentral chi-square distribution* $\chi^2_\nu(c)$ is that of a sum of ν squared independent normal variables of unit variance whose squared means add to the noncentrality c. A *square system* has as many rows as entries and no redundancy. An estimate is *admissible* when it satisfies the physical conditions the model assumes, such as positive conductances and loads. A *stopping rule* decides when an iteration has converged. *Levenberg–Marquardt damping* adds a multiple of a positive diagonal matrix to $\mathbf{J}^\mathsf{T}\mathbf{J}$ to shorten steps taken far from the minimum.

Key idea

The MAP state of a plant minimizes the sum of squared normalized rows, readings and priors together, and the Gauss–Newton iteration finds it by solving one linear least-squares problem per step. The MAP state is the answer to the estimation question only together with the checks that say the model and the readings agree: the standardized residuals of every row, and the minimized objective, whose double is chi-square with as many degrees of freedom as there are redundant rows. Without redundant rows the fit reproduces any readings exactly and the check is empty; with them, the check sees a reading only as far as the redundancy lies on that reading.

18.1.1 The fit and the question it answers

The estimation question that Chapter 17 posed for this part of the book asks what the two conductances, the two loads and the supply temperature are, given the readings. Its object is the posterior of the state, $p(\mathbf{z} \mid \mathbf{y})$, and the first summary of that posterior is its highest point, the MAP state $\hat{\mathbf{z}}$, with the curvature of the objective there as its spread. For the three-entry hall state, whose rows are all linear, Chapter 4 reached it in one Gauss–Newton step, the solution of one set of normal equations. With five entries and two rows that multiply a load by an exponential of a log-conductance there is no closed form: the MAP state is the solution of five nonlinear equations, and it is computed by iteration, the rows linearized at each iterate.

The MAP state is also the point at which the later chapters of this part are evaluated. Chapter 19 reads from the curvature of the objective there which entries the readings determine, Chapter 20 propagates the covariance to engineering outputs through gradients taken there, and the chapters on virtual sensors, bad data and sensor placement start from the same point.

An error in the fit is therefore inherited by every answer built on it.

18.1.2 Three checks on one fit

A minimizer always exists for a sum of squares that grows without bound, and an iteration will return one whether or not the model describes the plant. The number returned is the answer to the estimation question only if the readings and the model agree, and three checks decide that. The first is the standardized residual of each row at the fit, which states, in units of the row's own standard deviation, how far the fitted state is from that reading or that prior. The second is the minimized objective itself: for a model whose rows are correct and whose standard deviations are right, twice its value is distributed as chi-square with as many degrees of freedom as there are rows beyond the entries, and a value far in the upper tail of that distribution says that some row disagrees with the rest. The third is physical: the fitted conductances and loads must be positive and the fitted hall temperatures must lie above the supply temperature, since heat flows from the halls to the chilled water.

18.1.3 The plant and its state (step 0)

The plant is the graph of Figure 17.1 of Chapter 17: two hall nodes, each with a heat load as its source and a heat-removal path to the supply header, the header carrying the supply temperature, and the chiller as the source of the chilled water. The fit is carried on the reduced state of Chapter 17,

$$\mathbf{z} = \begin{bmatrix} g_1 \\ g_2 \\ Q_1 \\ Q_2 \\ T_s \end{bmatrix}, \qquad g_i = \ln(G_i/G_{\mathrm{ref}}), \quad G_{\mathrm{ref}} = 1\,\mathrm{kW/K}, \tag{18.1}$$

with the log-conductances g_i dimensionless, the loads Q_i in kW and the supply temperature T_s in °C; $n_z = 5$ entries. The plant has three physical rows (equations (17.4) and (17.5)), the steady energy balance of each hall, $Q_i = G_i(T_i - T_s)$, and the balance of the supply header, $P = Q_1 + Q_2$, with the chiller's heat duty P in kW. Chapter 17 wrote them as physics rows with model tolerances, σ_p in kelvin for the hall balances and $\sigma_{p,P}$ in kilowatts for the header (neither is the chiller meter's σ_P, which belongs to a reading), on a full state of eight entries that also carries T_1, T_2 and P, and projected that state once onto the null directions $\mathbf{N}$ of the physics rows' Jacobian (equation (17.14)), so that the reduced state has the Jacobian $\mathbf{J} = \mathbf{J}_{\mathrm{e}}\mathbf{N}$ and, as the tolerances vanish, the free-entry marginal of the full covariance (equations (17.18) and (17.19)). On the reduced state the hall temperatures and the duty are functions of the state,

$$T_i(\mathbf{z}) = T_s + Q_i\,\mathrm{e}^{-g_i}, \qquad i = 1, 2, \qquad P(\mathbf{z}) = Q_1 + Q_2, \tag{18.2}$$

in °C and kW, with $G_{\mathrm{ref}} = 1\,\mathrm{kW/K}$ absorbed in the exponential, and they are written as such wherever they appear below.

18.1.4 The questions (step 1)

Table 18.1 lists the questions asked of the fit: the first column states each in engineering language, the second gives the general form of its category (Chapter 2), the third its mathematical form over the state, the fourth its category, and the last the subsection where it is answered. The first is the estimation question, answered at the highest point of the posterior with the spread of each entry. The second asks whether the evidence could have disagreed with the fit: its event is that a record drawn afresh from the model misfits at least as much as the one observed, a

tail of a predictive distribution, taken row by row through the normalized residuals and for the whole fit through the minimized objective. The third asks whether the fitted state satisfies the physical conditions the model assumes, a joint event on the state. The atlas of Appendix A lists all three.

Table 18.1: The questions asked of the two-hall fit, the general form of their category, their mathematical form over the reduced state, their category and where each is answered. $\mathbf{y}$ is the five readings and the three priors of Table 18.2; $\mathbf{y}_{\mathrm{rep}}$ a record drawn afresh from the model and $X^2(\cdot)$ its misfit statistic; r_k^{N} the normalized residual of row k; $\nu = 3$ the number of redundant rows.

Engineering question	General form	Mathematical form	Category	Answered in
What are the conductances, the loads and the supply temperature?	$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y})$, $f(\mathbf{z}) = z_i$	$\hat{\mathbf{z}} = \arg\max_{\mathbf{z}} p(\mathbf{z} \mid \mathbf{y})$, with $\sqrt{\Sigma_{ii}}$	estimation	18.3.1
Do the readings agree with the physics and with each other?	$\Pr(X^2(\mathbf{y}_{\mathrm{rep}}) \ge X^2(\mathbf{y}))$ under the predictive of $\mathbf{y}_{\mathrm{rep}}$	r_k^{N} against $\mathcal{N}(0,1)$; $\Pr(\chi^2_\nu > 2\mathcal{L}(\hat{\mathbf{z}}))$	prediction under a threshold	18.3.2
Is the estimate physically admissible?	$\Pr(f(\mathbf{z}) \ge f^* \mid \mathbf{y})$ for a joint event	$\Pr(G_i > 0,\ Q_i > 0,\ T_i > T_s,\ i = 1, 2 \mid \mathbf{y})$	prediction under a threshold	18.3.3

How much the handbook priors decide the point is not a question about the plant but a check of the answer, its sensitivity to the prior (Section 18.4.5), with the fit without priors as its limit (Section 18.4.6). How many iterations the fit needs, and whether a second start reaches the same state, belong to the computation of the posterior (Section 18.2.3) and its checks (Section 18.4.1).

18.1.5 The evidence as rows (step 2)

Step 2 attaches the evidence to the reduced state as rows: five readings and three handbook priors, all Gaussian, each written below in normalized form, with a residual and a Jacobian row.

The modeling choices. Four choices make the rows and the posterior, and each buys something and can fail. The readings are taken as independent given the state, each with a normal error of the stated standard deviation: this is what makes the likelihood a product of Gaussian factors and the objective a sum of squared rows, and it fails if two sensors share a calibration or a stated standard deviation is wrong, which the chi-square check of the second question sees only in part (Section 18.6.5). The three balances are taken in the limit of vanishing tolerances, $\sigma_p \to 0$ for the halls and $\sigma_{p,P} \to 0$ for the header, on the reduced state of Section 18.1.3: this keeps as the state the five entries the operators ask for, with the balances held, and it fails for a hall not yet at steady state or a balance that misses by a physical amount, which Section 18.4.3 checks by fitting the full state at three tolerances. The handbook conductance enters as a normal prior on the log-conductance and the set point as a normal prior on the supply temperature, while the loads carry none: this lets the handbook's relative spread of 35 % be a row like a reading, and it misleads if the handbook describes coils unlike these, which Section 18.4.5 measures by narrowing and widening the prior and Section 18.4.6 by removing it. The spread is read from the Gauss–Newton curvature $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ at the MAP state rather than from the exact Hessian: this needs only the Jacobian already formed for the fit, and

it would mislead for a fit whose residuals are large on strongly curved rows, which Section 18.4.2 checks.

Table 18.2 lists the eight rows in normalized form, each divided by its standard deviation so that every row has unit weight. Each line gives the row's number k, what it is, its expression $r_k(\mathbf{z})$, the reading or prior value it compares with, and its standard deviation σ_k. Rows 1 and 2 compare the hall temperature readings with the hall temperatures $T_i(\mathbf{z})$ of equation (18.2); row 3 compares the supply temperature reading with the entry T_s; row 4 compares the chiller's heat meter with the duty $P(\mathbf{z}) = Q_1 + Q_2$; row 5 compares the electrical meter on hall 1's load with Q_1; rows 6 and 7 are the handbook priors on the two log-conductances, centered on a conductance of 5 kW/K with a spread of 0.3 in the logarithm, about 35 %; row 8 is the prior on the supply temperature, centered on the set point. The loads carry no prior row. There are $n_r = 8$ rows, and

$$\nu = n_r - n_z = 8 - 5 = 3 \tag{18.3}$$

rows are redundant, the count of equation (17.27): the same three as the eleven rows of the full state over its eight entries, since each of T_1, T_2 and P brings one entry and one physics row. The readings are the illustrative values of Chapter 17; the hall-1 reading, printed there to six decimals as 17.654545°C, is rounded to 17.655°C in this chapter's text, and every number of this chapter is computed with its unrounded value.

Table 18.2: The eight rows of the two-hall plant in normalized form (Chapter 17). $\sigma_T = 0.5\,\mathrm{K}$ for the hall sensors, $\sigma_s = 0.3\,\mathrm{K}$ for the supply sensor, $\sigma_P = 2\,\mathrm{kW}$ for the chiller meter, $\sigma_Q = 1.5\,\mathrm{kW}$ for the hall-1 meter, $\sigma_g = 0.3$ for the conductance priors and $\sigma_0 = 1\,\mathrm{K}$ for the set-point prior; $g^0 = \ln 5$ and $T_s^0 = 12°\mathrm{C}$.

k	Row	$r_k(\mathbf{z})$	Reading or prior	σ_k
1	hall-1 temperature	$(T_1(\mathbf{z}) - y_{T_1})/\sigma_T$	$y_{T_1} = 17.655°\mathrm{C}$	0.5 K
2	hall-2 temperature	$(T_2(\mathbf{z}) - y_{T_2})/\sigma_T$	$y_{T_2} = 17.700°\mathrm{C}$	0.5 K
3	supply temperature	$(T_s - y_{T_s})/\sigma_s$	$y_{T_s} = 12.10°\mathrm{C}$	0.3 K
4	chiller heat meter	$(P(\mathbf{z}) - y_P)/\sigma_P$	$y_P = 52.5\,\mathrm{kW}$	2 kW
5	hall-1 load meter	$(Q_1 - y_{Q_1})/\sigma_Q$	$y_{Q_1} = 31.0\,\mathrm{kW}$	1.5 kW
6	prior on g_1	$(g_1 - g^0)/\sigma_g$	$g^0 = \ln 5 = 1.6094$	0.3
7	prior on g_2	$(g_2 - g^0)/\sigma_g$	$g^0 = \ln 5 = 1.6094$	0.3
8	prior on T_s	$(T_s - T_s^0)/\sigma_0$	$T_s^0 = 12°\mathrm{C}$	1 K

Which row acts on which entry is a property of the graph and its sensors, and it fixes where the Jacobian of the rows can be nonzero, marked $*$, with the columns in the order of equation (18.1) and the rows labeled at the right:

$$\begin{bmatrix} * & 0 & * & 0 & * \\ 0 & * & 0 & * & * \\ 0 & 0 & 0 & 0 & * \\ 0 & 0 & * & * & 0 \\ 0 & 0 & * & 0 & 0 \\ * & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * \end{bmatrix} \begin{matrix} \text{hall-1 reading} \\ \text{hall-2 reading} \\ \text{supply reading} \\ \text{chiller heat meter} \\ \text{hall-1 load meter} \\ \text{prior on } g_1 \\ \text{prior on } g_2 \\ \text{prior on } T_s \end{matrix} \tag{18.4}$$

A hall row carries three entries because the balance of its hall carries the reading to the hall's conductance, its load and the supply temperature; the heat meter carries the two loads through the header balance, and every other row carries one entry. The first column says that the conductance of hall 1 is seen by its temperature reading and its prior and by nothing else, and the second the same for hall 2. The fourth says that the second load is seen only through

the hall-2 reading and the heat meter, each time together with another entry. The fifth is the supply temperature, shared by both hall readings, its own sensor and the set point, the entry that couples the two halls.

The residual vector and its Jacobian. Stacked in the order of Table 18.2, the eight rows form the residual vector

$$\mathbf{r}(\mathbf{z}) = \begin{bmatrix} (T_s + Q_1 \mathrm{e}^{-g_1} - y_{T_1})/\sigma_T \\ (T_s + Q_2 \mathrm{e}^{-g_2} - y_{T_2})/\sigma_T \\ (T_s - y_{T_s})/\sigma_s \\ (Q_1 + Q_2 - y_P)/\sigma_P \\ (Q_1 - y_{Q_1})/\sigma_Q \\ (g_1 - g^0)/\sigma_g \\ (g_2 - g^0)/\sigma_g \\ (T_s - T_s^0)/\sigma_0 \end{bmatrix}, \tag{18.5}$$

eight dimensionless numbers for each candidate state, with $T_i(\mathbf{z})$ and $P(\mathbf{z})$ written out. Their derivatives form the 8×5 Jacobian $\mathbf{J}(\mathbf{z}) = \partial \mathbf{r}/\partial \mathbf{z}^\mathsf{T}$, tabulated at the prior point in Table 17.5, with the nonzero entries of the pattern (18.4). Only the two hall rows are nonlinear. Their nonzero derivatives are

$$\frac{\partial r_i}{\partial g_i} = -\frac{Q_i \mathrm{e}^{-g_i}}{\sigma_T} = -\frac{T_i(\mathbf{z}) - T_s}{\sigma_T}, \qquad \frac{\partial r_i}{\partial Q_i} = \frac{\mathrm{e}^{-g_i}}{\sigma_T}, \qquad \frac{\partial r_i}{\partial T_s} = \frac{1}{\sigma_T}, \qquad i = 1, 2, \tag{18.6}$$

and every other row has constant derivatives: $1/\sigma_s$ on T_s for row 3, $1/\sigma_P$ on each load for row 4, $1/\sigma_Q$ on Q_1 for row 5, $1/\sigma_g$ and $1/\sigma_0$ on their entries for the priors. The derivative of a hall row with respect to its log-conductance is minus the temperature rise of the hall in units of the sensor's standard deviation, the property that makes a reading informative about a conductance.

With every row normalized the information matrix is $\mathbf{J}^\mathsf{T}\mathbf{J}$; written with the unnormalized rows, whose Jacobian is $\tilde{\mathbf{J}}$, and the weights $\mathbf{W} = \mathrm{diag}(1/\sigma_k^2)$, it is the same matrix $\tilde{\mathbf{J}}^\mathsf{T}\mathbf{W}\tilde{\mathbf{J}}$ (equation (4.26)).

18.2 The posterior of the state (step 3)

In the book's notation, $p(\mathbf{z} \mid \mathbf{y})$ is the posterior density of the state, $f(\mathbf{z})$ a quantity derived from it and f^* a threshold on that quantity. The *MAP state* $\hat{\mathbf{z}}$ is the minimum of the objective that the eight rows put on the reduced state; the *information matrix* $\mathbf{\Lambda} = \mathbf{J}^\mathsf{T}\mathbf{J}$ is the Gauss–Newton curvature of that objective at the minimum, with $\mathbf{J}$ the Jacobian of the normalized rows; and the *Laplace covariance* $\mathbf{\Sigma} = \mathbf{\Lambda}^{-1}$ makes the posterior of the state approximately normal, $\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}, \mathbf{\Sigma})$. The problem admits no conjugate closed form, since the two hall rows are curved in the state.

All three questions are read from one posterior: the posterior of the reduced state that the eight rows of step 2 give, approximately the Laplace normal about the MAP state. This section derives it and computes it for the two halls; Section 18.3 reads every answer from it.

18.2.1 The posterior from the rows

Bayes' theorem for the state reads

$$p(\mathbf{z} \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid \mathbf{z})\, p(\mathbf{z})}{p(\mathbf{y})}, \tag{18.7}$$

with $p(\mathbf{y})$ the evidence, which does not depend on $\mathbf{z}$. Write each reading as $y_k = h_k(\mathbf{z}) + \varepsilon_k$, $k = 1, \ldots, 5$, as Chapter 17 wrote its measurement models, where $h_k(\mathbf{z})$ is the value the model

predicts for the reading ($T_1(\mathbf{z})$ for row 1, $P(\mathbf{z})$ for row 4, and so on) and the errors $\varepsilon_k \sim \mathcal{N}(0, \sigma_k^2)$ are independent. Given the state, the readings are then independent, and the likelihood is the product of their densities,

$$p(\mathbf{y} \mid \mathbf{z}) = \prod_{k=1}^{5} \frac{1}{\sqrt{2\pi}\,\sigma_k} \exp\left[-\frac{\big(h_k(\mathbf{z}) - y_k\big)^2}{2\sigma_k^2}\right] = \prod_{k=1}^{5} \frac{1}{\sqrt{2\pi}\,\sigma_k}\, \mathrm{e}^{-r_k(\mathbf{z})^2/2},$$

since $r_k = (h_k - y_k)/\sigma_k$ is the normalized row. The prior is the product of independent Gaussian densities on g_1, g_2 and T_s, each written in the same way through its row, and a flat density on the two loads, constant in $\mathbf{z}$:

$$p(\mathbf{z}) \propto \prod_{k=6}^{8} \frac{1}{\sqrt{2\pi}\,\sigma_k}\, \mathrm{e}^{-r_k(\mathbf{z})^2/2}.$$

Multiplying the two in equation (18.7),

$$p(\mathbf{z} \mid \mathbf{y}) \propto \prod_{k=1}^{8} \mathrm{e}^{-r_k(\mathbf{z})^2/2} = \exp\left[-\frac{1}{2}\sum_{k=1}^{8} r_k(\mathbf{z})^2\right],$$

and taking the negative logarithm,

$$-\ln p(\mathbf{z} \mid \mathbf{y}) = \frac{1}{2}\sum_{k=1}^{8} r_k(\mathbf{z})^2 + \sum_{k=1}^{8} \ln(\sqrt{2\pi}\,\sigma_k) + \ln p(\mathbf{y}) + c = \mathcal{L}(\mathbf{z}) + \text{const}, \tag{18.8}$$

with c minus the logarithm of the flat prior's constant. The objective is

$$\mathcal{L}(\mathbf{z}) = \frac{1}{2}\sum_{k=1}^{n_r} r_k(\mathbf{z})^2 = \frac{1}{2}\,\mathbf{r}(\mathbf{z})^{\mathsf{T}}\mathbf{r}(\mathbf{z}), \tag{18.9}$$

the form $\frac{1}{2}\sum_k w_k r_k^2$ of the notation chapter with every weight equal to one because the rows are normalized. The constant collects the standard deviations, 2π and the evidence; none depends on the state, so maximizing the posterior is minimizing $\mathcal{L}$:

$$\hat{\mathbf{z}} = \arg\max_{\mathbf{z}}\, p(\mathbf{z} \mid \mathbf{y}) = \arg\min_{\mathbf{z}}\, \mathcal{L}(\mathbf{z}). \tag{18.10}$$

For a three-entry state and four rows this is the objective of equation (4.24); the derivation above is the same product of Gaussian factors carried over eight rows.

The condition for the maximum. At an interior minimum of $\mathcal{L}$ its gradient vanishes. Differentiating equation (18.9) row by row,

$$\frac{\partial \mathcal{L}}{\partial z_i} = \sum_{k=1}^{n_r} r_k(\mathbf{z})\,\frac{\partial r_k}{\partial z_i},$$

which in matrix form is

$$\nabla\mathcal{L}(\mathbf{z}) = \mathbf{J}(\mathbf{z})^{\mathsf{T}}\,\mathbf{r}(\mathbf{z}) = \mathbf{0}. \tag{18.11}$$

With the derivatives (18.6) the five equations read

$$\begin{aligned}
g_1: &\quad -\frac{Q_1 \mathrm{e}^{-g_1}}{\sigma_T}\, r_1 + \frac{r_6}{\sigma_g} = 0,\\
g_2: &\quad -\frac{Q_2 \mathrm{e}^{-g_2}}{\sigma_T}\, r_2 + \frac{r_7}{\sigma_g} = 0,\\
Q_1: &\quad \frac{\mathrm{e}^{-g_1}}{\sigma_T}\, r_1 + \frac{r_4}{\sigma_P} + \frac{r_5}{\sigma_Q} = 0,\\
Q_2: &\quad \frac{\mathrm{e}^{-g_2}}{\sigma_T}\, r_2 + \frac{r_4}{\sigma_P} = 0,\\
T_s: &\quad \frac{r_1}{\sigma_T} + \frac{r_2}{\sigma_T} + \frac{r_3}{\sigma_s} + \frac{r_8}{\sigma_0} = 0.
\end{aligned} \tag{18.12}$$

Each equation is a balance of pulls on one entry: every row that touches the entry, a nonzero of its column in the pattern (18.4), pulls on it with its misfit times its sensitivity to the entry, and at the MAP state the pulls cancel. The conductance of hall 1 is pulled by its temperature reading and by its prior and by nothing else; the supply temperature by both hall readings, its own sensor and the set point. The equations are nonlinear through the products $Q_i \mathrm{e}^{-g_i}$, which appear in the rows and in their derivatives, and they have no closed-form solution.

Gauss–Newton as repeated linearization. Chapter 17 took one step from the prior point by linearizing the rows there; the Gauss–Newton iteration repeats that step. Let $\mathbf{z}^{(m)}$ be the iterate after m steps, with residual vector $\mathbf{r}^{(m)} = \mathbf{r}(\mathbf{z}^{(m)})$ and Jacobian $\mathbf{J}^{(m)} = \mathbf{J}(\mathbf{z}^{(m)})$, evaluated at the iterate rather than at $\hat{\mathbf{z}}$. The first-order expansion of the rows at the prior point, equation (17.22), turned the objective into a quadratic whose minimizer solves the normal equations (17.23). The same derivation at $\mathbf{z}^{(m)}$ in place of the prior point gives the iteration

$$\left(\mathbf{J}^{(m)\mathsf{T}}\mathbf{J}^{(m)}\right)\Delta\mathbf{z}^{(m)} = -\mathbf{J}^{(m)\mathsf{T}}\mathbf{r}^{(m)}, \qquad \mathbf{z}^{(m+1)} = \mathbf{z}^{(m)} + \Delta\mathbf{z}^{(m)}. \tag{18.13}$$

When $\mathbf{J}^{(m)}$ has full column rank the step is unique. Two properties of the iteration follow. The right-hand side is minus the gradient (18.11) at the iterate, so a fixed point of the iteration, $\Delta\mathbf{z} = \mathbf{0}$, is a point where the gradient vanishes: the iteration stops exactly at a solution of the five equations (18.12). And the step is a descent direction, because

$$\nabla\mathcal{L}(\mathbf{z}^{(m)})^{\mathsf{T}}\Delta\mathbf{z}^{(m)} = -\Delta\mathbf{z}^{(m)\mathsf{T}}\mathbf{J}^{(m)\mathsf{T}}\mathbf{J}^{(m)}\Delta\mathbf{z}^{(m)} < 0$$

for any nonzero step, since $\mathbf{J}^{(m)\mathsf{T}}\mathbf{J}^{(m)}$ is positive definite: a short enough move along it lowers the objective. In computation the normal equations are not formed; the step is the least-squares solution of $\mathbf{J}^{(m)}\Delta\mathbf{z} \approx -\mathbf{r}^{(m)}$, obtained from an orthogonal factorization of $\mathbf{J}^{(m)}$, which gives the same step with less round-off.

The Laplace covariance. At the MAP state the matrix of the normal equations is the information matrix of the rows, and its inverse is the Laplace covariance of the reduced state,

$$\mathbf{\Lambda} = \mathbf{J}(\hat{\mathbf{z}})^{\mathsf{T}}\mathbf{J}(\hat{\mathbf{z}}), \qquad \mathbf{\Sigma} = \mathbf{\Lambda}^{-1} = \left(\mathbf{J}(\hat{\mathbf{z}})^{\mathsf{T}}\mathbf{J}(\hat{\mathbf{z}})\right)^{-1}, \tag{18.14}$$

the curvature of the objective in its Gauss–Newton form, evaluated once more at the reported state. Its diagonal gives the spreads $\sqrt{\Sigma_{ii}}$ of the first question, and through the hat matrix $\mathbf{K} = \mathbf{J}\mathbf{\Sigma}\mathbf{J}^{\mathsf{T}}$ the redundancy numbers of the second; Chapter 4 derived it as the Laplace approximation, and Chapter 19 reads from it which entries the readings determine.

18.2.2 The exact Hessian and the iteration

Newton's method, which Chapter 12 applied to a Poisson objective, steps with the exact Hessian of the objective. Differentiating the gradient $\sum_k r_k\, \partial r_k/\partial z_i$ once more,

$$\frac{\partial^2\mathcal{L}}{\partial z_i\,\partial z_j} = \sum_k \frac{\partial r_k}{\partial z_i}\frac{\partial r_k}{\partial z_j} + \sum_k r_k\,\frac{\partial^2 r_k}{\partial z_i\,\partial z_j},$$

which in matrix form is

$$\mathbf{H}(\mathbf{z}) = \nabla^2\mathcal{L}(\mathbf{z}) = \mathbf{J}^{\mathsf{T}}\mathbf{J} + \mathbf{S}, \qquad \mathbf{S}(\mathbf{z}) = \sum_{k=1}^{n_r} r_k(\mathbf{z})\,\nabla^2 r_k(\mathbf{z}). \tag{18.15}$$

The Gauss–Newton matrix $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ is the first term; the iteration of equation (18.13) is Newton's method with $\mathbf{S}$ dropped. In the two-hall plant only the hall rows have second derivatives. From equation (18.6),

$$\frac{\partial^2 r_i}{\partial g_i^2} = \frac{Q_i\mathrm{e}^{-g_i}}{\sigma_T} = \frac{T_i - T_s}{\sigma_T}, \qquad \frac{\partial^2 r_i}{\partial g_i\,\partial Q_i} = -\frac{\mathrm{e}^{-g_i}}{\sigma_T}, \qquad i = 1, 2, \tag{18.16}$$

with every other second derivative zero, so $\mathbf{S}$ has nonzero entries only in the (g_i, g_i) and (g_i, Q_i) positions, each multiplied by the residual r_i of its hall row. These second derivatives are the curvature of the hall balance itself, carried into the reading row by the reduced state: in the full state the readings are linear and the same term is the balance's curvature weighted by its residual, which does not vanish as $\sigma_p \to 0$, because the residual shrinks as σ_p while the curvature grows as $1/\sigma_p$, and in the limit it equals the term above.

Each entry of $\mathbf{S}$ is a residual times a curvature, so the dropped term is small when either factor is small: when the residuals are small at the fit, or when the rows are nearly linear. It matters chiefly when a row is both badly fitted and strongly curved. For a hall row the comparison can be made row by row. Its own Gauss–Newton term at (g_i, g_i) is the square of its derivative (18.6), $\left((T_i - T_s)/\sigma_T\right)^2$, and its contribution to $\mathbf{S}$ there is $r_i(T_i - T_s)/\sigma_T$, so

$$\frac{|S^{(i)}_{g_i g_i}|}{(\mathbf{J}^\mathsf{T}\mathbf{J})^{(i)}_{g_i g_i}} = \frac{|r_i|\,\sigma_T}{T_i - T_s}, \tag{18.17}$$

the residual divided by the temperature rise expressed in standard deviations of the sensor, with the superscript (i) marking the contribution of row i alone. A hall whose rise is ten sensor standard deviations keeps the dropped term below a tenth of the kept one unless its residual exceeds one standard deviation. Its effects on the iteration and on the answer differ. The two methods solve the same equation (18.11) and stop at the same point, so the MAP state does not depend on whether $\mathbf{S}$ is kept. What changes is the rate of approach: Newton's method converges quadratically near the minimum, while the Gauss–Newton iteration converges linearly, the error shrinking on each step by a factor close to the spectral radius of $(\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}\mathbf{S}$ at $\hat{\mathbf{z}}$, which is near zero for a small-residual fit. The covariance is the other place where the dropped term could matter, through $\boldsymbol{\Sigma}$ of equation (18.14) against the inverse of $\mathbf{H}(\hat{\mathbf{z}})$.

Two cases of Volume I had the term in a form that can be read off. The Poisson trend rows of Chapter 12 are count rows in the working form of Section 4.2.7, and because the log-rate is the canonical link of the Poisson count their curvature term vanishes at every state: $\mathbf{J}^\mathsf{T}\mathbf{J}$ is the exact Hessian there (equation (12.8)), and Newton and Gauss–Newton take the same step. The Weibull rows of Chapter 7 are count rows too, working residuals $(\Lambda_i - \delta_i)/\sqrt{\Lambda_i}$ (equation (7.33), the Poisson pair of equation (4.43) with $\mu = \Lambda_i$) with their terms $\mathcal{L}_i$ (equation (7.29)), but the log-shape is not the canonical link of those counts: the curvature term $\sum_i (\Lambda_i - \delta_i)\,\partial^2 \ln \Lambda_i / \partial \mathbf{z}\,\partial \mathbf{z}^\mathsf{T}$ of equation (7.34) does not vanish, and that chapter kept it, so that its information is $\mathbf{J}^\mathsf{T}\mathbf{J}$ plus the term, which at the MAP state adds the number of failures to the (v, v) entry (equation (7.41)). The term $\mathbf{S}$ of equation (18.15) is its Gaussian counterpart, the residual of each row times the row's own curvature, where the count rows carry the observed-minus-expected difference times the curvature of the log mean; Section 18.4.2 evaluates it at the MAP state of the two halls.

The stopping rule. An iteration is stopped on a step measured against the size of the state, not on an absolute quantity. The entries mix kilowatts, degrees and dimensionless log-conductances; the gradient norm is dominated by the stiffest columns, and an absolute tolerance on either the gradient or the step means a different relative accuracy for every entry and, for a large-magnitude entry, can sit below the round-off of the entry itself. The rule used here stops when every entry has moved by less than a fraction tol of its own size,

$$\max_i \frac{|\Delta z_i^{(m)}|}{1 + |z_i^{(m+1)}|} < \text{tol}, \qquad \text{tol} = 10^{-5}, \tag{18.18}$$

the 1 in the denominator guarding entries near zero.

The computation as an algorithm. The fit needs the rows, their Jacobian and a starting point, and it evaluates only linear least-squares solutions and the normal distribution function.

1. *Inputs.* The state with its units; the rows in normalized form, each with its reading or prior value and standard deviation (Table 18.2); a starting point $\mathbf{z}^{(0)}$, such as the prior point of equation (17.20); the tolerance tol (default 10^{-5}).
2. *Checks.* Require every $\sigma_k > 0$ and a Jacobian of full column rank at the start; report $\nu = n_r - n_z$ and stop with a warning that the chi-square check is empty if $\nu \le 0$.
3. *Step.* At $\mathbf{z}^{(m)}$ evaluate $\mathbf{r}^{(m)}$ and $\mathbf{J}^{(m)}$, and solve $\mathbf{J}^{(m)}\Delta\mathbf{z} \approx -\mathbf{r}^{(m)}$ by least squares, equation (18.13).
4. *Safeguard.* If $\mathcal{L}(\mathbf{z}^{(m)} + \Delta\mathbf{z}) \ge \mathcal{L}(\mathbf{z}^{(m)})$, halve the step until the objective decreases, which a descent direction guarantees; or, in the Levenberg–Marquardt form, solve $(\mathbf{J}^{\mathsf{T}}\mathbf{J} + \varpi\ \mathrm{diag}(\mathbf{J}^{\mathsf{T}}\mathbf{J}))\Delta\mathbf{z} = -\mathbf{J}^{\mathsf{T}}\mathbf{r}$ with a damping $\varpi > 0$ that is raised after a failed step and lowered after a successful one.
5. *Stop.* Accept the step and stop when the relative step rule (18.18) holds; otherwise return to step 3.
6. *Outputs.* The MAP state $\hat{\mathbf{z}}$, the objective $\mathcal{L}(\hat{\mathbf{z}})$, the standardized residuals $r_k(\hat{\mathbf{z}})$, the redundancy numbers $1 - K_{kk}$ and the normalized residuals, $2\mathcal{L}(\hat{\mathbf{z}})$ with ν and its tail $\Pr(\chi^2_\nu > 2\mathcal{L}(\hat{\mathbf{z}}))$, and $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ evaluated once more at the reported state for the covariance (18.14).
7. *Checks of the result.* The admissibility conditions of Section 18.3.3; a second run from a distant start; for a strongly nonlinear plant, the refits of records drawn from the model (Section 18.4.4).

18.2.3 The posterior of the two halls

The plant's inputs are the eight rows of Table 18.2 with the readings and priors listed there, the hall-1 reading taken unrounded, the starting point of Chapter 17 and the tolerance tol $= 10^{-5}$ of the stopping rule (18.18).

The iteration starts at the prior point of Chapter 17,

$$\mathbf{z}^{(0)} = (\ln 5,\ \ln 5,\ 28,\ 28,\ 12), \tag{18.19}$$

the handbook conductance for both halls, an even split of a total load near the chiller meter's reading, and the set point. There Chapter 17 found $\mathcal{L}(\mathbf{z}^{(0)}) = 3.6128$, carried almost entirely by the two meters, which the even split contradicts by 1.75 and 2.00 standard deviations (Table 17.4). The first step of equation (18.13) is the linearized step of Chapter 17; it moves the hall-1 load to 30.74 kW and the hall-2 load to 22.07 kW to satisfy both meters, and lowers the conductance of hall 2 to 4.12 kW/K, toward the value at which its smaller load still produces its reading; the objective falls to 0.4085.

Table 18.3 follows the iteration. Each row is one iterate m; the columns give the conductances $G_i = \mathrm{e}^{g_i}$ (the state carries g_i), the loads and the supply temperature, the objective, the Euclidean norm of its gradient (18.11), and the norm of the step taken from that iterate.

Table 18.3: Gauss–Newton iterations for the two-hall plant from the prior point $\mathbf{z}^{(0)} = (\ln 5, \ln 5, 28, 28, 12)$. Conductances in kW/K, loads in kW, T_s in °C; the gradient and step norms mix the units of the entries. Computed with the unrounded hall-1 reading.

m	G_1	G_2	Q_1	Q_2	T_s	$\mathcal{L}(\mathbf{z}^{(m)})$	$\lVert\nabla\mathcal{L}\rVert$	$\lVert\Delta\mathbf{z}^{(m)}\rVert$
0	5.000	5.000	28.000	28.000	12.000	3.6128	3.22	6.54
1	5.518	4.121	30.739	22.068	12.107	0.4085	3.06	0.175
2	5.491	4.057	30.687	22.234	12.112	0.3698	2.7×10^{-2}	3.7×10^{-3}
3	5.490	4.055	30.688	22.231	12.111	0.3698	7.0×10^{-4}	1.1×10^{-4}
4	5.490	4.055	30.688	22.231	12.111	0.3698	1.7×10^{-5}	2.3×10^{-6}

The objective falls $3.6128 \to 0.4085 \to 0.3698$ and stays there: three steps bring the state to its final value at the digits printed, and the objective to within 3×10^{-9} of its minimum. The gradient norm hardly changes on the first step, 3.22 to 3.06, although the objective falls by a factor of nine: it is dominated by the two conductance columns, whose curvature in $\mathbf{J}^\mathsf{T}\mathbf{J}$ is about 130, so that at $\mathbf{z}^{(1)}$ the remaining offset of 0.016 in g_2, a difference of 1.6 % in the conductance, gives with its coupling to the hall-2 load a gradient entry of 2.9, nearly the whole norm. After the second step it falls by a nearly constant factor, 0.027, 7.0×10^{-4}, 1.7×10^{-5}, a ratio of 0.021 to 0.026 per step, the linear rate of Section 18.2.2. The largest relative steps of the rule (18.18) are 0.26, 7.2×10^{-3}, 1.9×10^{-4} and 4.5×10^{-6} for $m = 0$ to 3: the fourth step meets the rule, and $\mathbf{z}^{(4)}$ is reported, within 2×10^{-6} of the minimum in every entry. No safeguard of the algorithm is triggered.

The Jacobian, the information and the covariance. The Jacobian at the prior point, where both rises are 5.6 K, is Table 17.5. At the MAP state, where the rises are 5.590 and 5.482 K, only the four curved entries of the hall rows have moved, by equation (18.6): the hall-1 row is $(-11.18,\ 0,\ 0.3643,\ 0,\ 2)$ and the hall-2 row $(0,\ -10.96,\ 0,\ 0.4932,\ 2)$, in the columns of equation (18.1), dimensionless for the log-conductances, in kW^{-1} for the loads and in K^{-1} for T_s; the other six rows are those of the table. At $\hat{\mathbf{z}}$ the information matrix and the Laplace covariance of equation (18.14), with rows and columns ordered as in equation (18.1), are

$$\mathbf{J}^\mathsf{T}\mathbf{J} = \begin{bmatrix} 136.08 & 0 & -4.0724 & 0 & -22.358 \\ 0 & 131.34 & 0 & -5.4081 & -21.930 \\ -4.0724 & 0 & 0.82715 & 0.25 & 0.72856 \\ 0 & -5.4081 & 0.25 & 0.49327 & 0.98646 \\ -22.358 & -21.930 & 0.72856 & 0.98646 & 20.111 \end{bmatrix},$$

$$\boldsymbol{\Sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1} = \begin{bmatrix} 0.01138 & -0.000417 & 0.06358 & -0.06274 & 0.01297 \\ -0.000417 & 0.01876 & -0.07977 & 0.2222 & 0.01199 \\ 0.06358 & -0.07977 & 2.109 & -1.949 & 0.002914 \\ -0.06274 & 0.2222 & -1.949 & 5.505 & -0.02688 \\ 0.01297 & 0.01199 & 0.002914 & -0.02688 & 0.07843 \end{bmatrix}, \tag{18.20}$$

in the units of the products of the entries' units.

The pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Table 18.3 and equation (18.20) is what step 4 uses: every answer of the chapter is computed from it. The exact Hessian (18.15) gives the spreads set beside the reported ones, and records drawn from the model at $\hat{\mathbf{z}}$ give the misfit tail set beside its chi-square value.

18.3 The answers (step 4)

The questions of Table 18.1 fall into two categories: estimation (the first) and prediction under a threshold (the second, a tail of the predictive distribution of a replicated record, and the third, a joint event on the state). Each is answered from the posterior of Section 18.2: its engineering question, the general form of its category, the form it takes for the two halls, the quantities of the posterior it uses, and the plant's numbers. Every answer is computed from the Gauss–Newton point $\hat{\mathbf{z}}$ and its Laplace covariance $\boldsymbol{\Sigma}$ of step 3, and that is the value the chapter reports; the exact Hessian (18.15), the curvature with the term that the Gauss–Newton matrix drops, gives the spreads set beside it, and refits of records drawn from the model give the misfit tail set beside its chi-square value, each with the gap.

18.3.1 The conductances, the loads and the supply temperature

Engineering question. What are the conductances, the loads and the supply temperature? The question asks for the five entries of the state given the five readings and the three priors,

and how well each is known.

General form. An estimation question about an entry asks for its posterior distribution, equation (2.9) with $f(\mathbf{z}) = z_i$,

$$\Pr(z_i \le f^* \mid \mathbf{y}) = \int \mathbf{1}\big[z_i \le f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

summarized by a named location and a spread, equation (2.10).

Specialized form. The question is equations (2.9) and (2.10) with the state $\mathbf{z} = (g_1, g_2, Q_1, Q_2, T_s)^\mathsf{T}$ of equation (18.1), the derived quantity $f(\mathbf{z}) = z_i$, each entry in turn (and T_1, T_2, P of equation (18.2) for the derived quantities), and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \mathbf{\Sigma})$ of equations (18.14) and (18.20). Under the Laplace posterior $\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}, \mathbf{\Sigma})$ the posterior of an entry is normal, with location $\hat{z}_i$ and spread $\sqrt{\Sigma_{ii}}$, and two entries are tied by the correlation $\Sigma_{ij}/\sqrt{\Sigma_{ii}\Sigma_{jj}}$. The location named is the mode of the posterior of the state, the MAP state, which under the Laplace posterior is also the mean and the median of every entry. The hall temperatures and the chiller's duty are derived quantities, $T_i(\hat{\mathbf{z}})$ and $P(\hat{\mathbf{z}})$ of equation (18.2) read at the fit, with the spreads $\sqrt{\mathbf{g}^\mathsf{T}\mathbf{\Sigma}\mathbf{g}}$ of the delta method; their gradients, in the columns of equation (18.1), are

$$\mathbf{g}_{T_1} = \big(-Q_1 \mathrm{e}^{-g_1},\ 0,\ \mathrm{e}^{-g_1},\ 0,\ 1\big)^\mathsf{T}, \qquad \mathbf{g}_{T_2} = \big(0,\ -Q_2 \mathrm{e}^{-g_2},\ 0,\ \mathrm{e}^{-g_2},\ 1\big)^\mathsf{T},$$
$$\mathbf{g}_P = (0,\ 0,\ 1,\ 1,\ 0)^\mathsf{T}.$$

Because the state carries g_i, the conductance reported from the fit, $\hat{G}_i = \mathrm{e}^{\hat{g}_i}$, is the MAP log-conductance transformed: it is the median of the Laplace posterior of G_i, not the mode of the posterior density of G_i, which the change of variable shifts lower by the factor $\mathrm{e}^{-\Sigma_{g_i g_i}}$. The prior on the conductances is Gaussian in g_i, not in G_i; written in G_i it would be log-normal, and the objective in G_i would carry the term $\ln G_i$ of the change of variable, which the log-conductance avoids. Readings and priors enter the objective alike: a prior row is a reading of an entry made by the handbook, with the handbook's spread as its standard deviation. The value beside each spread comes from the exact Hessian (18.15) at $\hat{\mathbf{z}}$, $\mathbf{J}^\mathsf{T}\mathbf{J} + \mathbf{S}$, inverted in place of $\mathbf{J}^\mathsf{T}\mathbf{J}$.

Building blocks. The MAP state $\hat{\mathbf{z}}$ from the Gauss–Newton iteration (Table 18.3) and the Laplace covariance $\mathbf{\Sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}$ at it, equation (18.20); the gradients of T_1, T_2 and P above; and, for the spreads beside them, the exact Hessian (18.15) with the second derivatives (18.16) of the hall rows and their residuals at $\hat{\mathbf{z}}$.

The MAP state. Table 18.4 states the MAP state. Each row is one entry, with its value in the state, its value in physical units and the units; the last rows give the hall temperatures $T_i(\hat{\mathbf{z}})$, the functions (18.2) of the state evaluated at the fit, and the minimized objective. Hall 1's path removes 5.49 kW/K and hall 2's 4.06 kW/K; the loads are 30.69 and 22.23 kW, a total of 52.92 kW; the supply water is at 12.11°C.

The entries read off with their spreads are $\hat{g}_1 = 1.7030$ with $\sqrt{\Sigma_{g_1 g_1}} = 0.1067$, $\hat{g}_2 = 1.3999$ with 0.1370, $\hat{Q}_1 = 30.69$ kW with 1.452 kW, $\hat{Q}_2 = 22.23$ kW with 2.346 kW and $\hat{T}_s = 12.111$°C with 0.280 K. The derived quantities, through the gradients of Section 18.3.1, have $\mathrm{sd}(T_1 \mid \mathbf{y}) = 0.480$ K, $\mathrm{sd}(T_2 \mid \mathbf{y}) = 0.481$ K and $\mathrm{sd}(P \mid \mathbf{y}) = 1.93$ kW: each hall temperature is known a little better than its sensor's 0.5 K, and the duty a little better than the meter's 2 kW. The (g_1, g_1) entry of the information, 136.08, is the square of the first column of $\mathbf{J}(\hat{\mathbf{z}})$, $11.18^2 = 125.0$ from the hall-1 reading and $3.333^2 = 11.1$ from the prior; the (Q_2, Q_2) entry, 0.493, is the smallest diagonal entry, because the second load is read only through the chiller meter and the hall-2

Table 18.4: The MAP state of the two-hall plant and the quantities derived from it. The conductances are $G_i = \mathrm{e}^{\hat{g}_i}$, the MAP log-conductances transformed. The spread is $\sqrt{\Sigma_{ii}}$ for an entry, in the units of the state, and $\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}$ for a derived quantity, in its physical units.

Entry or derived quantity	In the state	Physical value	Spread	Units
hall-1 log-conductance g_1	1.7030	$G_1 = 5.490$	0.1067 [–]	kW/K
hall-2 log-conductance g_2	1.3999	$G_2 = 4.055$	0.1370 [–]	kW/K
hall-1 load Q_1	30.688	30.69	1.452	kW
hall-2 load Q_2	22.231	22.23	2.346	kW
supply temperature T_s	12.1115	12.11	0.280	°C, K
hall-1 temperature $T_1(\hat{\mathbf{z}})$	–	17.701	0.480	°C, K
hall-2 temperature $T_2(\hat{\mathbf{z}})$	–	17.594	0.481	°C, K
chiller's duty $P(\hat{\mathbf{z}})$	–	52.92	1.93	kW
objective $\mathcal{L}(\hat{\mathbf{z}})$	0.3698	–	–	–

row. The largest correlation, $\Sigma_{g_2Q_2}/\sqrt{\Sigma_{g_2g_2}\Sigma_{Q_2Q_2}} = 0.691$, ties the hall-2 conductance to the hall-2 load, the two entries that the hall-2 row sees through one product. The mode of the posterior density of each conductance lies below $\hat{G}_i$ by the factor $\mathrm{e}^{-\Sigma_{g_ig_i}}$, about 1 and 2 % for the two halls.

The values beside them come from the exact Hessian: inverted, $\mathbf{J}^{\mathsf{T}}\mathbf{J} + \mathbf{S}$ gives the spreads 0.1063 and 0.1368 on g_1 and g_2, 1.448 and 2.318 kW on the loads and 0.280 K on T_s, and 0.478, 0.485 K and 1.91 kW on T_1, T_2 and P. The largest gap is on Q_2, 0.029 kW or 1.2 % narrower than the reported 2.346 kW; every other spread moves by less than 1 %. The term $\mathbf{S}$ is the residuals times the curvatures of the hall rows, and both factors are small at this fit.

18.3.2 The misfit of each row and of the whole fit

Engineering question. Do the readings agree with the physics and with each other? The question asks whether any row, and whether the fit as a whole, departs from the others by more than the errors of the readings and the spreads of the priors allow.

General form. An agreement question is a prediction under a threshold taken of a future observation. The future observation is a replicated record $\mathbf{y}_{\mathrm{rep}}$, drawn afresh from the model, and the event is that its misfit statistic X^2 is at least the one observed; its probability is the tail of the predictive distribution (2.13) of the statistic, compared with a level as in equation (2.12),

$$\Pr\big(X^2(\mathbf{y}_{\mathrm{rep}}) \geq X^2(\mathbf{y})\big) = \int \mathbf{1}\big[X^2(\mathbf{y}_{\mathrm{rep}}) \geq X^2(\mathbf{y})\big]\, p(\mathbf{y}_{\mathrm{rep}})\, \mathrm{d}\mathbf{y}_{\mathrm{rep}}, \tag{18.21}$$

a small value saying that the observed record misfits more than the model's own records do.

Specialized form. The question is equation (18.21) with the statistic $X^2 = 2\mathcal{L}(\hat{\mathbf{z}})$ for the whole fit and the normalized residual of a row for a single row, both evaluated at the MAP state of step 3, and $\mathbf{y}_{\mathrm{rep}}$ the eight rows redrawn from the model with the Jacobian $\mathbf{J}(\hat{\mathbf{z}})$ and the covariance $\boldsymbol{\Sigma}$ of equation (18.20). Two modeling choices enter here. The distribution of the misfit is derived for rows linear in the state: this gives it in closed form and makes it independent of the true state, and it becomes an approximation for the curved hall rows; the value beside it comes from refitting records drawn from the fitted model, which keeps the curvature. A prior row is treated as a reading of its entry, with the true state regarded as drawn from the prior: this lets the priors count among the redundant rows, and it misleads if the handbook spread is not a spread of real plants about the handbook value, in which case a large prior residual is no

evidence against a reading. The statistic is the minimized objective, $X^2 = 2\mathcal{L}(\hat{\mathbf{z}})$, for the whole fit, and the normalized residual of each row for a single row.

Where the redundancy of the two halls lies follows from the count of rows against entries and from the pattern (18.4), before any number is computed. The five readings face five entries, and the loads carry no prior: an error in a meter moves a load, the moved load moves a conductance through its hall row, and the rows that resist the chain are chiefly the priors on the conductances and on the supply temperature, with a share through the heat meter. The redundancy of this plant therefore lies mostly in the priors, and a reading is checked chiefly through the prior on the entry it moves; a small residual on a reading is then weak evidence that the reading is right, and the rows tied by one chain, a hall reading, the heat meter and the prior on that hall's conductance, share one normalized residual.

The linear model has normalized rows $\mathbf{r}(\mathbf{z}) = \mathbf{J}\mathbf{z} - \mathbf{b}$, with a constant $n_r \times n_z$ Jacobian $\mathbf{J}$ of full column rank and a vector $\mathbf{b}$ that holds the readings and prior means divided by their standard deviations. Under the model there is a true state $\mathbf{z}^{\text{true}}$ at which every normalized row is a standard normal error, $\zeta_k = \varepsilon_k/\sigma_k$ for a reading row,

$$\mathbf{b} = \mathbf{J}\mathbf{z}^{\text{true}} + \boldsymbol{\zeta}, \qquad \boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{n_r}) : \tag{18.22}$$

for a reading row this is its error model; for a prior row it is the statement that the true entry lies about the prior mean with the prior's spread, which holds when the true state is regarded as drawn from the prior. Entries without a prior row, the loads here, may take any fixed value. A replicated record is a fresh draw of $\boldsymbol{\zeta}$.

The residual at the fit. The minimizer of $\frac{1}{2}\|\mathbf{J}\mathbf{z} - \mathbf{b}\|^2$ solves the normal equations $\mathbf{J}^{\mathsf{T}}\mathbf{J}\,\hat{\mathbf{z}} = \mathbf{J}^{\mathsf{T}}\mathbf{b}$, so

$$\hat{\mathbf{z}} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}\mathbf{J}^{\mathsf{T}}\mathbf{b}, \qquad \mathbf{r}(\hat{\mathbf{z}}) = \mathbf{J}\hat{\mathbf{z}} - \mathbf{b} = -(\mathbf{I} - \mathbf{K})\,\mathbf{b}, \qquad \mathbf{K} = \mathbf{J}(\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}\mathbf{J}^{\mathsf{T}} = \mathbf{J}\boldsymbol{\Sigma}\mathbf{J}^{\mathsf{T}}.$$

The matrix $\mathbf{K}$, the hat matrix of least squares (written $\mathbf{K}$ here because P is the chiller's duty), projects onto the column space of $\mathbf{J}$: it is symmetric, it satisfies $\mathbf{K}^2 = \mathbf{K}$, and $\mathbf{K}\mathbf{J} = \mathbf{J}$. Substituting equation (18.22), the term in $\mathbf{z}^{\text{true}}$ vanishes because $(\mathbf{I} - \mathbf{K})\mathbf{J} = \mathbf{0}$:

$$\mathbf{r}(\hat{\mathbf{z}}) = -(\mathbf{I} - \mathbf{K})\,\boldsymbol{\zeta}, \qquad 2\mathcal{L}(\hat{\mathbf{z}}) = \mathbf{r}(\hat{\mathbf{z}})^{\mathsf{T}}\mathbf{r}(\hat{\mathbf{z}}) = \boldsymbol{\zeta}^{\mathsf{T}}(\mathbf{I} - \mathbf{K})\,\boldsymbol{\zeta}, \tag{18.23}$$

using $(\mathbf{I} - \mathbf{K})^{\mathsf{T}}(\mathbf{I} - \mathbf{K}) = \mathbf{I} - \mathbf{K}$. The residuals at the fit do not depend on the true state; they are the part of the errors that the columns of the Jacobian cannot absorb.

The count of degrees of freedom. A symmetric matrix with $\mathbf{M}^2 = \mathbf{M}$ has eigenvalues 0 or 1, and their number of ones is its trace. The trace of $\mathbf{K}$ is

$$\operatorname{tr}\mathbf{K} = \operatorname{tr}\big[(\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}\mathbf{J}^{\mathsf{T}}\mathbf{J}\big] = \operatorname{tr}\mathbf{I}_{n_z} = n_z,$$

by the cyclic property of the trace, so $\mathbf{I} - \mathbf{K}$ has $\nu = n_r - n_z$ eigenvalues equal to one and n_z equal to zero. Write $\mathbf{I} - \mathbf{K} = \mathbf{U}\mathbf{D}\mathbf{U}^{\mathsf{T}}$ with $\mathbf{U}$ orthogonal and $\mathbf{D}$ diagonal with those ones and zeros, and rotate the errors, $\mathbf{u} = \mathbf{U}^{\mathsf{T}}\boldsymbol{\zeta}$. An orthogonal rotation of independent standard normal variables gives independent standard normal variables, since $\operatorname{Cov}(\mathbf{u}) = \mathbf{U}^{\mathsf{T}}\mathbf{I}\mathbf{U} = \mathbf{I}$, and therefore

$$2\mathcal{L}(\hat{\mathbf{z}}) = \mathbf{u}^{\mathsf{T}}\mathbf{D}\,\mathbf{u} = \sum_{j=1}^{\nu} u_j^2 \sim \chi^2_\nu, \qquad \mathbb{E}[2\mathcal{L}(\hat{\mathbf{z}})] = \nu, \qquad \operatorname{Var}[2\mathcal{L}(\hat{\mathbf{z}})] = 2\nu, \tag{18.24}$$

the mean and variance of χ^2_ν following from $\mathbb{E}[u_j^2] = 1$ and $\operatorname{Var}(u_j^2) = 2$. The distribution does not depend on the true state, so the tail (18.21) of the whole fit is the same whatever predictive of the state the replicated record is averaged over: it is $\Pr(\chi^2_\nu > 2\mathcal{L}(\hat{\mathbf{z}}))$, and a small value says that the misfit is larger than correct rows with the stated standard deviations would produce.

The redundancy number of each row. The diagonal entry $1-K_{kk}$ of $\mathbf{I}-\mathbf{K}$ is the variance of the k-th standardized residual at the fit, by equation (18.23), and so the expectation of $r_k(\hat{\mathbf{z}})^2$; state-estimation practice calls it the *redundancy number* of the row. The redundancy numbers sum to the trace, $\sum_k (1-K_{kk}) = \nu$, so they say where the redundancy of the plant lies. A row with a redundancy number near zero is fitted almost exactly whatever its error, because the state can follow it; a row with a number near one is checked by the others. Dividing each standardized residual by the root of its redundancy number gives a variable of unit variance under the model,

$$r_k^{\mathrm{N}} = \frac{r_k(\hat{\mathbf{z}})}{\sqrt{1-K_{kk}}}, \tag{18.25}$$

the *normalized residual* of state estimation, the statistic of bad-data detection and of the tail (18.21) for a single row: under the model it is standard normal, so a row disagrees when $|r_k^{\mathrm{N}}|$ lies in the tails of $\mathcal{N}(0,1)$. For a prior row the redundancy number has a direct reading. The row of $\mathbf{J}$ of a prior on entry i is $1/\sigma_i$ in column i and zero elsewhere, so $K_{kk} = \Sigma_{ii}/\sigma_i^2$ with $\mathbf{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$, and

$$1 - K_{kk} = 1 - \frac{\Sigma_{ii}}{\sigma_i^2}\,: \tag{18.26}$$

the fraction of the prior variance of the entry that the other rows remove. A prior that the readings overrule is well checked; a prior on an entry the readings cannot reach has a redundancy number near zero and checks nothing.

The tail for three degrees of freedom. For $\nu = 3$ the tail has a closed form. The density of χ_3^2 is $\sqrt{x}\,\mathrm{e}^{-x/2}/\sqrt{2\pi}$; integrating by parts with $u = \sqrt{t}$ and $\mathrm{d}v = \mathrm{e}^{-t/2}\mathrm{d}t$,

$$\Pr(\chi_3^2 > x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \sqrt{t}\,\mathrm{e}^{-t/2}\,\mathrm{d}t = \frac{2\sqrt{x}\,\mathrm{e}^{-x/2}}{\sqrt{2\pi}} + \frac{1}{\sqrt{2\pi}} \int_x^\infty \frac{\mathrm{e}^{-t/2}}{\sqrt{t}}\,\mathrm{d}t,$$

and the substitution $t = s^2$ turns the last integral into $2\int_{\sqrt{x}}^\infty \mathrm{e}^{-s^2/2}\mathrm{d}s$, so that

$$\Pr(\chi_3^2 > x) = \sqrt{\frac{2x}{\pi}}\,\mathrm{e}^{-x/2} + 2\big[1 - \Phi(\sqrt{x})\big]. \tag{18.27}$$

The linearization caveat. For the two halls the hat matrix and the redundancy numbers are evaluated with the Jacobian at $\hat{\mathbf{z}}$, and the chi-square result is approximate: it holds when the rows are close to linear over the region in which the fit moves when the errors are redrawn, the same condition under which the Laplace approximation of Chapter 19 holds.

Building blocks. The residuals $r_k(\hat{\mathbf{z}})$ of the fit, the Jacobian $\mathbf{J}(\hat{\mathbf{z}})$ and the covariance $\mathbf{\Sigma}$ (Section 18.2.1); the hat matrix $\mathbf{K} = \mathbf{J}\mathbf{\Sigma}\mathbf{J}^{\mathsf{T}}$, its diagonal, the redundancy numbers and the normalized residuals (18.25); $2\mathcal{L}(\hat{\mathbf{z}})$ with $\nu = 3$ of equation (18.3); the closed-form tail (18.27) with Φ; and, for the tail beside it, the rows redrawn from the model at $\hat{\mathbf{z}}$ and refitted by the Gauss–Newton iteration. The residuals, the redundancy numbers and the tail follow, for the observed record and for a record with a faulty hall-2 sensor.

The residuals of the rows. Table 18.5 lists the eight rows at the MAP state. Each line gives the row, the value the fitted state predicts for it, the reading or prior value, the standardized residual $r_k(\hat{\mathbf{z}})$, its redundancy number $1-K_{kk}$, evaluated with the Jacobian at $\hat{\mathbf{z}}$, and the normalized residual r_k^{N} of equation (18.25). The squared standardized residuals sum to $2\mathcal{L}(\hat{\mathbf{z}}) = 0.7396$, and the redundancy numbers to 3.000, the redundancy ν.

Figure 18.1 draws the residuals as bars. The horizontal axis lists the eight rows, the five readings in blue and the three priors in gold; the vertical axis is the standardized residual in

Table 18.5: The eight rows at the MAP state: the fitted value, the reading or prior value, the standardized residual, the redundancy number and the normalized residual. Temperatures in °C, loads in kW, log-conductances dimensionless.

k	Row	Fitted value	Reading or prior	$r_k(\hat{\mathbf{z}})$	$1 - K_{kk}$	r_k^{N}
1	hall-1 temperature	17.701	17.655	0.093	0.078	0.33
2	hall-2 temperature	17.594	17.700	−0.212	0.073	−0.78
3	supply temperature	12.111	12.10	0.038	0.129	0.11
4	chiller heat meter	52.92	52.5	0.209	0.071	0.78
5	hall-1 load meter	30.69	31.0	−0.208	0.063	−0.83
6	prior on g_1	1.703	1.609	0.312	0.874	0.33
7	prior on g_2	1.400	1.609	−0.698	0.792	−0.78
8	prior on T_s	12.111	12.0	0.111	0.922	0.12
	sums			0.7396 (squares)	3.000	

units of each row's standard deviation, and the dashed lines mark ±1. No row lies beyond one standard deviation. The largest misfit is the prior on the hall-2 conductance, −0.70: the fitted conductance, 4.06 kW/K, is 19 % below the handbook's 5, and the prior's spread of 0.3 in the logarithm allows that comfortably. The readings are all within 0.21 of their predictions.

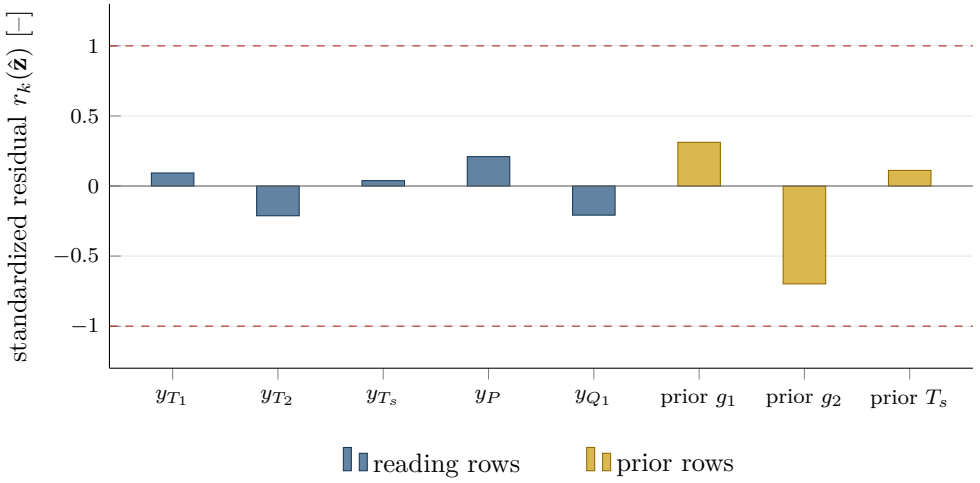


Figure 18.1: The standardized residuals of the eight rows at the MAP state, Table 18.5: readings in blue, priors in gold. The dashed lines mark ±1 standard deviation of each row.

The redundancy numbers of Table 18.5 temper that reading. Every reading row has a number between 0.06 and 0.13, so a reading's standardized residual is a small fraction of its error, and the priors hold the redundancy, as the count of Section 18.3.2 said they would. Their numbers follow directly from equation (18.26): with the posterior variances $\Sigma_{g_1 g_1} = 0.01138$, $\Sigma_{g_2 g_2} = 0.01876$ and $\Sigma_{T_s T_s} = 0.07843$ of equation (18.20), $1 - 0.01138/0.09 = 0.874$, $1 - 0.01876/0.09 = 0.792$ and $1 - 0.07843/1 = 0.922$, the fractions of the prior variances that the readings remove. Together the priors hold 2.59 of the three degrees of freedom, so they are expected to carry $2.59/3 = 86\,\%$ of $2\mathcal{L}(\hat{\mathbf{z}})$, and they carry 0.597 of the 0.740, 81 %. The normalized residuals of the last column put every row on the same footing; all lie within 0.83 of zero, and the rows tied by one chain, the hall-2 reading, the chiller meter and the hall-2 prior, share the same magnitude, 0.78. The hall-2 reading that the handbook comparison put 2.6 standard deviations high misses the fit by 0.21: the fit has moved the hall-2 conductance, the entry on which the readings and the handbook disagree most, to 4.06 kW/K, and the prior on it carries the largest residual. Section 18.3.2

shows what the same rows do when the hall-2 sensor is in fact faulty.

The chi-square of the redundant rows. By equation (18.24), if the rows and their standard deviations are correct, $2\mathcal{L}(\hat{\mathbf{z}})$ is approximately χ^2_3, of mean 3 and variance 6. The fit gives $2\mathcal{L}(\hat{\mathbf{z}}) = 0.7396$, of which 0.142 comes from the five reading rows and 0.597 from the three prior rows. In the closed form (18.27), at $x = 0.7396$, $\sqrt{x} = 0.8600$, $\Phi(0.8600) = 0.8051$ and $\mathrm{e}^{-0.3698} = 0.6909$:

$$\Pr(\chi^2_3 > 0.7396) = 0.6862 \times 0.6909 + 2 \times 0.1949 = 0.4741 + 0.3898 = 0.864.$$

The lower tail is $\Pr(\chi^2_3 < 0.7396) = 0.136$. The value beside it comes from records drawn from the model: of 4000 sets of the eight rows drawn at $\hat{\mathbf{z}}$, the readings from their error distributions and the prior means from their spreads, each refitted by Gauss–Newton, the fraction with $2\mathcal{L} > 0.740$ is 0.858, 0.006 below the chi-square tail and within the sampling error of the draws. The curved hall rows do not move the reference for a record without a fault.

Figure 18.2 draws the reference distribution. The horizontal axis is the value of $2\mathcal{L}(\hat{\mathbf{z}})$ and the vertical axis the density of χ^2_3; the vertical red line marks the observed 0.74, the shaded area to its right is the upper tail, and the dotted line marks the 95 % point, 7.81. The observed value lies left of the bulk of the distribution: 86 % of record sets drawn from a correct model would show a larger misfit. The readings agree with the model, with each other and with the priors. The lower tail, 0.136, is not small enough to suggest the opposite failure, standard deviations stated too large, which would make every misfit look small.

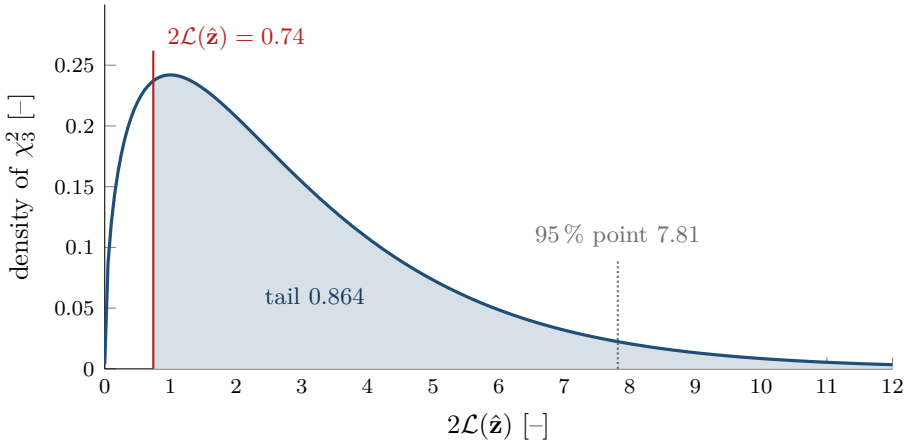


Figure 18.2: The chi-square density with three degrees of freedom, $\sqrt{x}\,\mathrm{e}^{-x/2}/\sqrt{2\pi}$, the reference distribution of $2\mathcal{L}(\hat{\mathbf{z}})$ for the two-hall fit. Red line: the observed value 0.740; shaded: the upper tail $\Pr(\chi^2_3 > 0.740) = 0.864$, equation (18.27); dotted: the 95 % point.

A faulty hall-2 sensor. The same residuals and tail are computed for a record with a known fault. The hall-2 sensor is made to read 22.0°C instead of 17.70°C, 4.3 K or 8.6 of its standard deviations high, the other readings and priors unchanged, and the plant is refitted. The fit moves the hall-2 conductance from 4.055 to 2.452 kW/K and its load from 22.23 to 23.66 kW, with $T_2(\hat{\mathbf{z}}) = 21.79$°C. The misfit is $2\mathcal{L}(\hat{\mathbf{z}}) = 6.687$, and by equation (18.27) $\Pr(\chi^2_3 > 6.687) = 0.083$: the check does not reject at the 5 % level. The standardized residuals are −0.41 on the faulty reading, −2.38 on the prior of g_2, 0.67 on the chiller meter and −0.55 on the hall-1 meter: the reading's error is absorbed by g_2, the entry it moves, and appears on the prior that checks that entry, which holds 5.64 of the 6.69. The faulty reading's redundancy number at this fit is 0.0256, below the 0.073 of the clean fit, because the larger hall-2 rise steepens the g_2 column of its row; its normalized residual is $-0.410/\sqrt{0.0256} = -2.56$, tied in magnitude with those

of the prior of g_2 (-2.56) and the chiller meter ($+2.56$), the rows of one chain. The residuals point at hall 2; the normalized residuals flag the chain but cannot isolate the reading, and they do not separate a faulty sensor from a degraded path.

A second hall-2 sensor of the same standard deviation, reading the correct 17.70°C, adds a ninth row, the same function of the state as row 2, and $\nu = 4$. Refitted with the faulty first sensor, the plant has $\hat{G}_2 = 3.029\,\mathrm{kW/K}$, $\hat{Q}_2 = 23.08\,\mathrm{kW}$ and $T_2(\hat{\mathbf{z}}) = 19.76°\mathrm{C}$, between the two readings. The standardized residuals of the two sensors are -4.48 and $+4.12$, each with redundancy number 0.51, so their normalized residuals are -6.28 and $+5.76$, and the prior of g_2 falls to -1.67; $2\mathcal{L}(\hat{\mathbf{z}}) = 40.4$ on four degrees of freedom, a tail of 3.6×10^{-8}. The fault is detected, and the two sensors on one balance are named as disagreeing; the faulty one carries the larger normalized residual, but the pair alone does not prove which is wrong. With both sensors correct at 17.70°C the misfit is $2\mathcal{L}(\hat{\mathbf{z}}) = 0.763$ on four degrees of freedom, tail 0.94.

18.3.3 The physical admissibility of the estimate

Engineering question. Is the estimate physically admissible? The model assumes that heat flows from each hall through a path of positive conductance into chilled water colder than the hall, carried by positive loads, and the question asks whether the fitted state satisfies those conditions.

General form. A requirement on several quantities at once is the threshold form (2.12) with the indicator of an intersection,

$$\Pr(f_1(\mathbf{z}) \geq f_1^*, \ldots, f_m(\mathbf{z}) \geq f_m^* \mid \mathbf{y}) = \int \mathbf{1}\left[f_1(\mathbf{z}) \geq f_1^*, \ldots, f_m(\mathbf{z}) \geq f_m^*\right] p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}.$$

Specialized form. The question is equation (2.12), with the indicator of the intersection above, for the state $\mathbf{z} = (g_1, g_2, Q_1, Q_2, T_s)^{\mathsf{T}}$, the derived quantities $f(\mathbf{z})$ the conductances G_i, the loads Q_i and the rises $T_i(\mathbf{z}) - T_s$, the thresholds $f^* = 0$, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of equation (18.20). One modeling choice enters here: the conductances are carried on the logarithmic scale and the loads on their natural scale. The first makes a positive conductance a property of every state; the second leaves a positive load to the evidence, so that a fit pulled far from the readings could return a negative load, which the values below look for.

For the two halls the events are $G_i > 0$, $Q_i > 0$ and $T_i(\mathbf{z}) > T_s$, $i = 1, 2$. The first holds at every state, since $G_i = \mathrm{e}^{g_i}$. The third follows from the second, because the rise $T_i(\mathbf{z}) - T_s = Q_i\,\mathrm{e}^{-g_i}$ is positive exactly when the load is; it is listed because it is the one an operator reads first, and a fit that violated it would point to a reading assigned to the wrong hall or to a sign error in a meter. The event therefore reduces to the two loads, and under the Laplace posterior each is a normal tail read from an entry and its spread,

$$\Pr(Q_i > 0 \mid \mathbf{y}) \approx \Phi\left(\frac{\hat{Q}_i}{\sqrt{\Sigma_{Q_iQ_i}}}\right), \qquad i = 1, 2, \tag{18.28}$$

so that the distance of each fitted load from zero, in posterior standard deviations, states the answer. Two comparisons of the same kind are made at the fit, the fitted total load against the chiller meter and the fitted supply temperature against the set point, each in the standard deviation of its row. The ratio of the two conductances has no row that could contradict it, so its check is one of plausibility, against the handbook's own spread about a common value: a ratio within that spread is plausible, and a ratio of ten would point to an error in the model or the readings.

Building blocks. The fitted entries $\hat{G}_i$, $\hat{Q}_i$, $\hat{T}_s$ and the hall temperatures $T_i(\hat{\mathbf{z}})$ (Table 18.4); the spreads $\sqrt{\Sigma_{Q_iQ_i}}$ of equation (18.20); the normal distribution Φ. No exact route is computed: the tails are too small for draws to reach.

The admissibility of the fitted state. Table 18.6 checks each condition at the MAP state. Each row is one condition, the fitted values that test it and the verdict.

Table 18.6: Physical admissibility of the MAP state.

Condition	Fitted values	Verdict
positive conductances, $G_i > 0$	5.490 and 4.055 kW/K	holds by the log scale
positive loads, $Q_i > 0$	30.69 and 22.23 kW	holds
halls above the supply, $T_i > T_s$	17.701 and 17.594°C against 12.111°C; rises 5.590 and 5.482 K	holds
total load against the chiller meter	52.92 kW against 52.5 ± 2 kW	within $0.21\,\sigma_P$
supply temperature against the set point	12.11°C against 12 ± 1°C	within $0.11\,\sigma_0$
conductance ratio	$G_1/G_2 = 1.354$	plausible

The first condition holds by the log scale. The loads carry no such guarantee, and their positivity is a finding: with the posterior spreads $\sqrt{\Sigma_{Q_1Q_1}} = 1.45$ kW and $\sqrt{\Sigma_{Q_2Q_2}} = 2.35$ kW of equation (18.20), the fitted loads lie about 21 and 9 standard deviations above zero. By equation (18.28),

$$\Pr(Q_1 \le 0 \mid \mathbf{y}) \approx \Phi\left(-\frac{30.688}{1.452}\right) = \Phi(-21.13) = 2.0 \times 10^{-99},$$
$$\Pr(Q_2 \le 0 \mid \mathbf{y}) \approx \Phi\left(-\frac{22.231}{2.346}\right) = \Phi(-9.475) = 1.3 \times 10^{-21},$$

so the probability of the joint event is at least one less the sum of the two tails, $1 - 1.3 \times 10^{-21}$: the estimate is admissible with posterior probability one to every digit an engineer would print, and the rises of 5.590 and 5.482 K are positive with the loads. The conductance ratio of 1.35 says that hall 2 removes about a quarter less heat per kelvin than hall 1, with nearly the same temperature rise at a smaller load; a ratio of this size lies within the handbook's own spread of 35 % about a common value.

18.4 The checks (step 5)

Each check sets the reported fit beside a calculation that does not share one of its assumptions.

18.4.1 A second starting point

A single run of an iteration does not show that its result is the minimum of the objective rather than a point where the iteration happened to stop. A second run from a distant start is the simplest check. The start

$$\mathbf{z}_B^{(0)} = (\ln 2,\ \ln 2,\ 40,\ 40,\ 12)$$

assumes conductances of 2 kW/K and loads of 40 kW in each hall, which predict hall temperatures of $12 + 40/2 = 32$°C, fourteen kelvin above the readings, and a total load of 80 kW against the chiller meter's 52.5; its objective is 942.48. Table 18.7 follows the iteration, with the columns of Table 18.3.

Table 18.7: Gauss–Newton iterations for the two-hall plant from the distant start $\mathbf{z}_B^{(0)} = (\ln 2, \ln 2, 40, 40, 12)$. Columns and units as in Table 18.3.

m	G_1	G_2	Q_1	Q_2	T_s	$\mathcal{L}(\mathbf{z}^{(m)})$	$\lVert\nabla\mathcal{L}\rVert$	$\lVert\Delta\mathbf{z}^{(m)}\rVert$
0	2.000	2.000	40.000	40.000	12.000	942.48	1640	19.9
1	3.294	2.659	30.866	22.335	12.140	49.154	183	0.563
2	4.936	3.769	30.694	22.369	12.124	1.7187	21.3	0.182
3	5.466	4.039	30.689	22.235	12.112	0.3722	0.83	7.3×10^{-3}
4	5.490	4.055	30.688	22.231	12.111	0.3698	1.1×10^{-2}	1.8×10^{-4}
5	5.490	4.055	30.688	22.231	12.111	0.3698	2.1×10^{-4}	4.0×10^{-6}
6	5.490	4.055	30.688	22.231	12.111	0.3698	4.9×10^{-6}	1.1×10^{-7}

The first step repairs the loads almost entirely, to 30.87 and 22.34 kW, because the two meters are linear rows that one linearized step satisfies; the conductances, which enter through the curved hall rows, need three more steps, 3.29, 4.94, 5.47 and 5.49 kW/K for hall 1. The objective falls $942.48 \to 49.15 \to 1.719 \to 0.3722 \to 0.3698$, and the stopping rule (18.18) is met by the sixth step, whose largest relative move is 8.3×10^{-7}.

Figure 18.3 draws the objective and the gradient norm of both runs on logarithmic axes. In panel (a) the horizontal axis is the iterate m and the vertical axis the objective; the blue circles start from the prior point, the gold squares from the distant start, and the dashed line is the minimum $\mathcal{L}(\hat{\mathbf{z}}) = 0.3698$. Panel (b) draws the gradient norm on the same horizontal axis. Both starts reach the dashed line within two or four steps, and in panel (b) both settle onto parallel straight lines, the signature of a constant convergence factor, which Section 18.4.2 traces to the dropped term.

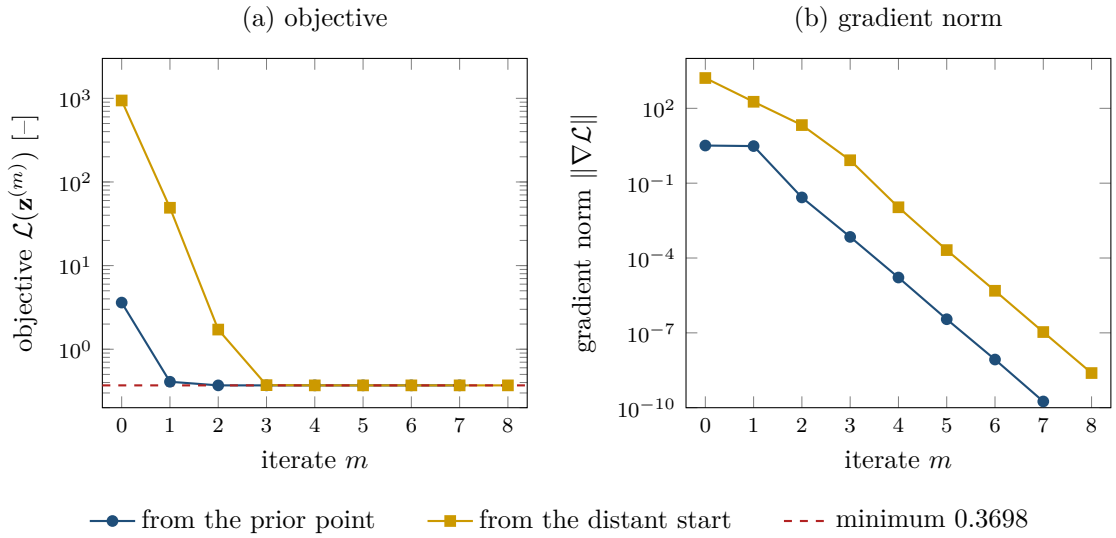


Figure 18.3: Gauss–Newton iterations for the two-hall plant from the prior point $(\ln 5, \ln 5, 28, 28, 12)$ (blue circles) and from the distant start $(\ln 2, \ln 2, 40, 40, 12)$ (gold squares), Tables 18.3 and 18.7, continued past the stopping rule (18.18) to show the rate. (a) The objective $\mathcal{L}(\mathbf{z}^{(m)})$; the dashed line is its minimum 0.3698. (b) The norm of the gradient $\mathbf{J}^{\mathsf{T}}\mathbf{r}$; after the first steps both fall by a factor of about 0.025 per step.

Figure 18.4 draws both paths over the objective. The horizontal axis is the log-conductance g_1 of hall 1 and the vertical axis its load Q_1 in kW. The gray closed curves are contours of the objective minimized over the other three entries at each point, the profile of $\mathcal{L}$ over (g_1, Q_1), at 0.5, 2, 8, 32 and 128 above the minimum, from the inside out; the innermost, a rise of $\frac{1}{2}$, is

where a Gaussian posterior would place its one-standard-deviation contour. The blue circles are the iterates from the prior point and the gold squares those from the distant start, each projected onto the plane; the red star is the MAP state. The path from the prior point reaches the innermost contour in one step. The distant path reaches the right load in its first step, the work of the two linear meter rows, and then moves along the valley in g_1 over three more steps, the work of the curved hall-1 row. The two runs end at the same point: the states at which the stopping rule halts them, $\mathbf{z}^{(4)}$ and $\mathbf{z}_B^{(6)}$, agree to 2×10^{-6} in every entry, the accuracy the rule asks for; iterated past the rule, to $m = 7$ and $m = 8$, they agree to better than 10^{-10}. A quasi-Newton minimization of the same objective, run as an independent check, returns the same state to seven digits.

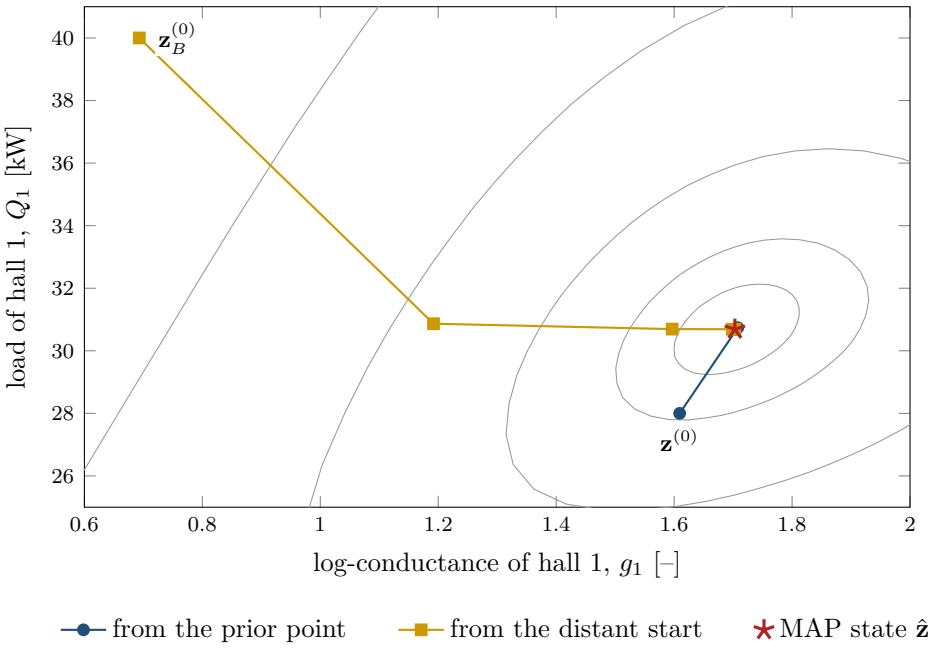


Figure 18.4: The two Gauss–Newton paths projected onto the plane of hall 1's log-conductance g_1 and load Q_1. Gray: contours of the profile objective, $\mathcal{L}$ minimized over (g_2, Q_2, T_s) at each point, at 0.5, 2, 8, 32 and 128 above the minimum 0.3698. Blue circles: iterates from $\mathbf{z}^{(0)} = (\ln 5, \ln 5, 28, 28, 12)$; gold squares: iterates from $\mathbf{z}_B^{(0)} = (\ln 2, \ln 2, 40, 40, 12)$; red star: the MAP state, $\hat{g}_1 = 1.7030$, $\hat{Q}_1 = 30.69\,\text{kW}$.

The plain iteration, without any safeguard, converged from a start whose objective was more than two thousand times its minimum. That is a property of this plant, whose only nonlinearity is the product of a load and an exponential, not a guarantee; the safeguards of Section 18.2.2 are for plants where it fails.

18.4.2 The exact Hessian against the Gauss–Newton matrix

The term $\mathbf{S}$ of equation (18.15) at $\hat{\mathbf{z}}$ has four distinct nonzero entries. With the temperature rises $T_1(\hat{\mathbf{z}}) - \hat{T}_s = 5.590\,\text{K}$ and $T_2(\hat{\mathbf{z}}) - \hat{T}_s = 5.482\,\text{K}$, the second derivatives (18.16) are $5.590/0.5 = 11.18$ and 10.96 on the diagonal and $-0.1821/0.5 = -0.3643$ and $-0.2466/0.5 = -0.4932$ off it. Multiplied by the residuals $r_1 = 0.0930$ and $r_2 = -0.2123$,

$$S_{g_1 g_1} = 1.039, \qquad S_{g_1 Q_1} = -0.0339, \qquad S_{g_2 g_2} = -2.328, \qquad S_{g_2 Q_2} = 0.1047, \tag{18.29}$$

against the Gauss–Newton entries $(\mathbf{J}^\mathsf{T}\mathbf{J})_{g_1 g_1} = 136.08$, $(\mathbf{J}^\mathsf{T}\mathbf{J})_{g_1 Q_1} = -4.072$, $(\mathbf{J}^\mathsf{T}\mathbf{J})_{g_2 g_2} = 131.34$ and $(\mathbf{J}^\mathsf{T}\mathbf{J})_{g_2 Q_2} = -5.408$. Row by row, the bound (18.17) gives $0.0930/11.18 = 0.0083$ for hall 1

and $0.2123/10.96 = 0.019$ for hall 2. The dropped term is below 2 % of the kept one entry by entry, and its spectral norm is 1.6 % of that of $\mathbf{J}^\mathsf{T}\mathbf{J}$. The spectral radius of $(\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}\mathbf{S}$ is 0.025, which is the factor by which the gradient norm of Table 18.3 fell on each late step, 0.021 to 0.026: the linear convergence of the iteration is the size of the dropped term. Inverted, the exact Hessian $\mathbf{J}^\mathsf{T}\mathbf{J} + \mathbf{S}$ changes the spreads $\sqrt{\Sigma_{ii}}$ of Section 18.3.1 by at most 1.2 %, for Q_2; this is the justification for building the reported covariance, and that of Chapter 19, on $\mathbf{J}^\mathsf{T}\mathbf{J}$. The residuals of this fit are small and its rows are mildly curved, so both factors of each entry of $\mathbf{S}$ are small.

18.4.3 The limit of exact balances

The reduced state is the limit of vanishing tolerances of the full state of eight entries and eleven rows, and the limit is checked numerically. Fitting the full state at $\sigma_p = \sigma_{p,P} = 0.1$, 0.01 and 0.001 ($\sigma_p$ in kelvin for the halls, $\sigma_{p,P}$ in kilowatts for the header) gives a hall-2 conductance of 4.058, 4.055 and 4.055 kW/K, a spread $\sqrt{\Sigma_{g_2 g_2}}$ of 0.1378, 0.1370 and 0.1370 and a misfit $2\mathcal{L}(\hat{\mathbf{z}})$ of 0.7373, 0.7396 and 0.7396, against the reduced fit's 4.055, 0.1370 and 0.7396 (Sections 18.3.1 and 18.3.2); the five-entry block of the full covariance differs from the reduced $\mathbf{\Sigma}$ by at most 10^{-4} at $\sigma_p = 0.01$ and 10^{-6} at 0.001, shrinking as σ_p^2. By $\sigma_p = 0.001$ every number of this chapter is reproduced at its printed precision.

A tolerance of physical size is a different model, a balance that may fail by a model discrepancy, and the reduced state then carries it in the reading rows. Because T_i appears only in its own reading and its own balance, and P only in the heat meter's row and the header balance, the five-entry marginal at finite tolerances is the reduced fit with the hall readings' standard deviation $\sqrt{\sigma_T^2 + \sigma_p^2}$ in place of σ_T and the heat meter's $\sqrt{\sigma_P^2 + \sigma_{p,P}^2}$ in place of σ_P, the result of equation (4.33). At $\sigma_p = \sigma_{p,P} = 0.1$ the reduced fit with both standard deviations inflated gives the full fit's hall-2 conductance, 4.0576 kW/K, and its spread 0.1378; a hall discrepancy of 0.2 K turns 0.5 K into 0.539 K.

18.4.4 The chi-square reference against the draws

The chi-square distribution of the misfit, equation (18.24), is exact for linear rows and approximate for the curved hall rows. It is tested by drawing 4000 sets of the eight rows from the model at $\hat{\mathbf{z}}$, the readings from their error distributions and the prior means from their spreads, refitting each by Gauss–Newton and recording $2\mathcal{L}$ at its minimum. Table 18.8 collects the check. Its rows give the counts of rows and entries, the redundancy, the statistic and its split between reading and prior rows, the reference values of χ^2_3, the two tails, and the simulation.

Table 18.8: The chi-square check of the two-hall fit. The simulation draws 4000 sets of the eight rows from the model at $\hat{\mathbf{z}}$, the readings from their error distributions and the prior means from their spreads, refits each by Gauss–Newton and records $2\mathcal{L}$ at its minimum.

Quantity	Value
rows n_r, entries n_z, redundancy ν	8, 5, 3
$2\mathcal{L}(\hat{\mathbf{z}})$	0.740
from the five reading rows; from the three prior rows	0.142; 0.597
mean, variance and 95 % point of χ^2_3	3, 6, 7.81
upper tail $\Pr(\chi^2_3 > 0.740)$	0.864
lower tail $\Pr(\chi^2_3 < 0.740)$	0.136
simulation: mean and variance of $2\mathcal{L}$ at the refit	2.97, 5.87
simulation: fraction of refits with $2\mathcal{L} > 0.740$	0.858

The simulation reproduces the mean, the variance and the tail of χ^2_3 within its sampling error, about 0.006 on the fraction: 2.97 and 5.87 against 3 and 6, and 0.858 against 0.864. For records without a fault the linearization caveat of Section 18.3.2 does not bite on this plant; a faulty reading moves the fit far enough along the curved hall row for it to matter, as Sections 18.3.2 and 18.6.2 show.

18.4.5 What the priors contribute

The prior rows on the conductances carry the handbook's statement that each path conducts about 5 kW/K, within 35 %. To see how much of the MAP state that statement decides, the fit is repeated with the prior's spread σ_g halved, doubled and quadrupled, the supply prior unchanged. Table 18.9 lists the results. Each row is one value of σ_g; the columns give the fitted conductances, loads and supply temperature, $2\mathcal{L}(\hat{\mathbf{z}})$ and its χ^2_3 tail. The last row removes all three prior rows, the square system of Section 18.4.6.

Table 18.9: The MAP state against the spread σ_g of the conductance priors, with the supply prior held at $\mathcal{N}(12, 1^2)$. The last row has no prior rows. Conductances in kW/K, loads in kW, T_s in °C.

σ_g	$\hat{G}_1$	$\hat{G}_2$	$\hat{Q}_1$	$\hat{Q}_2$	$\hat{T}_s$	$2\mathcal{L}(\hat{\mathbf{z}})$	$\Pr(\chi^2_3 > 2\mathcal{L})$
0.15	5.349	4.396	30.22	23.26	12.137	1.849	0.604
0.30	5.490	4.055	30.69	22.23	12.111	0.740	0.864
0.60	5.549	3.899	30.91	21.72	12.098	0.224	0.974
1.20	5.567	3.851	30.98	21.56	12.093	0.065	0.996
no prior rows	5.581	3.839	31.00	21.50	12.100	0	–

Figure 18.5 draws the two fitted conductances against σ_g. The horizontal axis is the prior's spread on a logarithmic scale and the vertical axis the conductance in kW/K; the blue curve is hall 1 and the gold curve hall 2, with the four values of the table marked, the dashed lines are the values without priors, and the dotted line is the handbook's 5 kW/K. As the prior widens, both curves leave the handbook value and settle onto the values the readings alone imply; by $\sigma_g = 0.6$ they are within 2 % of them.

Halving the prior's spread raises the hall-2 conductance from 4.055 to 4.396 kW/K, 8.4 %, and lowers hall 1's from 5.490 to 5.349, 2.6 %; doubling it moves them by −3.9 % and +1.1 %; removing the priors, by −5.3 % and +1.7 %. The moves are small because the readings dominate the conductances through the temperature rises. The diagonal entries of $\mathbf{J}^\mathsf{T}\mathbf{J}$ in equation (18.20) split by row: of the 136.08 on g_1, the hall-1 reading's line of $\mathbf{J}$ contributes $(5.590/0.5)^2 = 125.0$ through its derivative (18.6) and the prior $1/0.3^2 = 11.1$; of the 131.34 on g_2, $(5.482/0.5)^2 = 120.2$ and 11.1. Each reading carries about 92 % of the precision on its conductance, given the other entries. The prior's pull is larger for hall 2 because the readings put that conductance further from the handbook value. The fitted loads move with the conductances, within the meters' standard deviations, and the supply temperature by at most 0.03 K.

The check moves more than the point. The statistic $2\mathcal{L}(\hat{\mathbf{z}})$ falls from 1.85 to 0.065 as σ_g grows from 0.15 to 1.2, and the tail rises from 0.60 to 0.996. Every spread passes, but the meaning of passing changes: with a loose prior the prior rows can absorb almost any disagreement, and a check whose redundancy lies mostly in the priors, as Section 18.3.2 found, becomes weak. The degrees of freedom stay three for every finite spread, and drop to zero only when the priors are removed.

18.4.6 The square system without priors

Without the three prior rows the plant has five rows and five entries, a 5×5 Jacobian of full rank: every row is satisfied exactly, at the readings-only solution of Section 17.3.1, the

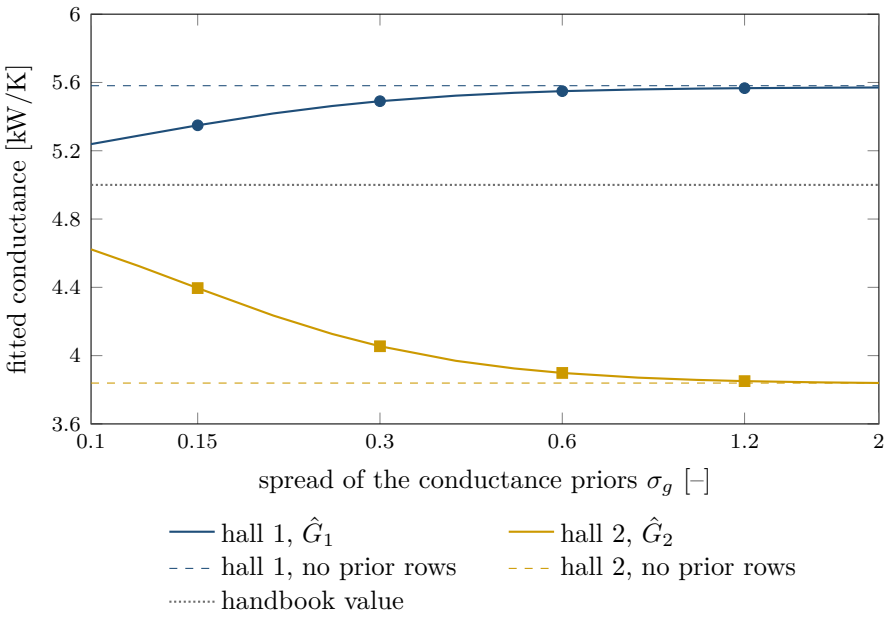


Figure 18.5: The fitted conductances of the two halls against the spread σ_g of the log-conductance priors, Table 18.9. Solid curves: MAP values (marks at $\sigma_g = 0.15$, 0.3, 0.6, 1.2); dashed: the fit without prior rows, 5.581 and 3.839 kW/K; dotted: the handbook value 5 kW/K.

values of the last line of Table 18.9, and $\mathcal{L} = 0$. The Gauss–Newton iteration from the prior point reaches that solution in four steps. The derivation of Section 18.3.2 says why the zero carries no information. With $n_r = n_z$ and a square Jacobian of full rank, the hat matrix is $\mathbf{K} = \mathbf{J}(\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}\mathbf{J}^\mathsf{T} = \mathbf{J}\mathbf{J}^{-1}(\mathbf{J}^\mathsf{T})^{-1}\mathbf{J}^\mathsf{T} = \mathbf{I}$, every redundancy number is zero, and $\mathbf{r}(\hat{\mathbf{z}}) = -(\mathbf{I} - \mathbf{K})\boldsymbol{\zeta} = \mathbf{0}$ whatever the errors: the chi-square check has no degrees of freedom and $2\mathcal{L}(\hat{\mathbf{z}}) = 0$ identically.

A reading in error is then not detected; it is absorbed into the entry it determines, the hall-2 reading of 22.0°C of Section 17.3.1 becoming a conductance of 2.172 kW/K with every residual zero. With the priors, the three redundant rows carry the check, and Section 18.3.2 showed how much of that fault they see. The same loss of redundancy has a second cost, in the spread of the estimate, which Chapter 19 derives.

Common mistake

Reading a small residual as agreement. A row of small redundancy number is fitted almost exactly whatever its error, because the state can follow it, so its standardized residual says little about the reading. Agreement is shown by the misfit of the redundant rows against its chi-square reference, and the residual of a single row is read together with its redundancy number, through the normalized residual (18.25).

18.5 The report (step 6)

The hall-2 sensor that the reading-by-reading comparison put 2.6 standard deviations high agrees with the rest of the record. Fitted together with the handbook's spread and the meters' errors, the hall-2 path has a conductance of 4.06 kW/K, 19 % below the handbook value; the hall-2 reading misses the fit by 0.21 of its standard deviation, no row misses by more than 0.70, and the misfit of the three redundant rows, 0.74 against a chi-square reference of mean three, is exceeded by 86 % of the records a correct model would produce. The suspicion moves from the sensor to the path. The same fit is a weak judge of any single reading: the readings' redundancy

numbers are 0.06 to 0.13, the priors hold 2.59 of the three degrees of freedom, and a hall-2 sensor that really reads 4.3 K high leaves the misfit at a tail of 0.083 (Section 18.3.2), a pass at the 5 % level. The answers below are given in the units in which the questions were asked.

18.5.1 The answers and their bands

Table 18.10 gives the answers. Each row is a question of Table 18.1; the second column gives the answer computed from $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$, which the chapter reports, the third the value beside it (the exact Hessian for the spreads, the drawn records for the misfit tail), and the fourth its band: the posterior spread $\sqrt{\Sigma_{ii}}$ for an entry, and for a check the statistics on which it rests.

Table 18.10: The answers to the questions of the two-hall fit. The conductances are the MAP log-conductances transformed; the spreads of the log-conductances are dimensionless.

Question	Answer (state route)	Beside it	Band
What are the conductances, the loads and the supply temperature?	$G_1 = 5.49$, $G_2 = 4.06$ kW/K; $Q_1 = 30.69$, $Q_2 = 22.23$ kW; $T_s = 12.11$°C	exact Hessian: every spread within 1.2 % (2.32 kW on Q_2)	$\sqrt{\Sigma_{ii}}$: 0.107 and 0.137 on $\hat{g}_1$, $\hat{g}_2$; 1.45 and 2.35 kW on the loads; 0.28 K on T_s; $\mathrm{corr}(g_2, Q_2) = 0.69$
Do the readings agree with the physics and with each other?	yes, for the plant as a whole; $2\mathcal{L}(\hat{\mathbf{z}}) = 0.74$ on 3 degrees of freedom, tail 0.86	drawn records: tail 0.858	no standardized residual beyond 0.70, no normalized residual beyond 0.83; redundancy numbers of the readings 0.06 to 0.13
Is the estimate physically admissible?	yes, with probability $1 - 1.3 \times 10^{-21}$	–	positive conductances; loads 21 and 9 standard deviations above zero; rises 5.59 and 5.48 K; $G_1/G_2 = 1.35$

18.5.2 The verdict on the readings

The checks support the reported point and its band. Two widely separated starts reach the same state; the exact Hessian changes no spread by more than 1.2 %; the full state reproduces the reduced fit at its printed precision by tolerances of 0.001; the chi-square reference holds against 4000 refits of drawn records; and halving or doubling the handbook spread moves the hall-2 conductance by +8.4 or −3.9 %, while the tail of the check stays above 0.6. The band is that of the Gauss–Newton curvature at the reported state; Chapter 19 reads from it which entries the readings determine, and Chapter 20 carries it to the hall temperature at a future load.

18.6 The next measurement (step 7)

The report leaves one weakness open: the hall-2 reading, with a redundancy number of 0.073, is checked chiefly through the prior on g_2, with a share through the heat meter, and the same holds in degree for every reading. The candidate measurements are sensors added on the existing balances.

18.6.1 The candidate measurements

Three candidates are considered, each a row the plant does not yet have: a second temperature sensor in hall 2, of the same standard deviation 0.5 K, on the balance of row 2; an electrical

meter on hall 2's load, of standard deviation 1.5 kW like the hall-1 meter, reading the entry Q_2 directly; and a second temperature sensor in hall 1, on the balance of row 1. Each adds one row and raises the redundancy to $\nu = 4$. Their Jacobian rows at $\hat{\mathbf{z}}$, in the columns of equation (18.1), are

$$\mathbf{J}_a = (0,\ -10.96,\ 0,\ 0.4932,\ 2), \qquad (0,\ 0,\ 0,\ 0.667,\ 0), \qquad (-11.18,\ 0,\ 0.3643,\ 0,\ 2),$$

the hall rows of $\mathbf{J}(\hat{\mathbf{z}})$ (Section 18.2.3) repeated for the two sensors, and $1/\sigma_Q$ in the column of Q_2 for the meter.

18.6.2 The preposterior covariance and the redundancy

A linear-Gaussian candidate row leaves a posterior whose covariance does not depend on its outcome, the Laplace preposterior of equation (2.17); with the candidate's row stacked under the eight,

$$\boldsymbol{\Sigma}_a = (\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a)^{-1}, \qquad \mathbf{K}_a = \mathbf{J}_+\boldsymbol{\Sigma}_a\mathbf{J}_+^{\mathsf{T}}, \qquad \mathbf{J}_+ = \begin{bmatrix}\mathbf{J}(\hat{\mathbf{z}})\\ \mathbf{J}_a\end{bmatrix}, \tag{18.30}$$

with $\mathbf{K}_a$ the hat matrix of the nine rows, whose diagonal gives the redundancy numbers after the candidate. Two quantities are read from them for the questions already asked. For the first, the width the candidate removes from a spread, $\sqrt{\Sigma_{ii}} - \sqrt{\Sigma_{a,ii}}$. For the second, what the candidate does to the check of a faulty reading. If the hall-2 reading is biased by Δy, the errors of equation (18.22) gain the vector $\mathbf{d}$, equal to $\Delta y/\sigma_T$ in the row of that reading and zero elsewhere, and equation (18.23) becomes

$$2\mathcal{L}(\hat{\mathbf{z}}) = (\boldsymbol{\zeta} + \mathbf{d})^{\mathsf{T}}(\mathbf{I} - \mathbf{K})(\boldsymbol{\zeta} + \mathbf{d}) \sim \chi^2_\nu\Big((\Delta y/\sigma_T)^2(1 - K_{22})\Big), \tag{18.31}$$

a noncentral chi-square whose noncentrality is the squared bias in standard deviations weighted by the reading's redundancy number, with mean ν plus the noncentrality. The result holds for linear rows, with $\mathbf{K}$ fixed by the Jacobian at the fit. The hall rows are curved, and a bias moves the fit along them: the hall-2 fault of Section 18.3.2 steepens the g_2 column of the faulty row and lowers that row's redundancy number from 0.073 to 0.026, so that the noncentrality read at $\hat{\mathbf{z}}$ overstates what the check sees. For the curved rows the noncentrality is read instead as the misfit of the faulted record without noise, refitted, which is 3.18 for the present plant against the linear $(\Delta y/\sigma_T)^2(1 - K_{22}) = 73.96 \times 0.073 = 5.41$. The probability that the check flags the fault at the 5 % level is the design form (2.16) with u the indicator that the check rejects: the fraction of records drawn from the model at $\hat{\mathbf{z}}$, with the bias added to the hall-2 reading, whose refitted misfit exceeds the 95 % point of χ^2_ν. It is computed from 15 000 refitted records per candidate, with a sampling error of about 0.004; the tail of equation (18.31) at the linear noncentrality is an optimistic bound on it. For the fault of Section 18.3.2, $\Delta y = 4.3$ K and $(\Delta y/\sigma_T)^2 = 73.96$.

Table 18.11 ranks the candidates. Each row is a candidate, the first the present plant; the columns give the redundancy, the spreads after the candidate of the two hall-2 entries, the largest width removed, the redundancy number of the hall-2 reading at $\hat{\mathbf{z}}$, the noncentrality of the 4.3 K fault on the curved rows, and the probability that the chi-square check flags it at the 5 % level, from the refitted records and, as the optimistic bound, from the linear noncentrality.

The two kinds of value part company. The hall-2 load meter removes the most width, nearly half of the spread of Q_2 and nearly a fifth of that of g_2, from 0.1370 to 0.1114, because it reads directly the one entry that no reading reached alone; but it leaves the hall-2 reading's redundancy number at 0.080, and the 4.3 K fault would be flagged with probability 0.28, no better than the present 0.29. The second hall-2 sensor removes little width, 7.5 % of the spread of g_2, but gives the hall-2 reading a partner on the same balance: its redundancy number rises from 0.073 to 0.519, the noncentrality of the fault from 3.18 to 38.0, and the fault is flagged

Table 18.11: Candidate measurements for the two-hall plant. Spreads $\sqrt{\Sigma_{a,ii}}$ of g_2 (dimensionless) and Q_2 (kW) from the preposterior (18.30) at $\hat{\mathbf{z}}$; the largest width removed from a present spread; the redundancy number of the hall-2 reading at $\hat{\mathbf{z}}$; for a 4.3 K bias on that reading, the noncentrality on the curved rows (the refitted misfit of the faulted record without noise) and the probability that the check rejects at 5 %, from 15 000 refitted records and, as an optimistic bound, from equation (18.31).

Candidate	ν	g_2 [–]	Q_2 [kW]	Largest width removed	Redundancy, hall 2 [–]	Noncentrality [–]	Flagged, refits [–]	Flagged, linear bound [–]
none (present)	3	0.1370	2.346	–	0.073	3.18	0.29	0.47
second hall-2 sensor	4	0.1267	2.341	g_2: 0.0103, 7.5 %	0.519	38.0	1.00	1.00
hall-2 load meter	4	0.1114	1.264	Q_2: 1.082 kW, 46 %	0.080	3.46	0.28	0.46
second hall-1 sensor	4	0.1370	2.346	g_1: 0.0161, 15 %	0.073	3.18	0.26	0.43

in all but about three of every ten thousand records. The second hall-1 sensor does for hall 1 what the second hall-2 sensor does for hall 2, and nothing for the hall-2 reading; with one more degree of freedom over which to spread the same noncentrality, it flags the hall-2 fault with probability 0.26. The linear bound overstates every probability but the second hall-2 sensor's, where both are one to two decimals, and it ranks the candidates in the same order.

18.6.3 The recommendation

The next measurement is the second temperature sensor in hall 2. Hall 2 is the hall whose conductance the readings and the handbook dispute most (Section 18.3.2), and its reading is the one the present fit can least check. With both sensors correct, the record of Section 18.3.2 passes with a tail of 0.94; with the first of them reading 22.0°C, it fails with a tail of 3.6×10^{-8}, where the single sensor gave 0.083, and over records drawn with that fault the check flags it in all but a few, against about three in ten today. Were the purpose the spread of the second load, the hall-2 load meter would be the measurement; the ranking of candidate measurements by the width each removes from an engineering output is the subject of Chapter 20.

18.6.4 What the analysis showed

The first reading of the record compared each reading with the handbook and found the hall-2 sensor 2.6 standard deviations high, a sensor to recalibrate; the analysis finds that the readings agree with each other and with the physics, and that it is the hall-2 path, not its sensor, that departs from the handbook. Fitted together, the five readings and three priors put the hall-2 conductance at 4.06 kW/K, 19 % below the handbook value and within its spread, and the hall-2 reading within a quarter of its standard deviation of the fit; the misfit of 0.74 on three degrees of freedom is what a correct model produces. That agreement is weak evidence about any one reading. It rests mostly on the priors, which hold 2.59 of the three degrees of freedom, and a reading's own residual shows only a small fraction of its error, so a hall-2 sensor that really reads 4.3 K high would usually still pass. On site, the work order for the hall-2 sensor is not justified by these readings; the hall-2 path is the one to watch, and the hall-2 sensor is the reading to give a partner, since a second sensor on that balance turns a fault the present check misses about seven times in ten into one it flags almost always.

18.6.5 Limits of the model

The MAP state and its checks rest on assumptions that the checks themselves cannot all test. The rows are taken as independent and Gaussian with known standard deviations. A standard deviation misstated by a common factor for every row scales every misfit, and the chi-square check sees it through either tail. A single row whose true standard deviation is s_k times the stated one moves the expectation of $2\mathcal{L}(\hat{\mathbf{z}})$ by $(s_k^2 - 1)(1 - K_{kk})$, its error variance weighted by its redundancy number in equation (18.23): on a row of high number, such as a prior here, the check sees it; on a row of low number, such as any one reading here, it barely moves the statistic. A bias common to two readings, such as a shared calibration of the two hall sensors, violates the independence of equation (18.22) and can pass the check while shifting the state. The balances are steady and taken in the limit of vanishing tolerances; a hall that has not reached steady state, or a discrepancy of physical size (Section 18.4.3), enters the readings' error and is checked only through the redundant rows. The chi-square distribution is exact for linear rows and approximate otherwise; the simulation of Table 18.8 confirmed it for this plant on records without a fault, a faulty reading moves the fit far enough along the curved rows to lower what the check sees (Section 18.6.2), and a plant with strongly curved rows needs the same tests.

The checks also do not locate a fault. A large $2\mathcal{L}(\hat{\mathbf{z}})$ says that some row disagrees with the rest, not which one; the residuals point, but, as Section 18.3.2 showed, a faulty reading with a small redundancy number is absorbed by its entry and shows its error on another row, and the rows of one chain share one normalized residual. Locating a faulty reading is a diagnosis question, answered by comparing models in which each reading in turn is allowed a bias, the subject of a later chapter on bad data and sensor bias. The companion volume *Reliability Engineering* treats state estimation and bad-data detection in its own terms. Finally, two starting points that reach the same state show that both paths end in the same valley, not that the objective has no other; for a plant with more strongly curved rows, several starts spread over the admissible region are the practical check.

Common mistake

Judging convergence by an absolute tolerance on the gradient or the step. The gradient norm of Table 18.3 barely moved on the first step while the objective fell ninefold, and an absolute step tolerance treats a change of 10^{-6} in a load of 30 kW and in a log-conductance near 1.7 as the same accuracy. The stopping rule is a relative step per entry, equation (18.18), confirmed by the objective no longer decreasing.

The next chapter derives the Laplace covariance for a state of many entries and reads from it which entries of the state the readings determine at all.

18.7 What to retain

Definition 18.1 (The MAP state of a plant and its checks)**.** For independent Gaussian rows written in normalized form $r_k(\mathbf{z})$, readings and priors alike, the negative log posterior is $\mathcal{L}(\mathbf{z}) = \frac{1}{2}\sum_k r_k(\mathbf{z})^2$ plus a constant, and the MAP state $\hat{\mathbf{z}}$ minimizes it, solving $\mathbf{J}^\mathsf{T}\mathbf{r} = \mathbf{0}$. On the reduced state the physical rows are the limit $\sigma_p \to 0$ of rows with a model tolerance: the rows of the evidence are seen through $\mathbf{J} = \mathbf{J}_\mathrm{e}\mathbf{N}$, the quantities the balances determine are functions of the state, and a tolerance of physical size adds σ_p^2 to the variance of the readings those quantities feed. The Gauss–Newton iteration solves the linearized problem $(\mathbf{J}^\mathsf{T}\mathbf{J})\Delta\mathbf{z} = -\mathbf{J}^\mathsf{T}\mathbf{r}$ at each iterate; it drops the term $\mathbf{S} = \sum_k r_k \nabla^2 r_k$ of the exact Hessian, which sets its linear rate of convergence but not its limit, and $\boldsymbol{\Sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}$ at $\hat{\mathbf{z}}$ is the Laplace covariance. For linear rows with correct standard deviations, $2\mathcal{L}(\hat{\mathbf{z}}) \sim \chi^2_\nu$ with $\nu = n_r - n_z$, and the redundancy numbers

$1 - K_{kk}$ of the rows sum to ν; for a prior row the number is $1 - \Sigma_{ii}/\sigma_i^2$. For linear rows a bias Δy_k on row k adds the noncentrality $(\Delta y_k/\sigma_k)^2(1 - K_{kk})$ to the misfit; on curved rows the refitted misfit of the faulted record can be smaller.

Three conclusions carry beyond the two halls. The point estimate is the answer to the estimation question only with its checks: the standardized and normalized residuals of every row, the chi-square of the redundant rows and the physical admissibility of the estimate. Redundancy is what makes the check possible, and where it lies decides what the check can see: a reading with a small redundancy number is absorbed by the entry it moves, and a second sensor on the same balance is the measurement that gives it a check, while the measurement that removes the most width can leave the check as weak as before. And the fit and its curvature come together: the matrix $\mathbf{J}^\mathsf{T}\mathbf{J}$ at the MAP state is the information matrix from which the spread of every entry, and of every engineering output, is derived.

Table 18.12 lists the terms of the chapter with their expressions, units and the subsection that defines each.

Table 18.12: The terms of maximum a posteriori estimation and where each is defined.

Term	Expression	Units	Meaning	Defined in
Objective	$\mathcal{L} = \frac{1}{2}\sum_k r_k^2$, eq. (18.9)	–	minus log posterior	18.2.1
MAP state	$\hat{\mathbf{z}} = \arg\min \mathcal{L}$	varies	highest posterior density	18.2.1
Reduced-state rows	$\mathbf{J} = \mathbf{J}_\mathrm{e}\mathbf{N}$, $\sigma_p \to 0$	varies	balances held exactly in the limit	18.1.3
Physical tolerance	$\sigma_T^2 \to \sigma_T^2 + \sigma_p^2$, $\sigma_P^2 \to \sigma_P^2 + \sigma_{p,P}^2$	K^2, kW^2	balances with a discrepancy	18.4.3
Laplace covariance	$\mathbf{\Sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}$, eq. (18.14)	varies	spreads $\sqrt{\Sigma_{ii}}$	18.2.1
Stationarity	$\mathbf{J}^\mathsf{T}\mathbf{r} = \mathbf{0}$, eq. (18.12)	varies	balance of pulls	18.2.1
Normal equations	$\mathbf{J}^\mathsf{T}\mathbf{J}\,\Delta\mathbf{z} = -\mathbf{J}^\mathsf{T}\mathbf{r}$	varies	Gauss–Newton step	18.2.1
Exact Hessian	$\mathbf{J}^\mathsf{T}\mathbf{J} + \sum_k r_k \nabla^2 r_k$	varies	Newton's curvature	18.2.2
Dropped-term bound	$\lvert r_i\rvert\,\sigma_T/(T_i - T_s)$	–	size of $\mathbf{S}$ per hall row	18.2.2
Stopping rule	$\max_i \lvert\Delta z_i\rvert/(1 + \lvert z_i\rvert) < \mathrm{tol}$	–	relative step	18.2.2
Levenberg–Marquardt	$(\mathbf{J}^\mathsf{T}\mathbf{J} + \varpi\,\mathrm{diag}\,\mathbf{J}^\mathsf{T}\mathbf{J})\Delta\mathbf{z} = -\mathbf{J}^\mathsf{T}\mathbf{r}$	–	damped step	18.2.2
Hat matrix	$\mathbf{K} = \mathbf{J}(\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}\mathbf{J}^\mathsf{T}$	–	fitted part of the rows	18.3.2
Chi-square result	$2\mathcal{L}(\hat{\mathbf{z}}) \sim \chi^2_{n_r - n_z}$	–	misfit of redundant rows	18.3.2
Redundancy number	$1 - K_{kk}$, $\sum_k = \nu$; prior row $1 - \Sigma_{ii}/\sigma_i^2$	–	where redundancy lies	18.3.2
Standardized residual	$r_k(\hat{\mathbf{z}})$	–	misfit in its own σ	18.3.2
Normalized residual	$r_k(\hat{\mathbf{z}})/\sqrt{1 - K_{kk}}$	–	bad-data statistic	18.3.2
Chi-square tail ($\nu = 3$)	$\sqrt{2x/\pi}\,\mathrm{e}^{-x/2} + 2[1 - \Phi(\sqrt{x})]$	–	agreement check	18.3.2
Admissibility	$G_i, Q_i > 0$, $T_i > T_s$	–	physical check	18.3.3
Square system	$n_r = n_z$, $\mathcal{L}(\hat{\mathbf{z}}) = 0$	–	no check possible	18.4.6
Preposterior	$\mathbf{\Sigma}_a = (\mathbf{\Sigma}^{-1} + \mathbf{J}_a^\mathsf{T}\mathbf{J}_a)^{-1}$, eq. (18.30)	varies	after a candidate row	18.6.2
Fault noncentrality	$(\Delta y/\sigma_T)^2(1 - K_{kk})$, eq. (18.31)	–	what the check sees of a bias, linear rows	18.6.2
Faulty sensor	refit with $y_{T_2} = 22.0°\mathrm{C}$	–	what the residuals show of a fault	18.3.2

18.8 Exercises

Exercise 18.1. *Formulation.* (a) The supply temperature alone is estimated from its sensor, $y_{T_s} = 12.10°\mathrm{C}$ with $\sigma_s = 0.3\,\mathrm{K}$, and the set-point prior $\mathcal{N}(12, 1^2)$. Write the one-entry state, the residual vector of the two rows and its 2×1 Jacobian; solve the normal equation; state $\mathbf{\Sigma}$, the hat matrix's diagonal, the redundancy numbers and normalized residuals, and the reference distribution of $2\mathcal{L}(\hat{\mathbf{z}})$. (b) State the redundancy ν of the two-hall plant without the hall-1 load meter; with a second hall-1 temperature sensor added; and without the three priors and the hall-1 load meter. ((a) $\mathbf{z} = (T_s)$, $\mathbf{r} = ((T_s - 12.10)/0.3,\ (T_s - 12)/1)^\mathsf{T}$, $\mathbf{J} = (3.333,\ 1)^\mathsf{T}\,\mathrm{K}^{-1}$. The normal equation $\mathbf{J}^\mathsf{T}\mathbf{J}\,\hat{T}_s = \mathbf{J}^\mathsf{T}\mathbf{b}$ with $\mathbf{b} = (12.10/0.3,\ 12)^\mathsf{T}$ reads $12.111\,\hat{T}_s = 146.44$, so $\hat{T}_s = 12.092°\mathrm{C}$, one step from any start, and $\mathbf{\Sigma} = 1/12.111 = 0.0826\,\mathrm{K}^2$. The hat matrix $\mathbf{J\Sigma J}^\mathsf{T}$ has diagonal $(0.917,\ 0.083)$, so the redundancy numbers are 0.083 for the sensor and 0.917 for the prior, summing to $\nu = 1$. The residuals are $\mathbf{r}(\hat{\mathbf{z}}) = (-0.0275,\ 0.0917)^\mathsf{T}$, the normalized residuals ∓ 0.096, equal in magnitude as for any two rows on one entry, and $2\mathcal{L}(\hat{\mathbf{z}}) = 0.0092 \sim \chi^2_1$, tail 0.92; the same value is $(y_{T_s} - T_s^0)^2/(\sigma_s^2 + \sigma_0^2) = 0.01/1.09$. (b) $7 - 5 = 2$; $9 - 5 = 4$; four rows for five entries, $\nu = -1$: the fit is not unique, no chi-square check exists, and the rank of the design is the subject of Chapter 19.)

Exercise 18.2. *Derivation.* Derive the Gauss–Newton step as the minimizer of the linearized objective, and show that it is a descent direction of the true objective whenever $\mathbf{J}$ has full column rank and the gradient is not zero. ($\mathcal{L}_{\mathrm{lin}}(\Delta\mathbf{z}) = \frac{1}{2}\|\mathbf{r} + \mathbf{J}\Delta\mathbf{z}\|^2$ has gradient $\mathbf{J}^\mathsf{T}(\mathbf{r} + \mathbf{J}\Delta\mathbf{z})$ and constant Hessian $\mathbf{J}^\mathsf{T}\mathbf{J}$, positive definite under full column rank, so the unique minimizer solves $\mathbf{J}^\mathsf{T}\mathbf{J}\,\Delta\mathbf{z} = -\mathbf{J}^\mathsf{T}\mathbf{r}$. Then $\nabla\mathcal{L}^\mathsf{T}\Delta\mathbf{z} = (\mathbf{J}^\mathsf{T}\mathbf{r})^\mathsf{T}\Delta\mathbf{z} = -\Delta\mathbf{z}^\mathsf{T}\mathbf{J}^\mathsf{T}\mathbf{J}\,\Delta\mathbf{z} < 0$, nonzero because $\Delta\mathbf{z} = \mathbf{0}$ would require $\mathbf{J}^\mathsf{T}\mathbf{r} = \mathbf{0}$.)

Exercise 18.3. *Derivation.* For the hall-1 row $r_1 = (T_s + Q_1 \mathrm{e}^{-g_1} - y_{T_1})/\sigma_T$, derive its contribution to the exact Hessian's extra term $\mathbf{S}$, and bound it at $\hat{\mathbf{z}}$ against the row's own Gauss–Newton contribution. ($\partial^2 r_1/\partial g_1^2 = Q_1\mathrm{e}^{-g_1}/\sigma_T = (T_1 - T_s)/\sigma_T$, $\partial^2 r_1/\partial g_1 \partial Q_1 = -\mathrm{e}^{-g_1}/\sigma_T$, all others zero, so $r_1\nabla^2 r_1$ has $r_1(T_1 - T_s)/\sigma_T$ at (g_1, g_1) and $-r_1\mathrm{e}^{-g_1}/\sigma_T$ at (g_1, Q_1). The row's Gauss–Newton term at (g_1, g_1) is $((T_1 - T_s)/\sigma_T)^2$, so the ratio is $|r_1|\,\sigma_T/(T_1 - T_s)$, the residual over the rise in units of σ_T. At $\hat{\mathbf{z}}$: $0.0930 \times 11.18 = 1.039$ against $11.18^2 = 125.0$, a ratio of 0.0083; off the diagonal $-0.0930 \times 0.3643 = -0.0339$ against $(-11.18)(0.3643) = -4.072$. At this rise the term would reach a tenth of the kept one only for a residual of $|r_1| = 1.1$.)

Exercise 18.4. *Derivation.* A linear model has one reading $y = az + c + \varepsilon$, $\varepsilon \sim \mathcal{N}(0, \sigma^2)$, and one prior $z \sim \mathcal{N}(z^0, \sigma_z^2)$. For the one-entry state $\mathbf{z} = (z)$ write the residual vector and the 2×1 Jacobian, derive the MAP estimate from the normal equation as a weighted mean, the minimized objective, and the number of Gauss–Newton steps needed from any start; then take $a = 30$, $c = 18$ and identify $\mathbf{J}$ with a column of Chapter 4's Jacobian. ($\mathbf{r} = ((az + c - y)/\sigma,\ (z - z^0)/\sigma_z)^\mathsf{T}$, $\mathbf{J} = (a/\sigma,\ 1/\sigma_z)^\mathsf{T}$, and the normal equation $\mathbf{J}^\mathsf{T}\mathbf{J}\,\hat{z} = \mathbf{J}^\mathsf{T}\mathbf{b}$ with $\mathbf{b} = ((y - c)/\sigma,\ z^0/\sigma_z)^\mathsf{T}$ gives $\hat{z} = [a(y - c)/\sigma^2 + z^0/\sigma_z^2]/[a^2/\sigma^2 + 1/\sigma_z^2]$, the mean of $(y - c)/a$ and z^0 weighted by a^2/σ^2 and $1/\sigma_z^2$, the two entries of $\mathbf{J}$ squared; $2\mathcal{L}(\hat{z}) = (y - c - az^0)^2/(\sigma^2 + a^2\sigma_z^2) \sim \chi^2_1$. The rows are linear, $\mathbf{S} = \mathbf{0}$, and one step reaches $\hat{z}$ from any start. With $y = 23.2$, $\sigma = 0.5$, $z^0 = 0.160$, $\sigma_z = 0.025$: $\mathbf{J} = (60,\ 40)^\mathsf{T}$, which is the column $\mathbf{Jv}$ of Chapter 4's three-entry fit along its free direction $\mathbf{v}$, equation (4.30), read on the reading and prior rows; $\mathbf{J}^\mathsf{T}\mathbf{J} = 5200$, $\mathbf{J}^\mathsf{T}\mathbf{b} = 880$, and $\hat{z} = 0.16923\,\mathrm{K/kW}$, the first entry of equation (4.45), with $\mathcal{L}(\hat{z}) = 0.0985$, the minimum of Figure 4.3.)

Exercise 18.5. *Application.* The hall-2 sensor develops a fault and reads $y_{T_2} = 22.0°\mathrm{C}$ instead of $17.70°\mathrm{C}$, the other readings and priors unchanged. Refit the plant, state what the standardized residuals, the normalized residuals and the chi-square check say, and explain why the faulty reading's redundancy number is smaller than at the clean fit. (The refit is that of Section 18.3.2: $\hat{G}_2 = 2.452\,\mathrm{kW/K}$, $2\mathcal{L}(\hat{\mathbf{z}}) = 6.687$, tail 0.083, no rejection at 5 %; the reading's error of $8.6\,\sigma_T$

is absorbed by g_2 and appears on the prior of g_2, -2.38, while the reading itself shows -0.41; the normalized residuals of the reading, the prior of g_2 and the chiller meter are tied at 2.56 in magnitude, so they flag the chain but cannot isolate the reading, nor separate a faulty sensor from a degraded path. The redundancy number falls from 0.073 to 0.0256 because the fit raises the hall-2 rise, and with it the magnitude of the g_2 entry of row 2, minus the rise over σ_T by equation (18.6); the row's leverage K_{22} grows and the state follows the reading more closely. A later chapter on bad data compares the explanations by their evidence.)

Exercise 18.6. *Application.* A second hall-2 temperature sensor, $\sigma_T = 0.5\,\mathrm{K}$, is installed and reads 17.70°C while the first reads the faulty 22.0°C of the previous exercise. Write the added row, state the new redundancy, refit, and state what the residuals and the chi-square check now say. (The ninth row is $(T_2(\mathbf{z}) - 17.70)/\sigma_T = (T_s + Q_2 \mathrm{e}^{-g_2} - 17.70)/\sigma_T$, the same function of the state as row 2, so $n_r = 9$ and $\nu = 4$. The refit is that of Section 18.3.2: $\hat{G}_2 = 3.029\,\mathrm{kW/K}$ and $T_2(\hat{\mathbf{z}}) = 19.76°\mathrm{C}$, between the two readings; normalized residuals -6.28 and $+5.76$ on the two sensors; $2\mathcal{L}(\hat{\mathbf{z}}) = 40.4$ on four degrees of freedom, tail 3.6×10^{-8}. The fault is detected and the two sensors on one balance are named as disagreeing; the pair alone does not prove which is wrong, and a third row on that balance, or the comparison of a later chapter, settles it. With both sensors correct at 17.70°C, $2\mathcal{L} = 0.763$, tail 0.94.)

Chapter 19

The Hessian, Covariance and Identifiability

19.1 The spread of a fitted state and what the readings determine

The operators of the two data halls have the fit of Chapter 18: five readings and three handbook priors, a MAP state that puts hall 2's heat-removal conductance at 4.055 kW/K and its load at 22.2 kW, and residuals that pass the χ^2 check. Next month the hall-1 load meter and the chiller heat meter go out of service together for replacement, in the same shutdown that cleans both halls' cooling coils, and the facility engineer must decide whether hall 2, whose electrical feed is rated for 45 kW, can run that month without a temporary meter. The first reading comes from the fit itself. Refitted on the three temperature readings that remain, it still converges, puts hall 2's load at 28.0 kW with a standard deviation of 8.9 kW, and its Gaussian band gives a load above 45 kW a probability of 2.8 %.

The engineer wants to know what that number rests on. A fit returns a spread for every entry of the state whether or not any reading sees the entry, and a converged fit with small residuals says nothing about which quantities the readings determine, which only the handbook values hold, or whether a Gaussian centered at the minimum describes the posterior at all once a meter is gone. The question of this chapter is

how well the fitted state of the two halls is known, entry by entry and in combination; which entries and which combinations of entries the readings determine at all and which only the priors hold; what is lost when a meter fails; and whether the Gaussian band computed from the curvature can be trusted when it is.

The rows and the reduced state are those of Chapters 17 and 18. What changes is the object read from them: not the minimum of the objective, which Chapter 18 found, but its curvature there, and the design of rows that the curvature reflects.

Jargon: Laplace approximation, Hessian, Gauss–Newton curvature, observed information, marginal and conditional variance, Schur complement, squared multiple correlation, singular value decomposition, condition number, observability analysis, design columns, structural and practical identifiability

The *Laplace approximation* replaces a posterior by the Gaussian centered at its mode whose inverse covariance is the *Hessian*, the matrix of second derivatives of the negative log posterior there. The *Gauss–Newton curvature* $\mathbf{J}^\mathsf{T}\mathbf{J}$ is the part of that Hessian built from first derivatives of the rows only; the *observed information* is the exact Hessian of the negative log likelihood at the data, which for a lifetime row, or a count row off its canonical link, adds a curvature term to $\mathbf{J}^\mathsf{T}\mathbf{J}$. The *marginal variance* of an entry is its variance with the other entries unknown; its *conditional variance* is its variance with the others held fixed. The *Schur complement* of a block of a matrix is the quantity that links the two, and the *squared multiple correlation* of an entry is the fraction of its marginal variance that the other entries explain. The *singular value decomposition* writes a matrix as a rotation, a scaling along orthogonal directions and a second rotation; its *condition number* is the ratio of its largest to its smallest scaling. *Observability analysis*, in state-estimation practice, decides whether a set of measurements determines the state, by the rank of the measurement Jacobian (numerical observability) or from the measurement graph (topological observability). The *design columns* are the columns of the Jacobian of the reading rows alone. *Structural identifiability* asks whether noise-free readings would determine the unknowns at all; *practical identifiability* asks whether the actual readings determine them to a useful precision.

Key idea

The inverse curvature of the objective at the MAP state is the posterior covariance under the Laplace approximation. Its diagonal gives the spread of each entry, its off-diagonal the combinations the readings fix, and the rank and singular values of the reading rows' Jacobian say which entries and combinations the readings determine at all. All of it is a property of the design, which rows exist and where the Jacobian is taken, and not of the values read, so what a failed meter costs is known before it fails. The Gaussian band is an approximation with known failure conditions, and an unobservable direction held only by a log-scale prior is where it fails worst.

19.1.1 From the point to the spread

Chapter 4 obtained the spread of the equipment hall's state from the curvature of the objective, in state-vector form: for the state $\mathbf{z} = (\rho, T_1, T_2)^\mathsf{T}$ of the hall with two physical rows of tolerance σ_p, a reading and a prior, the information matrix $\mathbf{J}^\mathsf{T}\mathbf{J}$ was the posterior precision and its inverse $\mathbf{\Sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}$, equation (4.46), the posterior covariance, exactly for that linear-Gaussian model; the Laplace approximation (4.38) extended the construction to a nonlinear model as a local Gaussian. There the two physical rows fixed two of the three directions of the state at an information of order $1/\sigma_p^2$, and the third, the free direction $\mathbf{v}$ along the balances, was the eigenvector of $\mathbf{\Sigma}$ that carried nearly all of its variance. The curvature along that one direction, $\mathbf{v}^\mathsf{T}\mathbf{J}^\mathsf{T}\mathbf{J}\,\mathbf{v} = 5200$ of equation (4.36), was a one-line remark beside the matrix, the same number seen along a line.

For the two halls the projection of Chapter 17 has already set aside the directions that the physics rows fix, and on the five entries that remain the curvature is a five-by-five matrix, three of whose properties come to the front. The off-diagonal terms of its inverse are covariances: the readings may fix a combination of two entries better than either entry. The spread of an entry with the others unknown differs from its spread with the others held fixed. And the curvature has directions: it may be large along some combinations of the entries, small along others, and

zero along a combination that no reading touches. Exercise 4.2 posed the first in miniature, and Exercise 4.4 the last, a single reading that could not separate two unknowns.

19.1.2 Observability and identifiability

The owner of the two halls asks of each quantity: can my sensors see this parameter at all? State-estimation practice calls the question *observability analysis* when it is asked of states such as temperatures and flows, and parameter estimation calls it *identifiability* when it is asked of parameters such as conductances. The companion volume *Reliability Engineering* treats the two in separate chapters on observability and on identifiability, because there the states and the parameters are different objects. In this book the temperatures, the loads and the log-conductances are all entries of one state, and one test serves both: the rank of the design columns, the Jacobian of the reading rows with the priors left out, and the comparison of each entry's posterior spread with its prior spread. An entry the readings cannot see keeps the spread its prior gave it, whatever the fit returns.

The test is a property of the design: which rows exist, what they measure, and at which operating point the Jacobian is taken. It does not depend on the values the meters read, because the curvature at a fixed state does not, as Chapter 4 observed for one reading. The question can therefore be answered before a meter is installed, and before one fails, without a new reading: what is lost when the hall-1 load meter goes out of service is a calculation on the existing Jacobian with one row removed.

19.1.3 The plant and its reduced state (step 0)

The plant is the one of Chapter 17, whose graph is Figure 17.1. Two data halls receive heat loads Q_1 and Q_2 from their equipment and reject them through heat-removal paths of conductance G_1 and G_2 to one chilled-water supply header at temperature T_s; the chiller is the source that holds the header. The conductances are carried by the log-conductances $g_i = \ln(G_i/G_{\mathrm{ref}})$ with $G_{\mathrm{ref}} = 1\,\mathrm{kW/K}$ pinned. Chapter 17 wrote the full state of the plant with the hall temperatures T_1, T_2 and the chiller's heat among its entries and the steady balances, $Q_i = G_i(T_i - T_s)$ for each hall and the header balance, as physics rows of tolerance σ_p, and derived from the null directions of those rows the reduced state (Section 17.2.1), the free entries along which the balances keep holding:

$$\mathbf{z} = \begin{bmatrix} g_1 \\ g_2 \\ Q_1 \\ Q_2 \\ T_s \end{bmatrix}, \qquad T_i(\mathbf{z}) = T_s + Q_i\,\mathrm{e}^{-g_i}, \tag{19.1}$$

with g_i dimensionless, Q_i in kW and T_s in °C, and the hall temperature $T_i(\mathbf{z})$ the value the balance gives along the null directions, the rise $Q_i\mathrm{e}^{-g_i}$ in kelvin with $G_{\mathrm{ref}} = 1\,\mathrm{kW/K}$ absorbed. The five entries are the free quantities of the graph; the hall temperatures are functions of them.

19.1.4 The record

The evidence is five readings and three priors, those of Chapter 17: the hall temperatures $y_{T_1} = 17.655°\mathrm{C}$ and $y_{T_2} = 17.700°\mathrm{C}$ with $\sigma_T = 0.5\,\mathrm{K}$; the supply temperature $y_{T_s} = 12.10°\mathrm{C}$ with $\sigma_s = 0.3\,\mathrm{K}$; the chiller's heat meter $y_P = 52.5\,\mathrm{kW}$ for $Q_1 + Q_2$ with $\sigma_P = 2\,\mathrm{kW}$; an electrical meter on hall 1's load $y_{Q_1} = 31.0\,\mathrm{kW}$ with $\sigma_Q = 1.5\,\mathrm{kW}$; and the priors $g_i \sim \mathcal{N}(\ln 5,\ 0.3^2)$, a handbook conductance of $5\,\mathrm{kW/K}$ with a spread of about 35 %, and $T_s \sim \mathcal{N}(12,\ 1^2)$, the set point. The loads have no prior: their prior is flat. The subscript of σ_P names the chiller meter, which reads the duty $P = Q_1 + Q_2$; the lowercase σ_p is the tolerance of the physics rows, a different quantity, and $\sigma_{0,i}$ below names the prior standard deviation of entry i (0.3, 0.3, infinite,

infinite, 1 K). The numbers are illustrative, generated from $G_1 = 5.5$, $G_2 = 4.0\,\mathrm{kW/K}$, $Q_1 = 30$, $Q_2 = 24\,\mathrm{kW}$ and $T_s = 12^\circ\mathrm{C}$ with stated errors. The meter outage of the opening removes the rows of the hall-1 load meter and the chiller heat meter; the record of the outage month is written $\mathbf{y}'$, the three temperature readings with the three priors, the handbook priors on the conductances included, since the cleaning of the coils leaves no conductance fitted before the shutdown in force.

19.1.5 The questions (step 1)

The engineer asks five questions of the record, and Table 19.1 lists them. The first column is the question in engineering language; the second the general form of its category in the notation of Chapter 2; the third its mathematical form for the two halls; the fourth its category; and the last the subsection where it is answered. The third is the plant owner's question, quoted. The first three are estimation, of the spread of each entry, of the shape of the posterior and of the comparison of the posterior with the prior. The fourth is a design question asked in reverse: not which row to add, but what a row's loss removes. The fifth is a prediction under a threshold, the engineer's question for the outage month, asked of the record $\mathbf{y}'$. Which row to add next asks nothing new of the record; it is the next measurement for these five, Section 19.6. The atlas of Appendix A lists the first, the third and the fifth.

Table 19.1: The questions asked of the two halls' curvature, the general form of their category, their mathematical form, their category and where each is answered. $\mathbf{y}$ is the five readings and the three priors; $\mathbf{y}_{-Q_1}$ the same without the hall-1 load meter, and $\mathbf{y}'$ without that meter and the chiller heat meter; $\mathbf{J}_\mathrm{d}$ the Jacobian of the reading rows; $\mathbf{v}$ a unit vector in the space of the state; $\tilde{y}_a$ the outcome of a change a of the design.

Engineering question	General form	Mathematical form	Category	Answered in
How well are the conductances, the loads and the supply temperature known?	$\mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y})$, quantiles	$\mathrm{Var}(z_i \mid \mathbf{y})$, $\mathrm{Cov}(z_i, z_j \mid \mathbf{y})$ under $p(\mathbf{z} \mid \mathbf{y}) \approx \mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$	estimation (the spread)	19.3.1
Which direction of the state do the readings determine least?	$\mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y})$	the unit $\mathbf{v}$ that maximizes $\mathrm{Var}(\mathbf{v}^\mathsf{T}\mathbf{z} \mid \mathbf{y})$	estimation (the shape of the posterior)	19.3.1
Can my sensors see this parameter at all?	$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y})$ against $\Pr(f(\mathbf{z}) \le f^*)$	unobserved when $p(z_i \mid \mathbf{y}) = p(z_i)$; $\sqrt{\mathrm{Var}(z_i)/\,\mathrm{Var}(z_i \mid \mathbf{y})}$; rank of $\mathbf{J}_\mathrm{d}$	estimation (observability analysis)	19.3.2
What is lost if the hall-1 load meter fails? The chiller meter too? The supply sensor?	$U(a)$, the design form	$\mathrm{Var}(z_i \mid \mathbf{y}_{-Q_1})$, and the same with the other rows removed, against $\mathrm{Var}(z_i \mid \mathbf{y})$	design (observability by design)	19.3.3
With both meters out, how likely is hall 2's load above 45 kW?	$\Pr(f(\mathbf{z}) \ge f^* \mid \mathbf{y})$	$\Pr(Q_2 > 45\,\mathrm{kW} \mid \mathbf{y}')$	prediction under a threshold	19.3.4

19.1.6 The evidence as rows (step 2)

Step 2 attaches the record to the reduced state as rows: five readings and three priors, all Gaussian, each written below in normalized form, with a residual and a Jacobian row.

The modeling choices. Three choices make the posterior, and each buys something and can fail. The chapter works on the reduced state of Chapter 17, whose Jacobian is the projection $\mathbf{J} = \mathbf{J}_{\mathrm{e}}\mathbf{N}$ of the reading and prior rows onto the null-space basis $\mathbf{N}$ of the physics rows, equation (17.18), and whose covariance is the free-entry marginal of the full state's in the limit $\sigma_p \to 0$, equation (17.19): the exact balances are that limit. This buys a five-entry curvature with no direction of order $1/\sigma_p$ in it, and it fails if a balance itself misses by a physical amount, the discrepancy that Section 19.4.4 varies and Section 19.6.5 discusses. The posterior is replaced by its Laplace approximation, the Gaussian at the MAP state with the inverse curvature as covariance, and this is the route the chapter reports for the first four questions: it makes every spread, correlation and direction linear algebra on one matrix, and it fails where the objective departs from its quadratic within the posterior's own width, which Section 19.4.1 tests against the exact posterior and Section 19.4.5 lists condition by condition. The curvature is the Gauss–Newton matrix $\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$, without the residual term of the exact Hessian: it is the matrix the fit already formed at every iteration, and it misleads when the residuals at the MAP state are large, which the exact Hessian of Section 19.4.1 measures.

Written normalized, each divided by its standard deviation so that every weight is one, the eight rows stack into the residual vector

$$\mathbf{r}(\mathbf{z}) = \begin{bmatrix} (T_s + Q_1 \mathrm{e}^{-g_1} - y_{T_1})/\sigma_T & \text{reading, hall-1 temperature} \\ (T_s + Q_2 \mathrm{e}^{-g_2} - y_{T_2})/\sigma_T & \text{reading, hall-2 temperature} \\ (T_s - y_{T_s})/\sigma_s & \text{reading, supply temperature} \\ (Q_1 + Q_2 - y_P)/\sigma_P & \text{reading, chiller heat meter} \\ (Q_1 - y_{Q_1})/\sigma_Q & \text{reading, hall-1 load meter} \\ (g_1 - \ln 5)/0.3 & \text{prior on } g_1 \\ (g_2 - \ln 5)/0.3 & \text{prior on } g_2 \\ (T_s - 12)/1 & \text{prior on } T_s \end{bmatrix} \tag{19.2}$$

five reading rows and three prior rows, each hall reading compared with the hall temperature $T_i(\mathbf{z})$ along the null directions of the physics rows, with the objective $\mathcal{L}(\mathbf{z}) = \frac{1}{2}\,\mathbf{r}(\mathbf{z})^{\mathsf{T}}\mathbf{r}(\mathbf{z})$. All eight are Gaussian rows; none is a physics row on this state, since the balances act through the projection. Which row acts on which entry fixes where the Jacobian $\mathbf{J} = \partial\mathbf{r}/\partial\mathbf{z}^{\mathsf{T}}$, eight by five, can be nonzero, marked $*$:

$$\mathbf{J} = \begin{bmatrix} * & 0 & * & 0 & * \\ 0 & * & 0 & * & * \\ 0 & 0 & 0 & 0 & * \\ 0 & 0 & * & * & 0 \\ 0 & 0 & * & 0 & 0 \\ * & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * \end{bmatrix} \begin{matrix} r_1, \text{ hall-1 temperature} \\ r_2, \text{ hall-2 temperature} \\ r_3, \text{ supply temperature} \\ r_4, \text{ chiller heat meter} \\ r_5, \text{ hall-1 load meter} \\ r_6, \text{ prior on } g_1 \\ r_7, \text{ prior on } g_2 \\ r_8, \text{ prior on } T_s \end{matrix}, \qquad \text{columns } g_1, g_2, Q_1, Q_2, T_s. \tag{19.3}$$

Read down the columns, the pattern is the design. The log-conductance g_i reaches only its own hall's temperature reading. The supply temperature reaches both hall readings and its own sensor: one sensor and two physical rows share it. The load Q_1 reaches its hall's reading and both meters; Q_2 reaches only its hall's reading and the chiller meter. The two halls share one column, T_s, and one row, the chiller meter.

The rows and their Jacobian. The eight rows of equation (19.2) are Gaussian rows in working form already: each is a difference divided by its standard deviation, its term in the objective is $\mathcal{L}_k = \frac{1}{2}r_k^2$, and the Jacobian $\mathbf{J}$ has one row per row of the model and one column per entry of the state, $J_{ki} = \partial r_k/\partial z_i$, eight by five, with the pattern (19.3). The hall rows are

the only nonlinear ones. Differentiating r_1 entry by entry,

$$\frac{\partial r_1}{\partial g_1} = -\frac{Q_1 \mathrm{e}^{-g_1}}{\sigma_T} = -\frac{\Delta T_1}{\sigma_T}, \qquad \frac{\partial r_1}{\partial Q_1} = \frac{\mathrm{e}^{-g_1}}{\sigma_T} = \frac{1}{G_1 \sigma_T}, \qquad \frac{\partial r_1}{\partial T_s} = \frac{1}{\sigma_T}, \tag{19.4}$$

with $\Delta T_1 = T_1 - T_s$ the temperature rise of hall 1, and the same for r_2 with the index 2. These are the entries of $\mathbf{J}_\mathrm{e}\mathbf{N}$ of Chapter 17: the hall reading of the full state sees only its hall temperature, with the derivative $1/\sigma_T$, and the null direction of g_1 moves that temperature by $-\Delta T_1$, the direction of Q_1 by $1/G_1$ and the direction of T_s by one. The derivative of a hall reading with respect to the log-conductance is the temperature rise counted in standard deviations of the sensor, which is why the log scale suits the conductance: a hall resolves its log-conductance to about $\sigma_T/\Delta T_i$, a relative precision of the conductance that depends on the rise and the sensor, whatever the values. The other rows are linear, and their derivatives are the constants $1/\sigma$ of the row in the column of each entry it reads, 1/2 for each load in the chiller meter's row.

A count, demand or lifetime row would enter the same objective in the working form of Section 4.2.7, a working residual with its Jacobian row (equation (4.43); the instances of equations (6.22) and (7.33)) and its term $\mathcal{L}_k$ the negative log-probability of the count or the lifetime; $\mathbf{J}^\mathsf{T}\mathbf{J}$ is then the expected information, and the Hessian adds the curvature term $\mathbf{C}$ of equation (4.44), observed minus expected, as in equation (7.34), which is zero for a count row on its canonical link (the logit of a demand probability, the log of a rate) and not for a lifetime row. The two halls have no such row.

The rows split into the reading rows and the prior rows, and the Jacobian splits with them,

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_\mathrm{d} \\ \mathbf{J}_\pi \end{bmatrix}, \qquad \mathbf{J}^\mathsf{T}\mathbf{J} = \mathbf{J}_\mathrm{d}^\mathsf{T}\mathbf{J}_\mathrm{d} + \mathbf{\Lambda}_\mathrm{prior}, \qquad \mathbf{\Lambda}_\mathrm{prior} = \mathbf{J}_\pi^\mathsf{T}\mathbf{J}_\pi = \mathrm{diag}(11.11,\ 11.11,\ 0,\ 0,\ 1), \tag{19.5}$$

where $\mathbf{J}_\mathrm{d}$, the first five rows, holds the derivatives of the readings, $\mathbf{J}_\pi$ the prior rows (the subscript p is kept for the physics rows of Chapter 17), and $\mathbf{\Lambda}_\mathrm{prior}$ is the prior precision of equation (4.26), with $1/0.3^2 = 11.11$ for each log-conductance, $1/1^2$ for the supply temperature and zero for the loads, whose prior is flat. The columns of $\mathbf{J}_\mathrm{d}$ are the design columns: column i lists how every reading responds to entry i.

The graph is not extended for these questions; it is shrunk. The fourth question removes a reading row, or two, and asks what the remaining rows determine; the fifth removes the rows of the hall-1 load meter and the chiller heat meter, r_4 and r_5, and asks a tail of the posterior that the remaining six rows leave.

19.2 The posterior of the state (step 3)

In the book's notation $p(\mathbf{z} \mid \mathbf{y})$ is the posterior density of the state, $f(\mathbf{z})$ a quantity derived from the state and f^* a level on it.

Four objects of the fit recur in the specialized forms. The *MAP state* $\hat{\mathbf{z}}$ is the minimum of the objective that the eight rows put on the state, found by Chapter 18; the *information matrix* $\mathbf{\Lambda} = \mathbf{J}^\mathsf{T}\mathbf{J}$ is the Gauss–Newton curvature of that objective at the minimum, with $\mathbf{J}$ the 8×5 Jacobian of the normalized rows; the *Laplace covariance* $\mathbf{\Sigma} = \mathbf{\Lambda}^{-1}$ makes the posterior of the state approximately normal, $\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}, \mathbf{\Sigma})$; and the *design columns* $\mathbf{J}_\mathrm{d}$ of equation (19.5) carry what the readings alone can say.

All five questions are read from one posterior: the posterior of the reduced state that the eight rows of step 2 give, approximately the Laplace normal about the MAP state. This section derives it and computes it for the two halls; Section 19.3 reads every answer from it.

19.2.1 The posterior from the rows

By Bayes' theorem with independent Gaussian rows, as Chapter 18 derived, the posterior is

$$p(\mathbf{z} \mid \mathbf{y}) = \frac{\exp[-\mathcal{L}(\mathbf{z})]}{\int \exp[-\mathcal{L}(\mathbf{z}')]\,\mathrm{d}\mathbf{z}'}, \tag{19.6}$$

with every constant of the row densities absorbed into the normalizing integral. Chapter 18 minimized $\mathcal{L}$ by Gauss–Newton, solving the normal equations $(\mathbf{J}^\mathsf{T}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^\mathsf{T}\mathbf{r}$ at each iterate until the gradient $\mathbf{J}^\mathsf{T}\mathbf{r}$ vanished. Write $\boldsymbol{\delta} = \mathbf{z} - \hat{\mathbf{z}}$ for the departure from the MAP state. The Taylor expansion of the objective to second order about $\hat{\mathbf{z}}$ is

$$\mathcal{L}(\mathbf{z}) = \mathcal{L}(\hat{\mathbf{z}}) + \nabla\mathcal{L}(\hat{\mathbf{z}})^\mathsf{T}\boldsymbol{\delta} + \tfrac{1}{2}\,\boldsymbol{\delta}^\mathsf{T}\mathbf{H}\,\boldsymbol{\delta} + \mathcal{E}_3(\boldsymbol{\delta}), \qquad \mathbf{H} = \left.\frac{\partial^2\mathcal{L}}{\partial\mathbf{z}\,\partial\mathbf{z}^\mathsf{T}}\right|_{\hat{\mathbf{z}}}, \tag{19.7}$$

where $\mathbf{H}$ is the $n \times n$ Hessian and $\mathcal{E}_3$ the remainder, of third order in $\boldsymbol{\delta}$. The gradient vanishes at the minimum, $\nabla\mathcal{L}(\hat{\mathbf{z}}) = \mathbf{0}$, which is the condition the fit solved. Dropping the remainder and inserting the expansion in equation (19.6), the constant $\exp[-\mathcal{L}(\hat{\mathbf{z}})]$ cancels between numerator and denominator:

$$p(\mathbf{z} \mid \mathbf{y}) \approx \frac{\exp(-\tfrac{1}{2}\boldsymbol{\delta}^\mathsf{T}\mathbf{H}\boldsymbol{\delta})}{\int \exp(-\tfrac{1}{2}\boldsymbol{\delta}'^\mathsf{T}\mathbf{H}\boldsymbol{\delta}')\,\mathrm{d}\boldsymbol{\delta}'}.$$

A Gaussian density with mean $\mathbf{m}$ and covariance $\boldsymbol{\Sigma}$ is proportional to $\exp[-\tfrac{1}{2}(\mathbf{z}-\mathbf{m})^\mathsf{T}\boldsymbol{\Sigma}^{-1}(\mathbf{z}-\mathbf{m})]$. The two exponents agree for every $\mathbf{z}$ if and only if $\mathbf{m} = \hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}^{-1} = \mathbf{H}$, so

$$\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}},\ \boldsymbol{\Sigma}), \qquad \boldsymbol{\Sigma} = \mathbf{H}^{-1}, \tag{19.8}$$

the Laplace approximation (4.38) of Chapter 4, now with its matrix written out.

The normalizing integral. Write the symmetric matrix as $\mathbf{H} = \mathbf{E}\mathbf{D}\mathbf{E}^\mathsf{T}$ with $\mathbf{E}$ orthogonal, its columns the eigenvectors, and $\mathbf{D} = \mathrm{diag}(d_1, \ldots, d_n)$ the eigenvalues. In the rotated coordinates $\mathbf{x} = \mathbf{E}^\mathsf{T}\boldsymbol{\delta}$, whose change-of-variable factor is one, the exponent is $-\tfrac{1}{2}\sum_j d_j x_j^2$ and the integral factorizes into n one-dimensional Gaussian integrals,

$$\int \exp(-\tfrac{1}{2}\boldsymbol{\delta}^\mathsf{T}\mathbf{H}\boldsymbol{\delta})\,\mathrm{d}\boldsymbol{\delta} = \prod_{j=1}^{n}\sqrt{\frac{2\pi}{d_j}} = (2\pi)^{n/2}\,(\det\mathbf{H})^{-1/2}. \tag{19.9}$$

Each factor exists only if $d_j > 0$. The approximation therefore requires $\mathbf{H}$ to be positive definite: a direction with $d_j = 0$ is one along which the objective is flat to second order, the posterior does not decay, and no covariance exists. Such a direction is an unobservable one, in the sense of Section 19.3.2, held by no prior. The approximation further requires the remainder to be small where the posterior has its mass, $|\mathcal{E}_3(\boldsymbol{\delta})| \ll 1$ for $\boldsymbol{\delta}^\mathsf{T}\mathbf{H}\boldsymbol{\delta}$ up to about 11, the value within which a Gaussian of five entries holds 95 % of its probability (the 0.95 quantile of the χ^2 distribution with five degrees of freedom, 11.07); Section 19.4.5 lists the ways this fails.

The curvature as an information matrix. Chapter 18 derived the Hessian of the objective $\mathcal{L} = \tfrac{1}{2}\sum_k r_k^2$ in Section 18.2.2: by equation (18.15) it is the Gauss–Newton matrix $\mathbf{J}^\mathsf{T}\mathbf{J}$ plus the term $\mathbf{S} = \sum_k r_k\nabla^2 r_k$, the curvature of the rows weighted by their residuals, which vanishes for linear rows and is small when the residuals at the MAP state are small. With every row normalized the first term is the information matrix $\boldsymbol{\Lambda} = \mathbf{J}^\mathsf{T}\mathbf{J}$; with unnormalized rows and the weights $w_k = 1/\sigma_k^2$ it is $\mathbf{J}^\mathsf{T}\mathbf{W}\mathbf{J}$, $\mathbf{W} = \mathrm{diag}(w_k)$, equation (4.26). Dropping $\mathbf{S}$ gives the working form

$$\boldsymbol{\Sigma} \approx \boldsymbol{\Lambda}^{-1} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}, \tag{19.10}$$

the covariance computed from first derivatives only, which the fit has already formed at every iteration; the answer is read off it as an entry $\hat{z}_i$ with $\sqrt{\Sigma_{ii}}$, a correlation, or a derived quantity by $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$.

Precisions add. Writing $\mathbf{j}_k$ for row k of the Jacobian as a column vector, the matrix product is a sum over rows,

$$\mathbf{\Lambda} = \sum_k \mathbf{j}_k \mathbf{j}_k^{\mathsf{T}}. \tag{19.11}$$

Each row contributes one outer product, a matrix of rank one that is positive semidefinite, since $\boldsymbol{\delta}^{\mathsf{T}} \mathbf{j}_k \mathbf{j}_k^{\mathsf{T}} \boldsymbol{\delta} = (\mathbf{j}_k^{\mathsf{T}} \boldsymbol{\delta})^2 \geq 0$ for every $\boldsymbol{\delta}$. A prior row on entry i with standard deviation $\sigma_{0,i}$ adds $1/\sigma_{0,i}^2$ to the diagonal entry Λ_{ii} and nothing elsewhere. Adding a row can only add curvature, and removing one, equation (19.26), can only take it away; Exercise 19.3 derives from this that adding a row can never increase a posterior variance.

The exact posterior by weighted draws. The check of the Laplace posterior needs the posterior (19.6) itself, which for five entries has no closed form. It is computed by importance sampling: states $\mathbf{z}^{(m)}$ are drawn from a multivariate t distribution with six degrees of freedom centered at $\hat{\mathbf{z}}$ with the scale matrix $\mathbf{\Sigma}$, and each is given a weight proportional to $\exp[-\mathcal{L}(\mathbf{z}^{(m)})]/q(\mathbf{z}^{(m)})$, with q the density it was drawn from, the weights normalized to sum to one. Every posterior moment and quantile is then a weighted sum over the draws. The heavy tails of the t distribution keep the weights bounded where the posterior is wider than the Laplace normal, and the effective number of draws, the inverse of the sum of the squared weights, says how many equally weighted draws the sample is worth.

19.2.2 The posterior of the two halls

The plant's inputs are the eight rows of equation (19.2) with the record of Section 19.1.4 and the MAP state of Chapter 18.

Chapter 18 minimized the objective by Gauss–Newton and found the MAP state

$$\hat{\mathbf{z}} = (1.7030,\ 1.3999,\ 30.688\,\mathrm{kW},\ 22.231\,\mathrm{kW},\ 12.111^\circ\mathrm{C})^{\mathsf{T}}, \qquad \mathcal{L}(\hat{\mathbf{z}}) = 0.3698, \tag{19.12}$$

that is $G_1 = 5.490$ and $G_2 = 4.055\,\mathrm{kW/K}$, with the hall temperatures $T_1 = 17.701^\circ\mathrm{C}$ and $T_2 = 17.594^\circ\mathrm{C}$ and the temperature rises $\Delta T_1 = T_1 - T_s = 5.590\,\mathrm{K}$ and $\Delta T_2 = 5.482\,\mathrm{K}$. By equation (19.4), $\partial r_1/\partial g_1 = -5.590/0.5 = -11.179$, $\partial r_1/\partial Q_1 = 1/(5.490 \times 0.5) = 0.3643\,\mathrm{kW}^{-1}$ and $\partial r_1/\partial T_s = 2\,\mathrm{K}^{-1}$: hall 1 resolves its log-conductance to about $1/11.2 = 0.09$.

Table 19.2 lists the Jacobian at $\hat{\mathbf{z}}$. Each row of the table is a row of the model, with its kind; the five columns of numbers are the derivatives with respect to the five entries, in the inverse units of each entry, filling the pattern (19.3).

The product $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ of the eight rows is, with the entries ordered $(g_1, g_2, Q_1, Q_2, T_s)$,

$$\mathbf{\Lambda} = \begin{bmatrix} 136.085 & 0 & -4.0724 & 0 & -22.358 \\ 0 & 131.338 & 0 & -5.4081 & -21.930 \\ -4.0724 & 0 & 0.8271 & 0.2500 & 0.7286 \\ 0 & -5.4081 & 0.2500 & 0.4933 & 0.9865 \\ -22.358 & -21.930 & 0.7286 & 0.9865 & 20.111 \end{bmatrix}, \tag{19.13}$$

with entry (i, j) in the inverse units of $z_i z_j$. Each entry is a sum over rows by equation (19.11); for instance $\Lambda_{11} = 11.179^2 + 3.333^2 = 124.97 + 11.11$, the hall-1 reading and the prior. The conductances draw 92 % of their curvature from their halls' readings; Q_2 draws 0.243 from its hall's reading and 0.25 from the chiller meter, the smallest diagonal entry. The eigenvalues run from 0.156 to 142.6, $\mathrm{cond}(\mathbf{\Lambda}) = 913 = 30.2^2$.

Table 19.3 lists its inverse, the posterior covariance $\mathbf{\Sigma}$ of the Laplace approximation (10), together with the correlations. Rows and columns are the five entries in the same order. The diagonal and the entries below it are the variances and covariances, in the products of the entries' units; the entries above the diagonal, in italics, are the correlations $\Sigma_{ij}/\sqrt{\Sigma_{ii}\Sigma_{jj}}$, without units. Both halves are symmetric, so each is given once. The product $\mathbf{\Lambda\Sigma}$ reproduces the identity to round-off.

Table 19.2: The Jacobian of the normalized rows (19.2) at the MAP state (19.12). Units: per unit of g_i (dimensionless), per kW for the loads, per K for the supply temperature. The first five rows are the reading rows; their columns are the design columns.

Row	Kind	g_1 [–]	g_2 [–]	Q_1 [kW^{-1}]	Q_2 [kW^{-1}]	T_s [K^{-1}]
r_1, hall-1 temperature	reading	−11.179	0	0.3643	0	2.000
r_2, hall-2 temperature	reading	0	−10.965	0	0.4932	2.000
r_3, supply temperature	reading	0	0	0	0	3.333
r_4, chiller heat meter	reading	0	0	0.500	0.500	0
r_5, hall-1 load meter	reading	0	0	0.6667	0	0
r_6, prior on g_1	prior	3.333	0	0	0	0
r_7, prior on g_2	prior	0	3.333	0	0	0
r_8, prior on T_s	prior	0	0	0	0	1.000

Table 19.3: The posterior covariance $\mathbf{\Sigma} = \mathbf{\Lambda}^{-1}$ of the two halls at the MAP state (diagonal and below) and the correlations (above the diagonal, in italics). Units of the covariances: products of the entries' units (the log-conductances dimensionless, the loads in kW, the supply temperature in K); for instance $\Sigma_{Q_2Q_2}$ in kW^2 and $\Sigma_{Q_2g_2}$ in kW. The correlations have no units.

Entry	g_1	g_2	Q_1	Q_2	T_s
g_1	0.011382	*−0.029*	*0.410*	*−0.251*	*0.434*
g_2	−0.000417	0.018765	*−0.401*	*0.691*	*0.313*
Q_1	0.06358	−0.07977	2.1085	*−0.572*	*0.007*
Q_2	−0.06274	0.22219	−1.9490	5.5048	*−0.041*
T_s	0.012974	0.011989	0.002914	−0.026883	0.078433

The posterior on the outage record. The fifth question of Table 19.1 is asked of the record $\mathbf{y}'$, the same construction with the rows r_4 and r_5 of the two meters left out: six rows on the same five entries, fitted by the same iteration with $\mathbf{\Lambda}' = \mathbf{J}'^{\mathsf{T}}\mathbf{J}'$ and $\mathbf{\Sigma}' = \mathbf{\Lambda}'^{-1}$. Each load then appears in its hall's row alone and the fit meets both hall readings exactly, so that the conductances stay at the prior mean and the loads follow, $\hat{Q}'_i = \mathrm{e}^{\hat{g}'_i}(y_{T_i} - \hat{T}'_s)$:

$$\hat{\mathbf{z}}' = \begin{bmatrix} 1.6094 \\ 1.6094 \\ 27.816 \\ 28.041 \\ 12.092 \end{bmatrix}, \qquad \sqrt{\operatorname{diag} \mathbf{\Sigma}'} = \begin{bmatrix} 0.300 \\ 0.300 \\ 8.829 \\ 8.893 \\ 0.2873 \end{bmatrix}, \tag{19.14}$$

in the order and the units of equation (19.13), with $\Sigma'_{Q_2Q_2} = 79.08\,\mathrm{kW}^2$.

The pair $(\hat{\mathbf{z}}, \mathbf{\Sigma})$ is what step 4 uses for the first four questions, and $(\hat{\mathbf{z}}', \mathbf{\Sigma}')$ for the fifth: every answer of the chapter is computed from the posterior of the state. Where the exact posterior has been computed, by the weighted draws of Section 19.2.1 or by the one-dimensional integral of the fifth question, its value is set beside the answer, with the gap.

19.3 The answers (step 4)

The five questions of Table 19.1 fall into three categories of Chapter 2 and take four of its general forms: estimation, of a moment of an entry or of a combination of entries (the first two) and of the posterior set against the prior (the third); design, run in reverse as the removal of a row (the fourth); and prediction under a threshold, a tail of the posterior (the fifth). Each is answered from the posterior of Section 19.2: its engineering question, the general form of its category, the form it takes for the two halls, the quantities of the posterior it uses, and the plant's numbers. The questions name the meters removed by the fourth and the level 45 kW of the fifth. Every answer is computed from the MAP state and the covariance of step 3, and that is the value the chapter reports; the exact posterior gives the value set beside it, with the gap.

19.3.1 The spreads, the correlations and the weakest direction

Engineering questions. How well are the conductances, the loads and the supply temperature known? Which direction of the state do the readings determine least? The first asks for the spread of each entry and for the correlations of the pairs, which say which combinations the readings fix better than either entry; the second for the combination of entries that is worst known, the one a decision that rests on it must allow for.

General form. Estimation of a derived quantity, summarized by its location and spread, equation (2.10), and its central interval, equation (2.11); the spread is the variance

$$\operatorname{Var}(f(\mathbf{z}) \mid \mathbf{y}) = \int \left(f(\mathbf{z}) - \mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]\right)^2 p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}. \tag{19.15}$$

Specialized form. The questions are equations (19.15), (2.10) and (2.11) with the state $\mathbf{z}$ the five entries of $\hat{\mathbf{z}}$, the derived quantity $f(\mathbf{z})$ in turn an entry z_i, a linear combination $\mathbf{a}^{\mathsf{T}}\mathbf{z}$ and the combination $\mathbf{v}^{\mathsf{T}}\mathbf{z}$ along a unit vector, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \mathbf{\Sigma})$ of Section 19.2.2. One choice enters here. Directions are measured in the units of the state, the log-conductances dimensionless, the loads in kW and the temperature in K: this makes the singular values comparable within the chapter, units in which the prior and meter spreads all lie between 0.3 and 2, and it rules out any comparison across choices of units, the mistake that Section 19.6.5 names.

Under the Laplace posterior every entry is normal, $z_i \mid \mathbf{y} \approx \mathcal{N}(\hat{z}_i, \Sigma_{ii})$: the spread of entry i is $\sqrt{\Sigma_{ii}}$ and its central 95 % band $\hat{z}_i \pm 1.96\sqrt{\Sigma_{ii}}$, which for a conductance maps to $G_i \in [\mathrm{e}^{\hat{g}_i - 1.96\sqrt{\Sigma_{ii}}},\ \mathrm{e}^{\hat{g}_i + 1.96\sqrt{\Sigma_{ii}}}]$ in kW/K. The correlation of two entries is $\Sigma_{ij}/\sqrt{\Sigma_{ii}\Sigma_{jj}}$, in $[-1, 1]$ and free of units. A linear combination $f = \mathbf{a}^\mathsf{T}\mathbf{z}$, such as the total load $Q_1 + Q_2$, is a derived quantity with the gradient $\mathbf{g} = \mathbf{a}$, and its variance is $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$, exactly under the Laplace posterior. Two entries together have the two-dimensional normal marginal with the 2×2 block $\boldsymbol{\Sigma}_{2\times 2}$ of $\boldsymbol{\Sigma}$, whose contour holding the probability p is the ellipse $\boldsymbol{\delta}^\mathsf{T}\boldsymbol{\Sigma}_{2\times 2}^{-1}\boldsymbol{\delta} = \chi_p^2$ about the MAP state, with the radius χ_p given by $\chi_p^2 = -2\ln(1-p)$, the p-quantile of the χ^2 distribution with two degrees of freedom.

Which pairs are correlated is read off the Jacobian. A direction $\mathbf{b}$ of the state whose image $\mathbf{J}\mathbf{b}$ has zero or small entries in some rows is one that those rows cannot see; the objective rises slowly along it, the posterior is long along it, and the entries that $\mathbf{b}$ moves together are correlated with the sign of their product in $\mathbf{b}$. A pair of columns that meet the same row with entries in a fixed ratio is the signature: a common scaling of load and conductance leaves a hall reading unchanged, and a transfer of load from one hall to the other leaves the chiller meter unchanged.

Marginal against conditional. The marginal spread is the uncertainty of entry i when every other entry is also unknown; the conditional spread is its uncertainty if the other entries were known exactly. The first is what a report on entry i owes; the second is what the readings would give if the rest of the plant were pinned. Both are read off the Laplace posterior with the block-inversion formula of Section 3.5.3, which Chapter 3 derived for any partition of a Gaussian state, applied to the partition into entry i and the rest r. Order the state as $(z_i, \mathbf{z}_r)$ and write the information matrix in the same blocks,

$$\boldsymbol{\Lambda} = \begin{bmatrix} \Lambda_{ii} & \boldsymbol{\lambda}_{ri}^\mathsf{T} \\ \boldsymbol{\lambda}_{ri} & \boldsymbol{\Lambda}_{rr} \end{bmatrix}, \tag{19.16}$$

with Λ_{ii} a number, $\boldsymbol{\lambda}_{ri}$ the column of couplings of entry i to the rest, and $\boldsymbol{\Lambda}_{rr}$ the block of the rest. The precision of a Gaussian conditional on a block is the corresponding block of the precision, so with $\mathbf{z}_r$ held fixed

$$\mathrm{Var}(z_i \mid \mathbf{z}_r, \mathbf{y}) = \frac{1}{\Lambda_{ii}}, \qquad \mathbb{E}[z_i \mid \mathbf{z}_r, \mathbf{y}] = \hat{z}_i - \frac{\boldsymbol{\lambda}_{ri}^\mathsf{T}(\mathbf{z}_r - \hat{\mathbf{z}}_r)}{\Lambda_{ii}}, \tag{19.17}$$

read straight off the diagonal of the curvature, with no inversion. The precision of a marginal is instead the Schur complement of the block integrated out, so

$$\mathrm{Var}(z_i \mid \mathbf{y}) = \Sigma_{ii} = \frac{1}{\Lambda_{ii} - \boldsymbol{\lambda}_{ri}^\mathsf{T}\boldsymbol{\Lambda}_{rr}^{-1}\boldsymbol{\lambda}_{ri}}, \tag{19.18}$$

the (i,i) entry of $\boldsymbol{\Lambda}^{-1}$ written without inverting the whole matrix: the marginal precision of entry i is its own curvature less the part the other entries can absorb. Chapter 7 met the same identity for two entries as the curvature of the profile likelihood in the log-shape, and Exercise 19.2 writes it for a block of several entries. Because $\boldsymbol{\Lambda}_{rr}$ is positive definite, the subtracted term is nonnegative, so

$$\Sigma_{ii} \geq \frac{1}{\Lambda_{ii}}, \qquad \Sigma_{ii}\,\Lambda_{ii} = \frac{1}{1-\varrho_i^2}, \qquad \varrho_i^2 = \frac{\boldsymbol{\lambda}_{ri}^\mathsf{T}\boldsymbol{\Lambda}_{rr}^{-1}\boldsymbol{\lambda}_{ri}}{\Lambda_{ii}} = 1 - \frac{\mathrm{Var}(z_i \mid \mathbf{z}_r, \mathbf{y})}{\mathrm{Var}(z_i \mid \mathbf{y})}, \tag{19.19}$$

with equality only when entry i is uncoupled from the rest, $\boldsymbol{\lambda}_{ri} = \mathbf{0}$. The number ϱ_i^2 is the *squared multiple correlation* of entry i with the other entries under the posterior: not the correlation of two entries, but the largest squared correlation between z_i and any linear combination $\mathbf{a}^\mathsf{T}\mathbf{z}_r$ of the rest, attained by the combination that predicts z_i best, and equal to

the fraction of the marginal variance of entry i that the rest explain. The conditional spread understates the marginal by the factor $\sqrt{1-\varrho_i^2}$, and the difference is the whole effect of the correlations: an entry whose readings also involve other unknown entries is known less well than its own curvature suggests. For a single other entry, ϱ_i^2 is the square of their correlation.

The weakest direction. The marginal and conditional spreads concern the axes of the state. The curvature also has directions of its own, and they are read from the singular value decomposition of the normalized Jacobian, a standard factorization of any $m \times n$ matrix with $m \geq n$:

$$\mathbf{J} = \mathbf{F}\,\mathrm{diag}(s_1, \ldots, s_n)\,\mathbf{V}^{\mathsf{T}}, \qquad s_1 \geq s_2 \geq \cdots \geq s_n \geq 0, \tag{19.20}$$

where $\mathbf{F}$ is $m \times n$ with orthonormal columns, $\mathbf{V}$ is $n \times n$ orthogonal with columns $\mathbf{v}_j$, the right singular vectors, and the s_j are the singular values. Since $\mathbf{F}^{\mathsf{T}}\mathbf{F}$ is the identity,

$$\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J} = \mathbf{V}\,\mathrm{diag}(s_1^2, \ldots, s_n^2)\,\mathbf{V}^{\mathsf{T}}, \qquad \mathbf{\Sigma} = \mathbf{\Lambda}^{-1} = \mathbf{V}\,\mathrm{diag}(s_1^{-2}, \ldots, s_n^{-2})\,\mathbf{V}^{\mathsf{T}}. \tag{19.21}$$

The right singular vectors are the eigenvectors of the curvature and of $\mathbf{\Sigma}$, and the squared singular values the eigenvalues of the curvature. The combination of the entries along a unit vector $\mathbf{v}$ is the derived quantity $\mathbf{v}^{\mathsf{T}}\mathbf{z}$, with the gradient $\mathbf{v}$ and the variance $\mathbf{v}^{\mathsf{T}}\mathbf{\Sigma}\mathbf{v}$. Along a right singular vector, with $\mathbf{V}^{\mathsf{T}}\mathbf{v}_j$ the j-th unit vector,

$$\mathrm{Var}(\mathbf{v}_j^{\mathsf{T}}\mathbf{z} \mid \mathbf{y}) = \mathbf{v}_j^{\mathsf{T}}\mathbf{\Sigma}\,\mathbf{v}_j = \frac{1}{s_j^2}, \qquad \mathcal{L}(\hat{\mathbf{z}} + t\,\mathbf{v}_j) \approx \mathcal{L}(\hat{\mathbf{z}}) + \tfrac{1}{2}s_j^2 t^2. \tag{19.22}$$

The spread along $\mathbf{v}_j$ is $1/s_j$, and the objective rises by one half when the state moves $1/s_j$ along it. The variance $\mathbf{v}^{\mathsf{T}}\mathbf{\Sigma}\mathbf{v}$ over unit vectors is largest along $\mathbf{v}_n$, so the readings determine worst the combination $\mathbf{v}_n^{\mathsf{T}}\mathbf{z}$, and best $\mathbf{v}_1^{\mathsf{T}}\mathbf{z}$. The ratio of the widest to the narrowest spread is the condition number $\mathrm{cond}(\mathbf{J}) = s_1/s_n$; the curvature has $\mathrm{cond}(\mathbf{\Lambda}) = \mathrm{cond}(\mathbf{J})^2$. Entry by entry the same decomposition splits each variance over the directions (Exercise 19.4).

The columns of $\mathbf{J}$ carry the inverse units of the entries, and the choice of units named above is what the decomposition depends on. Expressing the loads in watts instead of kilowatts divides the two load columns by a thousand: the singular values on the load directions fall a thousandfold, the condition number rises a thousandfold, and nothing physical has changed. Singular values and singular vectors are therefore compared only within one choice of units, and the rank, which does not depend on that choice, is the unit-free part of the analysis; the companion volume scales each column by a reference value of its parameter to the same end.

Building blocks. The MAP state $\hat{\mathbf{z}}$ of Chapter 18; the Jacobian at $\hat{\mathbf{z}}$ from the derivatives of Section 19.1.6; $\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$ and $\mathbf{\Sigma} = \mathbf{\Lambda}^{-1}$ (Section 19.2.1); the Schur quadratic form of equation (19.18); the gradients of the combinations; and the singular value decomposition of $\mathbf{J}$, all at the numbers of Section 19.2.2. For the exact values beside them, the weighted draws of Section 19.2.1.

The spreads and the correlations. Table 19.4 sets the spreads of each entry side by side. Each row is an entry; the columns give its MAP value, its marginal standard deviation $\sqrt{\Sigma_{ii}}$, its conditional standard deviation $\Lambda_{ii}^{-1/2}$ with the other four entries held fixed, the ratio of the two, the squared multiple correlation ϱ_i^2 of equation (19.19), the standard deviation the readings alone would give (the covariance $(\mathbf{J}_{\mathrm{d}}^{\mathsf{T}}\mathbf{J}_{\mathrm{d}})^{-1}$ without the priors, Section 19.3.2), the prior standard deviation, and the ratio of prior to posterior spread.

A worked Schur complement. For the load of hall 2, the fourth entry, the diagonal of equation (19.13) gives $\Lambda_{44} = 0.4933\,\mathrm{kW}^{-2}$ and the conditional standard deviation $1/\sqrt{0.4933} = 1.424\,\mathrm{kW}$. Its couplings to the other four entries (g_1, g_2, Q_1, T_s) are $\boldsymbol{\lambda}_{r4} =$

Table 19.4: The spreads of the five entries at the MAP state: marginal $\sqrt{\Sigma_{ii}}$ from Table 19.3, conditional $\Lambda_{ii}^{-1/2}$ from equation (19.13), their ratio, the squared multiple correlation, the readings-only spread, the prior spread $\sigma_{0,i}$ and the ratio $\sigma_{0,i}/\sqrt{\Sigma_{ii}}$. Spreads in the units of the entry: dimensionless, kW, K.

Entry	MAP	Marginal sd	Conditional sd	Ratio [–]	ϱ_i^2 [–]	Readings only	Prior sd	Prior / posterior [–]
g_1	1.7030	0.1067	0.0857	1.245	0.354	0.1152	0.3	2.81
g_2	1.3999	0.1370	0.0873	1.570	0.594	0.1548	0.3	2.19
Q_1 [kW]	30.688	1.452	1.100	1.321	0.427	1.500	flat	∞
Q_2 [kW]	22.231	2.346	1.424	1.648	0.632	2.500	flat	∞
T_s [°C]	12.111	0.2801	0.2230	1.256	0.366	0.300	1.0	3.57

$(0,\ -5.4081,\ 0.2500,\ 0.9865)$, through the hall-2 reading to g_2 and T_s and through the chiller meter to Q_1. The quadratic form of equation (19.18) with the inverse of the four-by-four block of the rest is $\boldsymbol{\lambda}_{r4}^{\mathsf{T}}\boldsymbol{\Lambda}_{rr}^{-1}\boldsymbol{\lambda}_{r4} = 0.3116$, so

$$\Sigma_{44} = \frac{1}{0.4933 - 0.3116} = \frac{1}{0.1817} = 5.505\,\mathrm{kW}^2, \quad \sqrt{\Sigma_{44}} = 2.346\,\mathrm{kW}, \quad \varrho_4^2 = \frac{0.3116}{0.4933} = 0.632,$$

the entry of Table 19.3: with the other entries known the spread would be 1.42 kW, and with them unknown it is 2.35 kW.

The spreads. The conductances are known to ± 0.107 and ± 0.137 in the log, about 11 % and 14 %, with the 95 % bands $G_1 \in [4.45,\ 6.77]$ and $G_2 \in [3.10,\ 5.30]\,\mathrm{kW/K}$; the loads to 1.45 and 2.35 kW, a little better than the meters alone give, 1.5 and $\sqrt{2^2 + 1.5^2} = 2.5\,\mathrm{kW}$; the supply temperature to 0.28 K, against its sensor's 0.3 K. The squared multiple correlations run from 0.35 for g_1 to 0.63 for Q_2. The values beside them come from the exact posterior, by the weighted draws: the spreads (0.108, 0.139, 1.449 kW, 2.315 kW, 0.281 K), the largest gap 0.031 kW on Q_2, and for g_1 the band $[1.513,\ 1.938]$ against the Laplace $[1.494,\ 1.912]$, 0.02 lower at either end, the long side of large conductances.

Load and conductance of hall 2, +0.69. The columns of g_2 and Q_2 meet the hall-2 row as $-\Delta T_2/\sigma_T = -10.965$ and $1/(G_2\sigma_T) = 0.4932\,\mathrm{kW}^{-1}$. The direction $\mathbf{b} = (0,\ 1,\ 0,\ \hat{Q}_2,\ 0)^{\mathsf{T}}$, which raises the log-conductance by one unit and the load by $\hat{Q}_2 = 22.23\,\mathrm{kW}$, a common scaling of load and conductance, therefore has

$$\mathbf{Jb} = (0,\ 0,\ 0,\ 11.12,\ 0,\ 0,\ 3.333,\ 0)^{\mathsf{T}} :$$

the hall-2 reading is blind to it, and only the chiller meter (row 4) and the prior on g_2 (row 7) hold it. Hall 1 has the same direction, $(1,\ 0,\ \hat{Q}_1,\ 0,\ 0)^{\mathsf{T}}$, with the image $(0,\ 0,\ 0,\ 15.34,\ 20.46,\ 3.333,\ 0,\ 0)^{\mathsf{T}}$: its own load meter (row 5) also holds it, and the correlation is weaker, +0.41.

The two loads, −0.57. The columns of Q_1 and Q_2 meet the chiller meter's row as 0.5 and 0.5, so the direction $(0,\ 0,\ 1,\ -1,\ 0)^{\mathsf{T}}$ leaves row 4 unchanged and is seen only by the load meter and, weakly, by the hall rows. With the gradients $\mathbf{g}_+ = (0,\ 0,\ 1,\ 1,\ 0)^{\mathsf{T}}$ and $\mathbf{g}_- = (0,\ 0,\ 1,\ -1,\ 0)^{\mathsf{T}}$, the total has the spread $\sqrt{\mathbf{g}_+^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}_+} = 1.93\,\mathrm{kW}$, below the meter's own 2 kW, and the difference $\sqrt{\mathbf{g}_-^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}_-} = 3.39\,\mathrm{kW}$.

Conductance and supply temperature, +0.43 and +0.31. The supply temperature's column meets both hall rows with 2, and the direction and its image are

$$\mathbf{b} = (1/\Delta T_1,\ 1/\Delta T_2,\ 0,\ 0,\ 1)^{\mathsf{T}} = (0.179,\ 0.182,\ 0,\ 0,\ 1)^{\mathsf{T}},$$

$$\mathbf{Jb} = (0,\ 0,\ 3.333,\ 0,\ 0,\ 0.596,\ 0.608,\ 1)^{\mathsf{T}} :$$

both hall readings are blind to it, and only the supply sensor and the priors see it. The correlation is weaker in hall 2, whose load column, held only by the chiller meter, absorbs part of a change of T_s in place of g_2.

The chains. The remaining correlations combine these directions: g_2 and Q_1 correlate at -0.40 through the split and the scaling of hall 2, and the two conductances at -0.03, a positive link through T_s and a negative one through the loads nearly cancelling.

Figure 19.1 draws two of these pairs. In each panel the axes are two entries of the state; the closed curves are the contours of their two-dimensional marginal under the Laplace approximation that hold 50 % (dashed) and 95 % (solid) of the probability, of radius χ_p; the dot is the MAP state. The thick green bars along the bottom and the left side are the marginal 95 % bands of the two entries, $\hat{z}_i \pm 1.96\sqrt{\Sigma_{ii}}$; the red segment through the center is the 95 % band of the vertical entry with the horizontal one held at its estimate. Panel (a) shows the loads: the ellipse is tilted down, the negative correlation, and the band of Q_2 given Q_1 is shorter than its marginal band by the factor $\sqrt{1-0.572^2} = 0.82$. Panel (b) shows the load and the conductance of hall 2: tilted up, and longer along the diagonal, the direction of a common scaling of load and conductance that the hall reading cannot see and only the chiller meter and the prior limit.

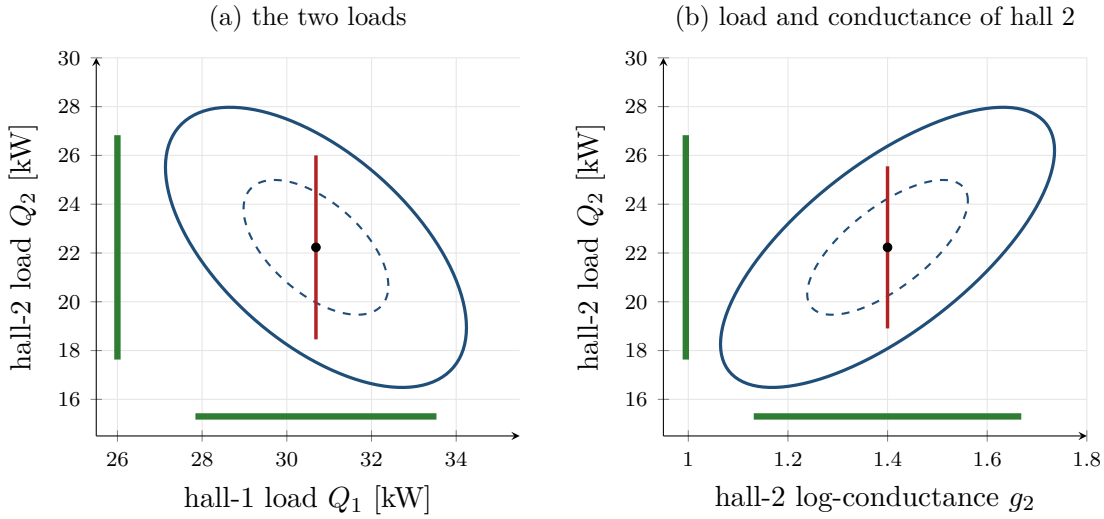


Figure 19.1: Two-dimensional marginals of the posterior of the two halls under the Laplace approximation, from Table 19.3. Ellipses: the contours holding 50 % (dashed, $\chi_p = 1.177$) and 95 % (solid, $\chi_p = 2.448$) of the probability, drawn as $\hat{\mathbf{z}} + \chi_p \mathbf{L}(\cos t,\ \sin t)^{\mathsf{T}}$ with $\mathbf{L}$ the Cholesky factor of the two-by-two block; dot: the MAP state. Green bars: marginal 95 % bands, (a) $Q_1 \in [27.84,\ 33.53]$ and $Q_2 \in [17.63,\ 26.83]$ kW, (b) $g_2 \in [1.131,\ 1.668]$ and the same band of Q_2. Red segment: the 95 % band of Q_2 given the horizontal entry at its estimate, (a) $[18.46,\ 26.00]$ kW, (b) $[18.91,\ 25.55]$ kW.

The singular directions. Table 19.5 lists the singular value decomposition (19.20) of the normalized Jacobian of Table 19.2. Each row is one direction; the columns give its singular value s_j, the spread $1/s_j$ along it by equation (19.22), the singular value of the design columns alone

(without the prior rows), the five entries of the right singular vector $\mathbf{v}_j$, and the combination it describes. Each vector has unit length in the units of the state and is determined up to its sign, which is chosen here so that its largest entry is positive, except the last, which is written so that it adds load to hall 1.

Table 19.5: The singular value decomposition (19.20) of the normalized Jacobian at the MAP state, with priors (s_j) and without (s_j of $\mathbf{J}_\mathrm{d}$), and the right singular vectors in the units of the state (log-conductances dimensionless, loads in kW, supply temperature in K).

j	s_j	$1/s_j$	s_j of $\mathbf{J}_\mathrm{d}$	g_1	g_2	Q_1	Q_2	T_s	Combination
1	11.940	0.0838	11.494	0.844	0.478	−0.026	−0.020	−0.240	both conductances, against T_s
2	11.544	0.0866	11.055	−0.506	0.859	0.015	−0.035	−0.067	contrast of the conductances
3	3.471	0.288	3.221	0.175	0.178	0.000	−0.001	0.968	T_s with both conductances
4	0.904	1.106	0.898	0.026	0.016	0.910	0.414	−0.007	the loads together
5	0.395	2.530	0.371	0.013	−0.037	0.414	−0.910	0.004	the split of the load

Reading the directions. The first two directions are the conductances, their combination and their contrast, each determined to about 0.085 in the log. The third is the supply temperature with both conductances in the proportion $0.18 \approx 1/\Delta T_i$ per kelvin, the direction of Section 19.3.1 that both hall rows are blind to, with the spread 0.288; the conductance priors supply about 6 % of s_3^2 ($11.11 \times (0.175^2 + 0.178^2) = 0.69$ of 12.05). The last two are the loads: their sum weighted toward hall 1, with the spread 1.11 kW, and $\mathbf{v}_5 = (0.01,\ -0.04,\ 0.41,\ -0.91,\ 0.00)^\mathsf{T}$, the split of the load weighted toward hall 2, the hall without a meter of its own, with the spread 2.53 kW, the combination the readings determine least. The condition number is $\mathrm{cond}(\mathbf{J}) = 11.940/0.395 = 30.2$, in these units. The design columns alone have singular values only slightly smaller, 11.49 down to 0.371.

Figure 19.2 draws the decomposition. Panel (a) has the index j of the direction on the horizontal axis and the singular value on a logarithmic vertical axis, with the prior rows (filled circles) and without them (open squares). Panel (b) draws the entries of the two weakest right singular vectors as bars over the five entries of the state. The singular values fall in two steps, a factor of three from the conductances to the supply temperature and a factor of four more to the loads, and the last two vectors have their weight almost entirely on Q_1 and Q_2.

The product $\mathbf{V}\,\mathrm{diag}(s_j^{-2})\mathbf{V}^\mathsf{T}$ reproduces Table 19.3 to 10^{-15}.

19.3.2 Observability and the ratio of spreads

Engineering question. Can my sensors see this parameter at all? The owner asks it of each of the five entries, conductances included, and the answer must hold whatever the readings happen to show.

General form. Estimation, the posterior distribution of a derived quantity, equation (2.9), set against the same distribution under the prior alone:

$$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) \le f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z} \quad \text{against}$$

$$\Pr(f(\mathbf{z}) \le f^*) = \int \mathbf{1}\big[f(\mathbf{z}) \le f^*\big]\, p(\mathbf{z})\, \mathrm{d}\mathbf{z};$$

where the two agree for every f^*, the readings have told nothing about f.

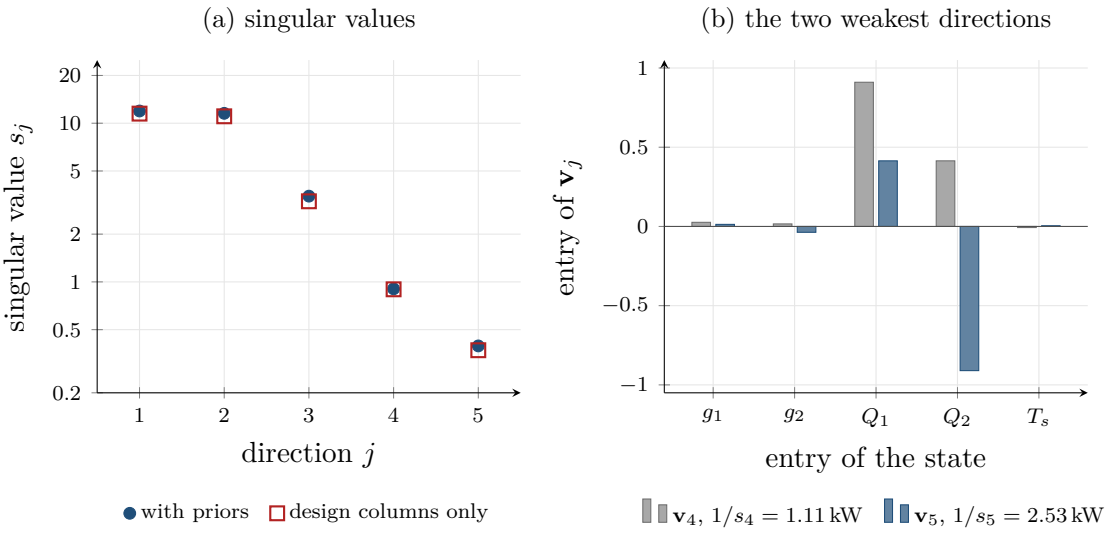


Figure 19.2: The singular value decomposition of the two halls' normalized Jacobian, Table 19.5. (a) The five singular values with the prior rows and for the design columns alone, logarithmic scale. (b) The entries of the two weakest right singular vectors: $\mathbf{v}_4$, the loads together, and $\mathbf{v}_5$, the split of the load between the halls.

Specialized form. The question is equation (2.9) with the state $\mathbf{z}$ the five entries, $f(\mathbf{z}) = z_i$, f^* every level of it, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 19.2.2, set against its prior. Two choices enter here. The test is made on the rows linearized at the MAP state: this buys a test that is a rank, decided by the design before any value is read, and it fails where a row is strongly nonlinear across an entry's own spread, so that the exact marginal moves away from the prior even along a direction the linearized readings cannot see, which Section 19.4.2 shows on this plant. The loads keep their flat prior: this is what lets the readings alone fix them, and it makes the ratio of prior to posterior spread uninformative for them, so that for the loads the rank of the design columns is the test.

Here $f(\mathbf{z}) = z_i$, and the design columns $\mathbf{J}_\mathrm{d}$ of equation (19.5) describe what the readings alone can say. A direction $\mathbf{b}$ with

$$\mathbf{J}_\mathrm{d}\,\mathbf{b} = \mathbf{0} \tag{19.23}$$

leaves every reading unchanged to first order when the state moves along it: no reading can distinguish $\hat{\mathbf{z}}$ from $\hat{\mathbf{z}} + t\mathbf{b}$. Such a direction is *unobservable*, and the readings determine the state completely only when there is none, that is when $\mathbf{J}_\mathrm{d}$ has rank n: the test of numerical observability analysis, applied here to every entry of the state, conductances included. The curvature along an unobservable direction comes from the priors alone, $\mathbf{b}^\mathsf{T}\boldsymbol{\Lambda}\mathbf{b} = \mathbf{b}^\mathsf{T}\boldsymbol{\Lambda}_\mathrm{prior}\mathbf{b}$ by equation (19.5), and if the priors are flat along it the curvature is singular and the posterior along $\mathbf{b}$ is as wide as the prior, without limit. A zero singular value of the design columns is the linearized test of *structural* identifiability: the direction is not identified by any noise-free readings of this design. A direction with a small but nonzero singular value is identified in principle, and its *practical* identifiability is judged by its spread $1/s_j$ of equation (19.22) against the precision the decision at hand needs.

When an entry keeps its prior. The rank says whether directions are lost; the entry-by-entry question needs one step more. Suppose an unobservable direction $\mathbf{b}$ moves entry i, $b_i = 1$, and otherwise moves only entries whose prior is flat, $b_j = 0$ for every other entry j with a proper prior. Change variables from $\mathbf{z}$ to z_i and the remaining coordinates $z'_j = z_j - z_i b_j$ for $j \neq i$, so that $\mathbf{z} = \mathbf{z}' + z_i\mathbf{b}$ with $\mathbf{z}'$ the vector of the z'_j and a zero in position i; the map is triangular with ones on its diagonal, so its change-of-variable factor is one. In the linearized

rows $\mathbf{J}_\mathrm{d}\mathbf{z} = \mathbf{J}_\mathrm{d}\mathbf{z}' + z_i\mathbf{J}_\mathrm{d}\mathbf{b} = \mathbf{J}_\mathrm{d}\mathbf{z}'$ by equation (19.23), so the reading rows depend on $\mathbf{z}'$ only; the prior on z_i depends on z_i only; every other entry with a proper prior has $b_j = 0$, so $z_j = z'_j$ and its prior depends on $\mathbf{z}'$ only; and the flat priors are constant. The posterior factorizes,

$$p(\mathbf{z} \mid \mathbf{y}) \propto p(z_i)\; h(\mathbf{z}'), \qquad p(z_i \mid \mathbf{y}) = p(z_i), \tag{19.24}$$

for a function h that collects the readings and the remaining priors: under the linearized rows of the Laplace approximation the posterior of entry i is its prior, and its posterior spread equals its prior spread. With the rows kept nonlinear the factorization is approximate.

The ratio of spreads. The practical test runs the other way, from the spreads to the diagnosis. Since a posterior is built from the prior by adding reading rows, and no row can increase a variance (Exercise 19.3), the ratio

$$\frac{\sigma_{0,i}}{\sqrt{\Sigma_{ii}}} \geq 1 \tag{19.25}$$

holds for every entry, and a ratio of one says that the readings have told nothing about entry i: the entry is unobserved, whatever value the fit returns for it. A ratio well above one says that the readings add information, not that they alone hold the entry. For an entry with a flat prior, $\sigma_{0,i} = \infty$, the ratio is infinite whenever the posterior is proper and certifies nothing: the entry may be held by the priors of other entries through an unobservable direction. For such an entry the test is the rank of the design columns, or its spread under the readings-only covariance $(\mathbf{J}_\mathrm{d}^\mathsf{T}\mathbf{J}_\mathrm{d})^{-1}$, which exists only when that rank is full.

> **Principle 19.1.** *An entry with a proper prior whose posterior spread equals its prior spread is unobserved: the readings have told nothing about it. A direction along which the design columns $\mathbf{J}_\mathrm{d}$, the reading rows alone, have a zero singular value is structurally unidentifiable: no reading changes when the state moves along it, and only a prior can hold it. A direction with a small singular value s_j of the full Jacobian $\mathbf{J}$, priors included, has the spread $1/s_j$, and it is weakly identified when that spread exceeds the precision the decision needs. Identifiability is read from these directions and ratios, the shape of the posterior, never from the marginal spreads alone.*

Building blocks. The design columns $\mathbf{J}_\mathrm{d}$ at $\hat{\mathbf{z}}$, their singular values, rank and null vectors; the readings-only covariance $(\mathbf{J}_\mathrm{d}^\mathsf{T}\mathbf{J}_\mathrm{d})^{-1}$; the spreads $\sqrt{\Sigma_{ii}}$ and the prior spreads $\sigma_{0,i}$.

The rank of the design columns. The design columns of the two halls, the first five rows of Table 19.2, form a square five-by-five matrix with determinant -136.2 in the inverse units of the entries; it is nonsingular, and $\mathbf{J}_\mathrm{d}$ has full rank five, with the smallest singular value 0.371. Every entry is observable from the readings without any prior. The covariance the readings alone would give, $(\mathbf{J}_\mathrm{d}^\mathsf{T}\mathbf{J}_\mathrm{d})^{-1}$, has the standard deviations listed in the readings-only column of Table 19.4, and with as many readings as entries each entry is fixed by one combination of them: T_s by its sensor, 0.300 K; Q_1 by its meter, 1.500 kW; Q_2 by the chiller meter less the load meter, $\sqrt{2^2 + 1.5^2} = 2.500$ kW; and each log-conductance by its hall reading given the rest, 0.115 and 0.155.

The ratios of prior to posterior spread are 2.81 and 2.19 for the conductances, 3.57 for the supply temperature and infinite for the loads. With five readings and five entries the readings have no redundancy among themselves: the χ^2 check of Chapter 18 rests on the three prior rows, and the loss of any reading leaves fewer readings than entries.

19.3.3 A meter taken out of the design

Engineering question. What is lost if the hall-1 load meter fails? The chiller meter too? The supply sensor? Each removes a reading row, or two, and the engineer wants the spreads, the entries that fall back on their priors and the combinations no longer seen.

General form. Design, the preposterior expectation of a quantity u of the posterior that a change a of the design would produce, equation (2.16):

$$U(a) = \int u(p(\mathbf{z} \mid \mathbf{y}, \tilde{y}_a))\, p(\tilde{y}_a \mid \mathbf{y})\, \mathrm{d}\tilde{y}_a,$$

here with a the removal of a set of rows and u the spread of an entry.

Specialized form. The question is equation (2.16) with a the removal of a set of rows, u the spread of entry i, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \mathbf{\Sigma})$ of Section 19.2.2, with the rows taken out of its curvature. Two choices enter here. The removal is evaluated at the same MAP state, without a new fit: this buys the answer before the meter fails, since the curvature at a fixed state depends on which rows exist and not on the values they carry, and it is approximate because a refit after a real failure moves the MAP state and, through the nonlinearity of the hall rows, the spreads with it; Section 19.3.3 gives the refitted values where they differ. The meters' errors are taken as independent: this is what lets a removal take out one row and nothing else, and it fails if two meters share a calibration or a transducer.

A removal brings no new reading, so the design form has no outcome to average over: $U(a)$ is u of the posterior without the rows. Each reading row is the linear Gaussian row of the design form of Chapter 2, $(\mathbf{h}_k\mathbf{z} - y_k)/\sigma_k$ with the Jacobian row $\mathbf{J}_k = \mathbf{h}_k/\sigma_k$ of equation (2.17) (a hall row linearized at $\hat{\mathbf{z}}$), and the covariance it leaves does not depend on the value it carried. With $\mathbf{j}_k = \mathbf{J}_k^{\mathsf{T}}$, row k of $\mathbf{J}$ written as a column, and a the set of rows removed,

$$\mathbf{\Lambda}_{-a} = \mathbf{\Lambda} - \sum_{k \in a} \mathbf{j}_k \mathbf{j}_k^{\mathsf{T}}, \qquad \mathbf{\Sigma}_{-a} = \mathbf{\Lambda}_{-a}^{-1}, \qquad u(a) = \sqrt{(\mathbf{\Sigma}_{-a})_{ii}}, \tag{19.26}$$

the curvature without those rows at the same $\hat{\mathbf{z}}$, the sum of row outer products of Section 19.2.1 with the removed terms left out. The covariance $\mathbf{\Sigma}_{-a}$ exists only while $\mathbf{\Lambda}_{-a}$ stays positive definite; when the removal leaves a direction that neither the remaining readings nor a prior hold, it has none. The directions the removal loses are the null vectors of the remaining design columns, $\mathbf{J}_{\mathrm{d},-a}\mathbf{b} = \mathbf{0}$, the test (19.23) on the shrunk design; an entry that a lost direction moves together with flat-prior entries only returns to its prior by the factorization (19.24). The weak directions of the full design announce the losses: the angle between a lost direction $\mathbf{b}$ and a right singular vector $\mathbf{v}_j$ of the full Jacobian is $\arccos(|\mathbf{b}^{\mathsf{T}}\mathbf{v}_j|/\|\mathbf{b}\|)$, and a lost plane is compared with the plane of two singular vectors by the principal angles between them.

Building blocks. The rows $\mathbf{j}_k$ of the Jacobian at $\hat{\mathbf{z}}$; the curvature and covariance (19.26) for each removal; the singular values and null vectors of the remaining design columns; the angles to the singular vectors of Section 19.3.1; and, for comparison, the refit on the remaining rows.

Meters removed one at a time. Figure 19.3 draws the design for five cases. Each grid is the design-column incidence of the pattern (19.3): the five reading rows down, the five entries across, a filled cell where the reading depends on the entry. The rows a case removes are grayed. Under each grid are the case and the rank of the remaining design columns.

Table 19.6 gives the result of equation (19.26) for each case. Each row is a case; the columns give the rank of the remaining design columns, the posterior standard deviation of each entry with the priors, the entries whose spread has returned to its prior, and the directions lost, the null directions (19.23) of the remaining design columns.

The lost directions. Each is a vector $\mathbf{b}$ with $\mathbf{J}_{\mathrm{d}}\mathbf{b} = \mathbf{0}$ on the remaining design rows of Table 19.2. Hall i's row is blind to every $\mathbf{b}$ with $b_{T_s} = 0$ and $b_{Q_i} = \hat{Q}_i\, b_{g_i}$, the common

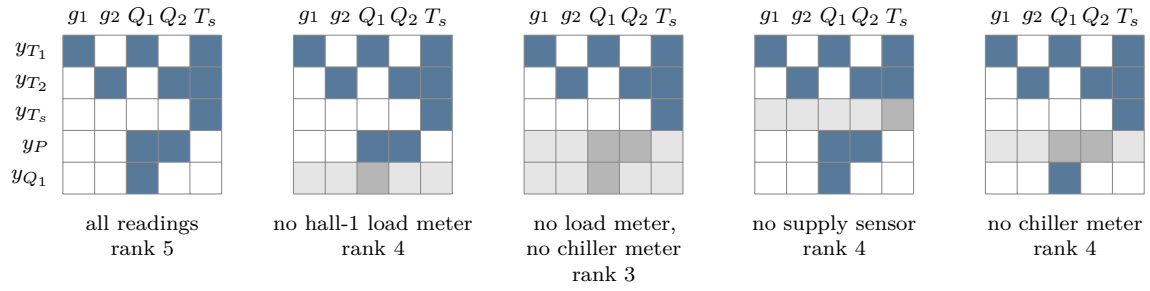


Figure 19.3: The design columns of the two halls in five cases. Rows: the readings of the hall-1 and hall-2 temperatures, the supply temperature, the chiller heat meter and the hall-1 load meter. Columns: the five entries of the state. Blue cell: the reading depends on the entry (nonzero in Table 19.2); gray row: removed. The rank of the remaining design columns is stated under each grid; five is full.

Table 19.6: The two halls with meters removed, evaluated at the MAP state (19.12): rank of the remaining design columns, posterior standard deviations with the priors, the entries back at their prior spread, and the lost directions in $(g_1, g_2, Q_1, Q_2, T_s)$, up to a common factor. Prior spreads: 0.3, 0.3, flat, flat, 1.0 K.

Case	Rank	g_1 [–]	g_2 [–]	Q_1 [kW]	Q_2 [kW]	T_s [K]	At prior	Lost directions
all readings	5	0.107	0.137	1.45	2.35	0.280	none	none
no hall-1 load meter	4	0.200	0.252	5.79	5.69	0.280	none	$(1/\hat{Q}_1,\ -1/\hat{Q}_2,\ 1,\ -1,\ 0)$
no load meter, no chiller meter	3	0.300	0.300	9.74	7.07	0.287	g_1, g_2	$(1/\hat{Q}_1,\ 0,\ 1,\ 0,\ 0)$ and $(0,\ 1/\hat{Q}_2,\ 0,\ 1,\ 0)$
no supply sensor	4	0.161	0.177	1.45	2.36	0.781	none	$(1/\Delta T_1,\ 1/\Delta T_2,\ 0,\ 0,\ 1)$
no chiller meter	4	0.107	0.300	1.48	7.06	0.284	g_2	$(0,\ 1/\hat{Q}_2,\ 0,\ 1,\ 0)$

scaling of Section 19.3.1. Without the load meter, the supply sensor still requires $b_{T_s} = 0$ and the chiller meter $b_{Q_1} + b_{Q_2} = 0$, which leaves $\mathbf{b} = (1/\hat{Q}_1,\ -1/\hat{Q}_2,\ 1,\ -1,\ 0)^{\mathsf{T}} = (0.0326,\ -0.0450,\ 1,\ -1,\ 0)^{\mathsf{T}}$ per kW, with

$$\mathbf{J}_{\mathrm{d}}\mathbf{b} = \big(-11.179 \times 0.0326 + 0.3643,\quad 10.965 \times 0.0450 - 0.4932,\quad 0,\quad 0.5 - 0.5\big)^{\mathsf{T}} = \mathbf{0},$$

load moved from one hall to the other, each conductance following its load. Without the chiller meter as well, each hall's scaling is lost on its own. Without the supply sensor, the meters fix both loads and the hall rows are blind to $b_{g_i} = b_{T_s}/\Delta T_i$, the direction $(0.179,\ 0.182,\ 0,\ 0,\ 1)^{\mathsf{T}}$ per kelvin. Without the chiller meter alone, only hall 2's scaling is lost.

What each loss costs. Figure 19.4 draws the standard deviations of Table 19.6 as grouped bars. Panel (a) holds the two log-conductances and the supply temperature, whose prior spreads are marked by the red dashed segments; panel (b) holds the two loads, which have no prior. In each group the five bars are the five cases in the order of the table.

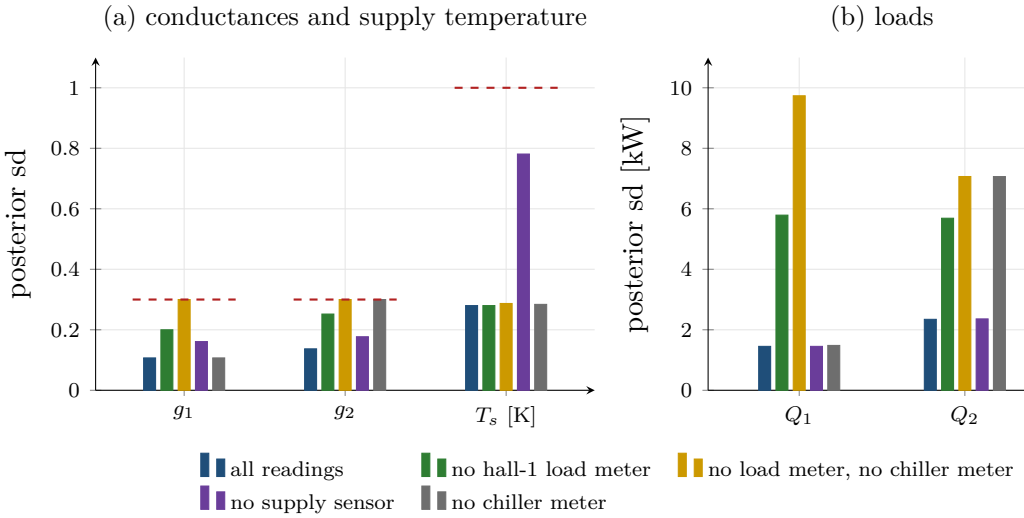


Figure 19.4: Posterior standard deviations of the five entries with meters removed, at the MAP state, Table 19.6. Red dashed segments: the prior standard deviations 0.3 of the log-conductances and 1.0 K of the supply temperature; the loads have a flat prior. A bar that reaches its red segment marks an entry the remaining readings do not see.

The hall-1 load meter. Without it the design loses the split of the load, and the loads widen to 5.79 and 5.69 kW; the conductances widen to 0.200 and 0.252, still inside their priors, because the chiller meter still sees each hall's scaling direction. The weakest direction of the remaining design, within 0.6° of the lost direction, has the singular value 0.125, a spread of 8.0 kW along it, held by the conductance priors. A refit without the meter moves the MAP loads to 26.19 and 26.40 kW and gives the spreads 5.91 and 5.91 kW there.

The load meter and the chiller meter. Without both, the design has rank three and loses each hall's scaling. Each moves a log-conductance with a load of flat prior, so by the factorization (19.24) both log-conductances sit at exactly their prior, 0.300. The loads, 9.74 and 7.07 kW, are then held by the conductance priors through the hall readings; this is the design of the outage, whose fifth question is computed in Section 19.3.4.

The supply sensor. Without it the lost direction is $\mathbf{v}_3$ of Table 19.5 to within 0.1°. The supply temperature widens to 0.781 K, short of its prior's 1.0, because the lost direction also moves the conductances, whose priors hold it in part; its ratio of prior to posterior spread is 1.28, not one, although no reading sees it. The conductances widen to 0.161 and 0.177, and the loads are unchanged.

The chiller meter. Without it g_2 returns exactly to its prior, 0.300, and Q_2 widens to 7.06 kW; hall 1, with its own meter, is untouched. A refit without it moves $\hat{Q}_2$ to 28.12 kW and gives the spread 8.91 kW there; a refit without both meters gives the spreads 8.83 and 8.89 kW at 27.82 and 28.04 kW, and one without the supply sensor the spread 0.777 K on T_s. A refit without it moves $\hat{Q}_2$ to 28.12 kW and gives the spread 8.91 kW there; a refit without both meters gives the spreads 8.83 and 8.89 kW at 27.82 and 28.04 kW, and one without the supply sensor the spread 0.777 K on T_s.

The weak directions announce the losses. The weakest direction $\mathbf{v}_5$ lies within 20.5° of the direction lost with the load meter and within 24.5° of the direction lost with the chiller meter; $\mathbf{v}_4$ and $\mathbf{v}_5$ together span, to within 0.55°, the plane lost when both meters go; and $\mathbf{v}_3$ is the direction lost with the supply sensor. On this plant each weak direction of the full design is held by a single meter, so the singular value decomposition of the present design maps which meter matters for what; a weak direction can equally be one that several readings hold together, each weakly.

Structural and practical identifiability on the plant. The full design is structurally identified, rank five. Its weakest direction is identified weakly in practice, $1/s_5 = 2.53$ kW on the split: enough for a cooling-margin question that needs each hall's load to ±5 kW at 95 %, not for billing the halls separately to ±1 kW.

19.3.4 Hall 2's load during the meter outage

Engineering question. With both meters out, how likely is hall 2's load above 45 kW? During the outage month only the three temperature readings and the three priors remain, the record $\mathbf{y}'$, and the engineer needs the probability that the load exceeds the rating of hall 2's electrical feed.

General form. A prediction under a threshold, a tail of the posterior, equation (2.12); here the event is an exceedance, the complement of the requirement,

$$\Pr(f(\mathbf{z}) > f^* \mid \mathbf{y}') = \int \mathbf{1}\big[f(\mathbf{z}) > f^*\big]\, p(\mathbf{z} \mid \mathbf{y}')\, \mathrm{d}\mathbf{z}. \tag{19.27}$$

Specialized form. The question is equation (19.27) with the state $\mathbf{z}$ the five entries, $f(\mathbf{z}) = Q_2$, $f^* = 45$ kW, and $p(\mathbf{z} \mid \mathbf{y}')$ the posterior of step 3 on the outage record, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}', \boldsymbol{\Sigma}')$ of equation (19.14). Under it the load is normal about its refitted value with the spread of the refitted curvature, and the tail is the normal tail

$$\Pr(Q_2 > 45\,\mathrm{kW} \mid \mathbf{y}') \approx 1 - \Phi\left(\frac{45 - \hat{Q}_2'}{\sqrt{\Sigma'_{Q_2 Q_2}}}\right). \tag{19.28}$$

Two choices enter here. The conductance prior of the outage month is the handbook prior: the shutdown cleans the coils, so the conductances fitted from the present record do not carry into the month, while the present temperature readings are taken to stand for the month's. This

is what makes the handbook spread the hold on each load, and it fails if the cleaning leaves a conductance where it was, in which case the present posterior of the conductance carries forward; Section 19.4.3 computes that case and varies the handbook spread. The loads' flat prior is kept, as in the full design: it is the prior the operators hold, and it is what makes the conductance prior, a log-scale prior, the only thing that holds each load, the situation Section 19.4.2 shows to be the worst case for the Gaussian.

The value beside it comes from the exact posterior, which here has a one-dimensional form: without the two meters each load appears in one row only, its hall's temperature row, linearly and with a flat prior, so that it integrates out in closed form; this holds as long as nothing else touches the load, and a nameplate prior on the load or a meter on either hall changes the integral. Integrating Q_i out over its flat prior, with the rise $\Delta T_i = Q_i \mathrm{e}^{-g_i}$ as the variable of integration and $\mathrm{d}Q_i = \mathrm{e}^{g_i}\mathrm{d}(\Delta T_i)$,

$$\int \exp\left[-\frac{(T_s + Q_i\mathrm{e}^{-g_i} - y_{T_i})^2}{2\sigma_T^2}\right]\mathrm{d}Q_i = \mathrm{e}^{g_i}\sqrt{2\pi}\,\sigma_T,$$

the change-of-variable factor e^{g_i}, which does not depend on T_s. The marginal of g_i is its prior times e^{g_i}, and completing the square, $-(g_i-\mu_0)^2/(2\sigma_0^2) + g_i = -(g_i - \mu_0 - \sigma_0^2)^2/(2\sigma_0^2) + \text{const}$ with $\mu_0 = \ln 5$ and $\sigma_0 = 0.3$,

$$g_i \mid \mathbf{y}' \sim \mathcal{N}(\ln 5 + 0.3^2,\ 0.3^2), \qquad Q_i \mid g_i, T_s, \mathbf{y}' \sim \mathcal{N}(\mathrm{e}^{g_i}(y_{T_i} - T_s),\ \mathrm{e}^{2g_i}\sigma_T^2), \tag{19.29}$$

with T_s normal from its sensor and its prior alone, since the hall rows integrate to constants in T_s: precisions add, $1/0.3^2 + 1/1^2$, and the mean is the precision-weighted mean of 12.10 and 12. The load is a product of a lognormal conductance and a normal rise. Given g_i, the supply temperature and the reading error combine into one normal, $Q_i = \mathrm{e}^{g_i}(y_{T_i} - T_s + \varepsilon_i)$ with $y_{T_i} - T_s + \varepsilon_i \sim \mathcal{N}(y_{T_i} - \mathbb{E}[T_s \mid \mathbf{y}'],\ \mathrm{Var}(T_s \mid \mathbf{y}') + \sigma_T^2)$, so the tail is the one-dimensional integral over the conductance

$$\Pr(Q_i > q \mid \mathbf{y}') = \int \left[1 - \Phi\left(\frac{q\,\mathrm{e}^{-g} - y_{T_i} + \mathbb{E}[T_s \mid \mathbf{y}']}{\sqrt{\mathrm{Var}(T_s \mid \mathbf{y}') + \sigma_T^2}}\right)\right]\mathcal{N}(g;\ \ln 5 + 0.3^2,\ 0.3^2)\,\mathrm{d}g, \tag{19.30}$$

and the quantiles of Q_i are the levels q at which it equals 0.975, 0.5 and 0.025.

Building blocks. The refitted MAP state $\hat{\mathbf{z}}'$ and the covariance $\mathbf{\Sigma}'$ of equation (19.14); and, for the exact values beside them, the posterior of T_s from its sensor and prior, the shifted normal of g_i, and the one-dimensional integral (19.30) with its quantiles by quadrature.

Hall 2's load above 45 kW during the outage. By equation (19.28), with $\hat{Q}_2' = 28.041\,\mathrm{kW}$ and $\sqrt{\Sigma'_{Q_2Q_2}} = 8.893\,\mathrm{kW}$ of equation (19.14),

$$\Pr(Q_2 > 45\,\mathrm{kW} \mid \mathbf{y}') \approx 1 - \Phi\left(\frac{45 - 28.041}{8.893}\right) = 1 - \Phi(1.907) = 0.028,$$

with the band $28.04 \pm 1.96 \times 8.893 = [10.6,\ 45.5]\,\mathrm{kW}$; for hall 1's load the band is $[10.5,\ 45.1]\,\mathrm{kW}$ and $\Pr(Q_1 < 10\,\mathrm{kW} \mid \mathbf{y}') = 0.022$.

The values beside them come from the exact posterior. With the rows of the two meters removed, the supply temperature has the posterior $T_s \mid \mathbf{y}' \sim \mathcal{N}(12.092,\ 0.287^2)$ from its sensor and its prior, and each log-conductance the shifted normal $g_i \mid \mathbf{y}' \sim \mathcal{N}(\ln 5 + 0.09,\ 0.3^2) = \mathcal{N}(1.6994,\ 0.3^2)$ of equation (19.29). The integral (19.30) gives for hall 2's load the median 30.5 kW, the 95 % band $[16.3,\ 56.8]\,\mathrm{kW}$ and

$$\Pr(Q_2 > 45\,\mathrm{kW} \mid \mathbf{y}') = 0.11,$$

and for hall 1's the median 30.3 kW, the band [16.2, 56.3] kW and $\Pr(Q_1 < 10\,\text{kW} \mid \mathbf{y}') = 0.0003$; two million draws of equation (19.29) reproduce these values within their sampling error. The gap on the tail is 0.082, a factor of four, and at the upper end of the band 11.3 kW. The state route errs here because the load is a product with a lognormal factor of spread 0.3, long in its upper tail, and the change-of-variable factor of the flat load prior moves the conductance 0.09 above the mode at which the state is centered (Section 19.4.2).

19.4 The checks (step 5)

The full calculation for all five questions is the exact posterior (19.6). The checks set each against the route reported, vary the priors, take the physics tolerance σ_p of the full state to its limit, and collect the conditions under which the inverse curvature is not the covariance.

19.4.1 The Laplace posterior against the exact posterior

The objective along the extreme directions. Figure 19.5 draws the objective along the best- and the worst-determined directions. The horizontal axis is the distance c from the MAP state along $\mathbf{v}_j$, counted in units of the Laplace spread $1/s_j$, so that the state is $\hat{\mathbf{z}} + (c/s_j)\mathbf{v}_j$; the vertical axis is the rise of the objective, $\mathcal{L} - \mathcal{L}(\hat{\mathbf{z}})$. The dashed parabola $\frac{1}{2}c^2$ is what the Laplace approximation assumes along every direction, by equation (19.22). Along the weakest direction $\mathbf{v}_5$ the objective follows the parabola to within 4 % at one spread (0.508 and 0.520 against 0.5) and 11 % at three. Along $\mathbf{v}_1$, the conductances, it is asymmetric: 0.529 and 0.474 at one spread, 5.34 and 3.84 at three. The asymmetry is the exponential e^{-g_i} of the hall rows: a smaller conductance raises the predicted temperature faster than a larger one lowers it, so the posterior is longer on the side of large conductances.

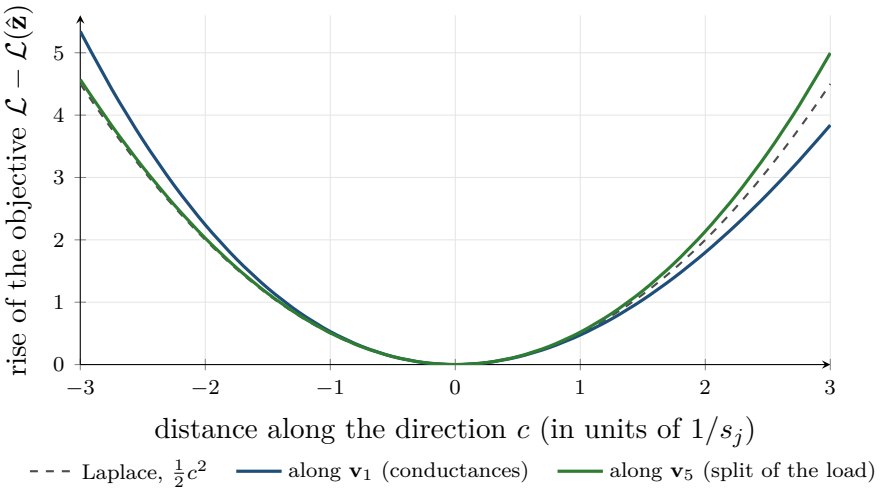


Figure 19.5: The objective of the two halls along the best-determined direction $\mathbf{v}_1$ (spread $1/s_1 = 0.0838$) and the worst-determined $\mathbf{v}_5$ (spread $1/s_5 = 2.530$) of Table 19.5, evaluated exactly at $\hat{\mathbf{z}} + (c/s_j)\mathbf{v}_j$ for c from -3 to 3, against the parabola $\frac{1}{2}c^2$ of the Laplace approximation. Positive c along $\mathbf{v}_1$ raises both conductances; along $\mathbf{v}_5$ it moves load from hall 2 to hall 1.

The exact Hessian. Chapter 18 evaluated the dropped term $\mathbf{S}$ of equation (18.15) at the MAP state, equation (18.29), and found in Section 18.4.2 that inverting $\mathbf{\Lambda} + \mathbf{S}$ instead of $\mathbf{\Lambda}$ changes no standard deviation by more than 1.2 %, in keeping with the per-row bound (18.17): the Gauss–Newton curvature is adequate here. On the full state of Chapter 17 the hall readings are linear and the same term comes from the two physics rows instead, $\sum_k r_k\, \partial^2 r_k / \partial \mathbf{z}\, \partial \mathbf{z}^{\mathsf{T}}$ over

its two physics rows: each normalized physics residual shrinks as σ_p while its second derivative grows as $1/\sigma_p$, so the product does not vanish in the limit. It is the curvature of the balance: on the eight-entry state of Section 19.4.4 its entries in the (g_i, Q_i) blocks are 1.039, -0.0339, -2.328 and 0.1047 at every tolerance from 0.01 to 0.001 K, the entries of equation (18.29), and the 1.2 % is the same number on either state.

The exact posterior. The weighted draws of Section 19.2.1, 4×10^5 states from the multivariate t with six degrees of freedom, have an effective number of 3.3×10^5. The exact standard deviations are (0.108, 0.139, 1.449 kW, 2.315 kW, 0.281 K), within 1.5 % of the Laplace values; the exact means differ from the MAP state by at most 0.13 standard deviations (g_1 higher by 0.014, T_s by 0.024 K); and the exact 95 % band of g_1 is [1.513, 1.938] against the Laplace [1.494, 1.912], the upper 2.5 % tail of the Laplace band holding 4.0 % of the exact posterior, the long side of large conductances seen along $\mathbf{v}_1$. The loads, whose rows are linear, have exact bands within about 0.13 kW of the Laplace bands at either end. For this plant, with all five readings, the Laplace covariance agrees with the exact posterior covariance to within 1.5 % in every spread.

19.4.2 The Laplace tail during the outage

The two halls without the load meter and the chiller meter show the first of the conditions of Section 19.4.5 on the state itself. The design has rank three, the loads have flat priors, and each hall's load is held only through its conductance prior. The MAP state refitted on $\mathbf{y}'$ has $g_i = \ln 5$ exactly, the prior mean, and the loads 27.81 and 28.04 kW; the curvature there gives the spreads 8.83 and 8.89 kW, against 9.74 and 7.07 kW at the MAP state of the full record (Table 19.6), and the Laplace bands $Q_2 \in [10.6,\ 45.5]$ and $Q_1 \in [10.5,\ 45.1]$ kW. The exact bands of Section 19.3.4 are [16.3, 56.8] and [16.2, 56.3] kW.

Both the center and the shape of the Laplace normal differ from the exact posterior. The center is shifted: the flat prior on the load is not flat in g_i, and the change-of-variable factor e^{g_i} of equation (19.29) moves the marginal of the log-conductance by $\sigma_0^2 = 0.09$ above the joint mode, the mechanism of Chapter 7's factor $1/\beta$. And the shape is wrong: a product with a lognormal factor of spread 0.3 has a long upper tail that the symmetric normal does not have. The probability that hall 2's load exceeds 45 kW is 0.11 exactly and 0.028 under the Laplace approximation, four times smaller; the probability that hall 1's is below 10 kW is 0.0003 exactly and 0.022 under it. An unobservable direction held by a log-scale prior is where the Laplace approximation is least to be trusted, and the rank test of Section 19.3.2 finds it before any fit is made.

19.4.3 Sensitivity to the priors

The priors of the record are the handbook conductance and the set point. Removing them altogether leaves the readings-only covariance $(\mathbf{J}_\mathrm{d}^\mathsf{T}\mathbf{J}_\mathrm{d})^{-1}$, the readings-only column of Table 19.4: the conductances widen from 0.107 and 0.137 to 0.115 and 0.155, the supply temperature from 0.280 to 0.300 K, and the loads from 1.45 and 2.35 to 1.50 and 2.50 kW. The priors sharpen the conductances by 7 and 12 % and the supply temperature by 7 %; with all five readings the answers do not rest on them.

During the outage they hold everything, and Table 19.7 varies them. Each row is a conductance prior for the outage month; the columns give its standard deviation in the log, the exact probability that hall 2's load exceeds 45 kW by the integral (19.30), the median and the 95 % band of the load. The first five rows keep the handbook mean $\ln 5$ and vary its spread; the last carries the present posterior of g_2, $\mathcal{N}(1.3999,\ 0.137^2)$ from Table 19.4, into the month in its place, the case of a cleaning that leaves the conductance where it was.

Table 19.7: The tail of hall 2's load during the outage, $\Pr(Q_2 > 45\,\mathrm{kW} \mid \mathbf{y}')$, its median and its central 95 % band, for five spreads of the handbook conductance prior about $\ln 5$ and for the present posterior of g_2 carried forward: the Laplace normal refitted under each prior, with the exact value in parentheses.

Conductance prior of the month	sd of g_2 [–]	$\Pr(Q_2 > 45\,\mathrm{kW})$	Median [kW]	95 % band [kW]
handbook	0.20	0.0036 (0.025)	28.0 (29.1)	[15.7, 40.4] ([18.6, 45.0])
handbook	0.25	0.013 (0.062)	28.0 (29.7)	[13.2, 42.9] ([17.4, 50.4])
handbook, the record's	0.30	0.028 (0.110)	28.0 (30.5)	[10.6, 45.5] ([16.3, 56.8])
handbook	0.35	0.049 (0.165)	28.0 (31.5)	[8.0, 48.1] ([15.4, 64.4])
handbook	0.40	0.072 (0.221)	28.0 (32.7)	[5.3, 50.7] ([14.5, 73.5])
present posterior carried forward	0.137	6×10^{-9} (3×10^{-5})	22.7 (23.1)	[15.1, 30.4] ([16.4, 32.1])

The tail grows twentyfold as the handbook spread goes from 0.20 to 0.40 (ninefold exactly), and the upper end of the band moves from 40 to 51 kW (45 to 74 kW exactly): the handbook spread is as much a part of the outage answer as the readings are. Carried forward, the present posterior would put the tail at 6×10^{-9} (3×10^{-5} exactly) and the band at 15 to 30 kW (16 to 32 kW exactly); the cleaning is what removes that knowledge for the month (Section 19.6.5).

19.4.4 The limit of exact balances and the conditioning

The reduced state is the limit $\sigma_p \to 0$ of the full state of Chapter 17, and the limit is checked here on the covariance at the MAP state. That state, equation (17.6), carries the hall temperatures and the chiller's heat as entries beside the five of this chapter, eight in all, with the eleven rows of equation (17.10): the five readings, the three priors and the three balances as physics rows, the hall balances with the tolerance σ_p in kelvin and the header balance with $\sigma_{p,P}$ in kilowatts, both given the same number here as in Table 17.7. The chiller meter's $\sigma_P = 2\,\mathrm{kW}$, which the table below sets beside the tolerance, is a reading's standard deviation and not a tolerance.

The singular value decomposition of the 11×8 Jacobian of the full state splits in two. Its three largest singular values belong to the directions that unbalance a hall or the header, the directions in which the covariance is nearly null, and grow as $1/\sigma_p$, from 58.3, 56.2 and 17.3 at $\sigma_p = 0.1$ to 5.8×10^3, 5.6×10^3 and 1.7×10^3 at 0.001: a tight tolerance makes any imbalance expensive and leaves those directions a spread of order σ_p. The other five, 3.471, 2.056, 2.052, 0.606 and 0.341, do not move with σ_p below about 0.1; their directions lie in the null space of the physics rows, and the rank, the weak directions and the observability statements of this chapter are statements about them. Measured in the coordinates of the five entries through the basis $\mathbf{N}$, the same five-dimensional subspace has the directions of $\mathbf{J} = \mathbf{J}_\mathrm{e}\mathbf{N}$, Table 19.5, with the same rank. Only the third, 3.471, has the same singular value on both states, because it hardly moves the carried entries; the conductance directions read 2.05 rather than 11.9 and 11.5 and the load directions 0.606 and 0.341 rather than 0.904 and 0.395, since a unit length on the full state also counts the kelvins of hall temperature that a change of conductance moves and the kilowatts of chiller heat that a change of load moves, the units choice of Section 19.3.1. Table 19.8 gives, for each tolerance, the ratio of the chiller meter's standard deviation to $\sigma_{p,P}$, the condition number of the eight-entry curvature, the largest relative change of the five spreads from Table 19.4 (the effect of the discrepancy), and the relative round-off error of the spreads obtained by inverting $\mathbf{\Lambda}$ rather than from the singular value decomposition of $\mathbf{J}$.

As the tolerance falls the marginal covariance of the five entries tends to Table 19.3: the change of the spreads falls as σ_p^2, from 60 % at 1 K, twice the hall sensors' error, where the balance is allowed to fail and widens the log-conductance of hall 1 most, through 1 % at 0.1 K

Table 19.8: The eight-entry full state of the two halls, equation (17.6), with the hall balances given the tolerance σ_p in kelvin and the header balance $\sigma_{p,P}$ in kilowatts, of the same number: condition number of the curvature, largest relative change of the five spreads from their values in the limit (Table 19.4), and the relative round-off error of the spreads from the inverse of $\mathbf{\Lambda}$ in double precision. σ_P is the chiller meter's standard deviation.

σ_p [K] $= \sigma_{p,P}$ [kW]	$\sigma_P/\sigma_{p,P}$ [–]	cond($\mathbf{\Lambda}$) [–]	Change of spreads [–]	Round-off of $\mathbf{\Lambda}^{-1}$ [–]
1	2	4.3×10^{2}	6.0×10^{-1}	1×10^{-15}
0.1	20	2.9×10^{4}	1.0×10^{-2}	2×10^{-14}
0.01	200	2.9×10^{6}	1.0×10^{-4}	1×10^{-12}
10^{-3}	2000	2.9×10^{8}	1.0×10^{-6}	8×10^{-11}
10^{-6}	2×10^{6}	2.9×10^{14}	below 10^{-10}	9×10^{-5}

to 10^{-4} at Chapter 17's 0.01 K, which is the check that the five-entry marginal of the full state reaches the covariance of this chapter. The condition number grows as $1/\sigma_p^2$: the physics rows' entries in $\mathbf{J}$ scale as $1/\sigma_p$, so their contribution to $\mathbf{\Lambda}$ scales as $1/\sigma_p^2$ while the readings' does not, which is the growth of the three largest singular values squared. The inverse of the curvature in finite precision loses accuracy in proportion: its relative error can reach cond($\mathbf{\Lambda}$) times the machine precision, 1.1×10^{-16} in double precision, a standard bound of numerical linear algebra. At 10^{-6} the inversion keeps only four digits, its error now six orders of magnitude larger than the modeling effect it was meant to remove, and a Gauss–Newton step solved from the same matrix can carry an error of the same relative size. At 0.01, about two decades below the meters and sensors, the discrepancy changes the spreads by 10^{-4} and the inversion keeps twelve digits. Rows are therefore given scales within about two decades of one another, and the covariance is computed from a factorization of $\mathbf{J}$, by the singular value decomposition or an orthogonal factorization, whose condition number is cond($\mathbf{J}$) rather than cond($\mathbf{J}$)^2, whenever that is large; the projection onto the null space of the physics rows, which this chapter works in, removes the three large singular values altogether. Expressing an entry in other units has the same effect on the columns that σ_p has on the rows, the units choice seen from the computation.

19.4.5 The conditions under which the inverse curvature fails

The derivation of Section 19.2.1 made two assumptions: that the remainder $\mathcal{E}_3$ is negligible over the posterior's own width, and that the curvature used is the Hessian of the objective. Four conditions break one or the other, and Volume I met each, with the exact route beside the approximate one. A fifth, mixed scales among the rows, breaks neither: the posterior is as Gaussian as before, but the computed inverse of the curvature loses accuracy, as Section 19.4.4 showed on the full state. Table 19.9 collects the five as a rule: each row is a condition, the second column the diagnostic that flags it, the third the case where it was shown with the Laplace and the exact numbers, the last the remedy used there.

Figure 19.6 draws four of the cases side by side with the data of those chapters' figures. In each panel the horizontal axis is the quantity of the case and the vertical axis its posterior density; the solid green curve is the exact density and the dashed blue curve the one the Laplace approximation, or the delta method built on it, implies.

Strong nonlinearity within the band. The crossing time of Chapter 12 divides by the trend, and the posterior of the trend has positive density at zero: within the band the quantity has a pole. Panel (c) shows the result. The Gaussian on the two entries of that chapter's state was adequate, its grid check moving the upper end of a band that reaches 28 years by 1.3 years, but the derived quantity bent across the band so sharply that no tangent could carry it: the

Table 19.9: The conditions under which $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}^{-1}$ (or $\mathbf{H}^{-1}$) is not to be trusted, the diagnostic that flags each, the chapter where each was shown with the Laplace and the exact numbers there, and the remedy. The thresholds are working values, not sharp limits.

Condition	Diagnostic	Evidence (Laplace against exact)	Remedy
Strong nonlinearity within the band	the objective along a weak direction, or a derived quantity's draws, departs from $\frac{1}{2}c^2$ by more than about 10 % at $c = 2$	Chapter 12: crossing time of a rate trend, delta band [4.44, 11.45] against exact [5.90, 28.34] years; this plant without the hall-1 load meter and the chiller heat meter: Q_2 band [10.6, 45.5] against [16.3, 56.8] kW	invert an event linear in the state; draws
Entry near a bound	the mode within about three standard deviations of a bound, on the natural scale	Chapter 15: logit of the common-cause share with one double failure, standard deviation 0.748 against 0.840, skewness -0.68; band on the share [0.034, 0.399] against [0.018, 0.329]	grid of the exact posterior
Few rows behind an entry	gamma shape (counts plus prior pseudo-counts) below about 10	Chapter 9: log rate with a gamma posterior of shape four, $\Pr(R(72) \geq 0.90 \mid \mathbf{y}) = 0.012$ against 0.042; Chapter 7: five failures, $\Pr(\beta > 1 \mid \mathbf{y}) = 0.762$ against 0.522	the exact gamma or one-dimensional integral
Rows not Gaussian; the joint mode is not the marginal mode	lifetime rows, or count rows off their canonical link, with a curvature term; for Gaussian rows, the residual term $\mathbf{S}$ above a few percent of Λ_{ii}	Chapter 7: curvature term of the Weibull working rows; mode of the log-shape $\ln \beta$ at 0.279 against 0.109 for the exact marginal	exact Hessian; integrate the other entries out
Mixed scales; an ill-conditioned curvature	$\mathrm{cond}(\boldsymbol{\Lambda}) \times 10^{-16}$ above the relative precision the spreads need	this chapter, on the full state: physics rows of tolerance 10^{-3} (K, kW) beside sensors of 0.5 K and meters of 2 kW, $\mathrm{cond}(\boldsymbol{\Lambda}) = 2.9 \times 10^{8}$, at 10^{-6} 2.9×10^{14}	rows within two decades; factorize $\mathbf{J}$, not $\boldsymbol{\Lambda}$

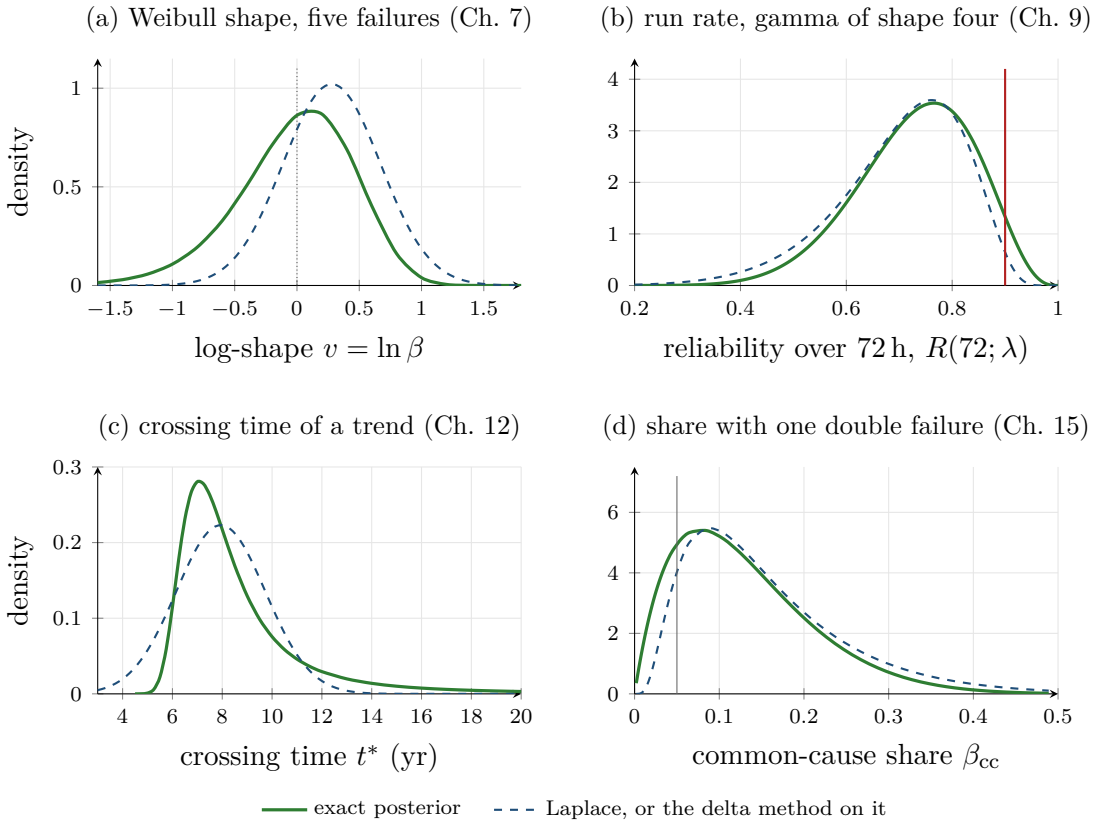


Figure 19.6: Four cases of Volume I in which the Gaussian at the mode failed, redrawn from the data of those chapters. (a) The log-shape of the Weibull motors under the flat prior: the exact marginal against the Laplace normal $\mathcal{N}(0.279,\ 0.391^2)$; dotted line $\beta = 1$. (b) The generator's 72 h run reliability: the transformed Gamma(4, 880 h) against the Laplace normal $\mathcal{N}(-5.394,\ 0.5^2)$ on $\ln \lambda$ mapped to the reliability; red line the requirement 0.90. (c) The crossing time of the limit by a rising failure rate: the closed-form density against the delta-method normal $\mathcal{N}(7.94,\ 1.79^2)$. (d) The common-cause share of the pump pair: the exact marginal against the Laplace normal $\mathcal{N}(-1.874,\ 0.748^2)$ on its logit mapped to the share; thin line 0.05.

delta-method normal is symmetric, starts years too early and stops decades too soon. The same condition on the state itself appears when a row is strongly nonlinear in an entry over the entry's own spread, which is what happens on this plant when the hall-1 load meter and the chiller heat meter are both lost (Section 19.4.2).

An entry near a bound. The common-cause share of Chapter 15 lies in $(0, 1)$, and with one double failure its posterior sits near zero. Carried as a logit it is unbounded, but the logit of a Beta-like variable with a small first parameter is skewed, and the normal at the mode, panel (d), is too thin near zero and too heavy in the upper tail. The transformation to an unbounded entry is necessary; it is not sufficient when the evidence is one event.

Few rows behind an entry. The curvature of a log rate informed by a gamma posterior of shape a is a, and the log of a gamma variable is near normal only when a is large. With shape four in Chapter 9, panel (b), the Laplace normal understates the probability that the rate is small by a factor of three and a half at the requirement. With five failures in Chapter 7, panel (a), the conditional posterior of $\eta^{-\beta}$ given the shape is a gamma of shape five, equation (7.44), and the shape's marginal is skewed and wider than the curvature says.

Rows that are not Gaussian. Chapter 7 wrote the Weibull rows as count rows in working form, equation (7.33), with $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ the expected information, and added the curvature term of equation (7.34), the observed-minus-expected difference, to obtain the Hessian at the data that the Laplace approximation calls for. It then showed a second effect that no Hessian removes: integrating the scale out weights each shape by $1/\beta$, so the mode of the marginal of the log-shape, 0.109, is not the joint MAP, 0.279. The Laplace approximation is centered at the joint mode, and a marginal report from it inherits the offset. For Gaussian rows with large residuals the residual term $\mathbf{S}$ of equation (18.15) is the analogous correction; on this plant it changes no spread by more than 1.2 % (Section 19.4.1).

Mixed scales. The fifth condition concerns the computation rather than the posterior: Table 19.8 shows it on the full state, where a tolerance of 10^{-6} leaves the inverse of the curvature four digits.

Remedies. Each condition has a remedy that Volume I used. *Reparameterize*: carry a positive quantity by its log and a probability by its logit, as every chapter of Volume I did, and center a trend at the middle of its record, as Chapter 12 did, so that the posterior in the new entries is closer to normal. *Use the exact Hessian* when the rows are not Gaussian or the residuals are large. *Draw*: when the approximation is in doubt, draws of the posterior, from the Laplace normal pushed through a derived quantity or importance-weighted to the exact posterior as in Section 19.4.1, test it directly. *Integrate*: when one or two entries carry the difficulty, the exact posterior on a grid or by a one-dimensional integral, as in Chapters 7, 9 and 15 and in equation (19.30), is cheap and settles the question. And *factorize the Jacobian* rather than the curvature when the rows differ in scale.

19.5 The report (step 6)

With every meter in service the Gaussian at the mode and the exact posterior agree to within 1.5 % in every spread. With the hall-1 load meter and the chiller heat meter out, the posterior of the state gives hall 2's load a probability of 2.8 % of exceeding 45 kW, and the exact posterior 11 %, four times as much: the split of the load, the direction the readings determine least, is then held by nothing but the handbook conductances, and a log-scale prior gives the load a

long upper tail that the normal does not have. The answers below are given in the units in which the questions were asked.

19.5.1 The answers and their bands

Table 19.10 gives the answers, one row per question of Table 19.1 in its order: the answer from the posterior of the state in the second column, the exact posterior's value beside it in the third, and its band or qualification in the last.

Table 19.10: The answers to the questions asked of the two halls' curvature. Spreads are posterior standard deviations under the Laplace approximation at the MAP state, on $\mathbf{y}'$ for the fifth; bands are central 95 %.

Question	Answer (state route)	Exact	Band and qualification
How well are the conductances, the loads and the supply temperature known?	$G_1 = 5.49$, $G_2 = 4.055$ kW/K to 11 and 14 %; $Q_1 = 30.7 \pm 1.45$, $Q_2 = 22.2 \pm 2.35$ kW; $T_s = 12.11 \pm 0.28$°C	spreads (0.108, 0.139, 1.449 kW, 2.315 kW, 0.281 K), within 1.5 %	$G_1 \in [4.45,\ 6.77]$, $G_2 \in [3.10,\ 5.30]$ kW/K; correlations 0.69 (g_2, Q_2), −0.57 (Q_1, Q_2), 0.43 (g_1, T_s)
Which direction is least determined?	the split of the load, $\mathbf{v}_5 = (0.01,\ -0.04,\ 0.41,\ -0.91,\ 0.00)$	the objective along it within 4 % of the parabola at one spread	spread $1/s_5 = 2.53$ kW; $\mathrm{cond}(\mathbf{J}) = 30.2$ in the stated units
Can my sensors see this parameter at all?	yes, every one: design columns of rank five	– (a linearized test)	prior/posterior spread 2.81, 2.19, ∞, ∞, 3.57; no redundancy among the readings themselves
What is lost with a meter?	load meter: the split (rank four); with the chiller meter: each hall's scaling (rank three); supply sensor: T_s with the conductances (rank four)	no exact posterior computed; refits: loads 5.91, 5.91 kW; both meters out, 8.83, 8.89 kW; T_s 0.777 K	loads 5.8, 5.7 kW; conductances at their prior 0.30, loads 9.7, 7.1 kW; T_s 0.78 against 0.28 K
With both meters out, how likely is hall 2's load above 45 kW?	$\Pr(Q_2 > 45\,\mathrm{kW} \mid \mathbf{y}') =$ 0.028; band [10.6, 45.5] kW	0.11; band [16.3, 56.8] kW	gap 0.082, a factor of four, from the long upper tail of a load held by a log-scale prior; 0.0036 to 0.072 (0.025 to 0.221 exact) for a handbook spread of 0.20 to 0.40

19.5.2 The report line

The two halls' heat-removal conductances are 5.49 and 4.055 kW/K, known to 11 and 14 %, their loads 30.7 ± 1.45 and 22.2 ± 2.35 kW, and the supply temperature 12.11 ± 0.28°C. The five readings determine every one of these without the handbook values, which sharpen the conductances by 7 and 12 % only, and the Gaussian bands are confirmed by the exact posterior to within 1.5 %. The quantity the readings determine least is how the total load divides between the two halls, known to 2.5 kW; it is held by the hall-1 load meter, and if that meter fails the division is no longer seen by any reading and rests on the handbook conductances, with the loads then known only to about 6 kW. During the planned outage of that meter and the chiller heat meter, hall 2's load is known only through its handbook conductance: its 95 % band is 11 to 45 kW and the probability that it exceeds 45 kW is 0.028 from the posterior of the state, or

16 to 57 kW and 0.11 from the exact posterior: the two differ by a factor of four, because the load is then a product with a lognormal factor, and the gap is reported with the answer.

Common mistake

Reading identifiability from the marginal spreads. A small marginal spread says that an entry is well known, not that the readings know it: an entry with a tight prior keeps a small spread whether or not any reading sees it. The test is the rank of the design columns, the singular directions, and the ratio of prior to posterior spread; for the two halls without the chiller heat meter the conductance of hall 2 reports ±0.30 under the Laplace approximation, a spread that looks like an estimate and is the prior returned unchanged.

19.6 The next measurement (step 7)

The report leaves one decision open, the outage month, during which hall 2's load rests on the handbook conductance. The candidate measurements are single rows added to the design before the outage begins: a meter or a sensor whose Jacobian row is known now, at the MAP state, because the curvature does not depend on the value the row will read.

19.6.1 The preposterior of a candidate row

A candidate reading a is the linear Gaussian row of the design form of Chapter 2, $(\mathbf{h}_a\mathbf{z} - \tilde{y}_a)/\sigma_a$, with the Jacobian row $\mathbf{J}_a = \mathbf{h}_a/\sigma_a$ of equation (2.17) (a hall sensor's row linearized at $\hat{\mathbf{z}}$). Written as the column $\mathbf{j}_a = \mathbf{J}_a^\mathsf{T}$, it adds the outer product $\mathbf{j}_a\mathbf{j}_a^\mathsf{T} = \mathbf{\Lambda}_a$ to $\mathbf{\Lambda}$ by equation (19.11), and the covariance after it is the one-row case of the preposterior,

$$\mathbf{\Sigma}_a = \left(\mathbf{\Sigma}^{-1} + \mathbf{j}_a\mathbf{j}_a^\mathsf{T}\right)^{-1}, \qquad \mathbf{\Sigma}_{a,\text{out}} = \left(\mathbf{\Lambda} - \mathbf{j}_4\mathbf{j}_4^\mathsf{T} - \mathbf{j}_5\mathbf{j}_5^\mathsf{T} + \mathbf{j}_a\mathbf{j}_a^\mathsf{T}\right)^{-1}, \tag{19.31}$$

the second being the same row added to the outage design, the removal (19.26) of the two meters' rows with the candidate kept. The width a candidate removes from an answer is the current spread less the preposterior one, and its expected information gain is $\frac{1}{2}\ln(\det\mathbf{\Sigma}/\det\mathbf{\Sigma}_a)$, equation (2.18). The removal analysis of Section 19.3.3 was this calculation run backwards, and the singular directions say where to look: a row whose $\mathbf{j}_a$ has a large component along $\mathbf{v}_5$ adds curvature where it is scarcest. Four candidates are compared, each with the standard deviation of a meter or sensor of the plant: a load meter on hall 2, $\mathbf{j}_a = (0,\ 0,\ 0,\ 1/1.5,\ 0)$; a second hall-1 load meter, $(0,\ 0,\ 1/1.5,\ 0,\ 0)$; a second hall-2 temperature sensor, the row r_2 of Table 19.2 again; and a second supply sensor, $(0,\ 0,\ 0,\ 0,\ 3.333)$.

19.6.2 The candidate rows

Table 19.11 ranks the four. Each row is a candidate, with the present design first; the columns give the spread of hall 2's load after it, the width it removes from the present 2.346 kW, the spread along the weakest direction $\mathbf{v}_5$, the spread of hall 2's load if the hall-1 load meter then fails, the same during the outage of the hall-1 load meter and the chiller heat meter, from equation (19.31), the expected information gain, and the probability that hall 2's load exceeds 45 kW during the outage with the candidate in service. The spreads and the gain are Laplace values at the MAP state (19.12); the probability is the normal tail (19.28) of the state refitted on the outage record with the candidate's row added, the candidate's reading set at the value its row predicts at $\hat{\mathbf{z}}$. The value in parentheses is exact, from the outage posterior with the same row, the loads integrated out in closed form as in equation (19.29) and the conductances and the supply temperature on a grid.

The hall-2 load meter is the only candidate that acts on the weak direction: it removes 1.08 kW from the spread of hall 2's load and 1.1 kW from the spread along the split, it keeps

Table 19.11: Candidate rows for the two halls: spread of Q_2 after the row, the width removed from the present 2.346 kW, the spread along $\mathbf{v}_5$, the spread of Q_2 with the hall-1 load meter failed and during the outage of both meters with the candidate in service, the expected information gain (Laplace preposterior at the MAP state), and the probability that Q_2 exceeds 45 kW during the outage, from the state refitted with the candidate's reading at its value predicted at $\hat{\mathbf{z}}$ (exact in parentheses).

Candidate	sd of Q_2 [kW]	Width removed [kW]	sd along $\mathbf{v}_5$ [kW]	No load meter [kW]	Outage [kW]	EIG [nat]	$\Pr(Q_2 > 45\,\text{kW})$, outage
present design	2.346	–	2.530	5.688	7.067	0	0.028 (0.110)
hall-2 load meter, 1.5 kW	1.264	1.082	1.425	1.450	1.467	0.619	$< 10^{-50}$ (both)
second hall-1 load meter, 1.5 kW	2.153	0.194	2.189	2.346	7.065	0.331	0.029 (0.111)
second hall-2 sensor, 0.5 K	2.341	0.006	2.524	5.607	6.920	0.328	0.023 (0.102)
second supply sensor, 0.3 K	2.345	0.001	2.529	5.685	7.021	0.313	0.027 (0.108)

hall 2's load seen when either of the other meters is lost, 1.45 kW without the hall-1 load meter and 1.47 kW during the outage, and it takes the outage tail from 0.028 (0.110 exact) to nothing measurable by either route. A second hall-1 load meter guards against the failure of the first, leaving the present 2.346 kW when one of the two fails, but it does nothing for the outage, which takes the chiller meter as well: the tail stays at 0.029 (0.111). The second hall-2 sensor gains as much information in total as a second hall-1 meter, 0.33 nat, and the second supply sensor 0.31 nat, but both spend it on the directions already well determined, the conductances and the supply temperature, and leave the outage tail near 0.025 (near 0.10 exactly). The ranking by width on hall 2's load and the ranking by information gain agree on the first place, and only the first ranking and the outage tail separate the candidates for the decision at hand.

19.6.3 The recommendation

The next measurement is a load meter on hall 2, of the class of the hall-1 meter, installed before the outage begins. It narrows hall 2's load from 2.35 to 1.26 kW and the split of the load from 2.53 to 1.43 kW now, and during the outage it keeps hall 2's load to 1.47 kW, read directly by its own row, in place of a band held by the handbook conductance alone, and the probability of exceeding the feed's 45 kW falls from 0.028 (0.11 exact) to nothing measurable; the gap between the two routes does not move the recommendation, which ranks first by both. Hall 1's load stays unseen during the outage with or without it, at 9.73 kW under the Laplace normal at the MAP state, and the same long upper tail that this chapter found for hall 2 would then apply to hall 1; if hall 1's load matters during the month, the replacement of its meter is better not scheduled together with the chiller meter's: with the hall-2 meter in place and only the hall-1 load meter out, hall 1's load is known to 2.40 kW.

19.6.4 What the analysis showed

The first reading of the outage, a fit that still converges on three temperature readings and a Gaussian band that gives hall 2's load a 2.8 % chance of exceeding 45 kW, treated the fit's spreads as knowledge from the readings; the analysis shows that during the outage they are the handbook's, and it sets beside the 2.8 % the 11 % of the exact posterior, whose gap is the long upper tail of a load held by a log-scale prior. With every meter in service the plant is well instrumented: the five readings alone determine all five quantities, the handbook values

add 7 to 12 %, and the Gaussian bands hold to 1.5 %. Its one soft spot is how the total load divides between the halls, known to 2.5 kW and held by a single meter, the hall-1 load meter, as the weakest singular direction showed before any meter was removed. Take that meter away and the division is no longer seen by any reading; take the chiller meter with it and each hall's load rests on a handbook conductance, where the Gaussian at the mode departs from the exact posterior in both center and tail. What the operators carry back is that the meter to protect is the one whose loss turns the weakest direction unobservable, and that a load meter on hall 2, installed before the outage, keeps hall 2's load within 1.5 kW through the month.

19.6.5 Limits of the model

One operating point. Every curvature of this chapter is taken at one MAP state of one steady operating condition, and the design statements depend on it through the Jacobian: a hall reading sees its log-conductance with the weight $\Delta T_i/\sigma_T$ of equation (19.4), so at a light load, with a small rise, the same sensors resolve the conductances less well. The rank holds at every operating point with positive rises; the singular values and the spreads do not.

The linearized observability test. The rank and the null directions are those of the rows linearized at the MAP state. Along an unobservable direction the nonlinear rows still carry information, the center shift of Section 19.4.2, and the factorization that returns an entry to its prior is then approximate.

Exact balances. The reduced state takes the balances as exact. A balance that fails by a physical amount, heat gained by the pipes or a load that does not reach the coil, is a discrepancy that adds its variance to every quantity the balance determines; a tolerance of 0.1 K in Table 19.8 already inflates the spreads by 1 %, and a real discrepancy belongs in the model as a tolerance of physical size, not in the limit.

The handbook prior. With every meter in service the priors matter little; during the outage they are the answer. The 35 % spread of the handbook conductance, the assumption that the two halls' conductances are independent of each other, and the assumption that the cleaning returns each coil to the handbook population set the exact band of each load (Table 19.7), and a cleaned coil of hall 2 that departed from the handbook would move that band without any reading to show it.

Independent meter errors. The removal and the candidate rows assume that each meter's error is its own. Two meters sharing a transducer, a calibration or a data path fail together, and their joint loss is the two-row removal of Table 19.6, not two single ones.

Common mistake

Comparing singular values or condition numbers across choices of units. Expressing the loads in watts multiplies the condition number of the two halls by a thousand without changing the posterior. Singular values are compared within one choice of units, stated with them; the rank and the prior-to-posterior ratios do not depend on the choice.

Chapter 20 carries $\boldsymbol{\Sigma}$ to engineering outputs, such as a hall's temperature under a future load, and ranks candidate rows by the width each removes from those outputs.

19.7 What to retain

Definition 19.1 (Laplace covariance and observability of a state)**.** Near its MAP state the posterior of a state of n entries is approximated by $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma} = \mathbf{H}^{-1}$, the inverse Hessian of the objective, and for Gaussian rows with small residuals $\boldsymbol{\Sigma} \approx \boldsymbol{\Lambda}^{-1} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$ on normalized rows ($\mathbf{J}^{\mathsf{T}}\mathbf{W}\mathbf{J}$ on unnormalized ones). The marginal variance of entry i is $\Sigma_{ii} = 1/(\Lambda_{ii} - \boldsymbol{\lambda}_{ri}^{\mathsf{T}}\boldsymbol{\Lambda}_{rr}^{-1}\boldsymbol{\lambda}_{ri})$, at least the conditional variance $1/\Lambda_{ii}$, and $\varrho_i^2 = 1 - 1/(\Lambda_{ii}\Sigma_{ii})$ is its squared multiple correlation with the other entries. With $\mathbf{J} = \mathbf{F}\,\mathrm{diag}(s_j)\mathbf{V}^{\mathsf{T}}$ the spread along the right singular vector $\mathbf{v}_j$ is $1/s_j$. The readings observe the state when the design columns $\mathbf{J}_\mathrm{d}$, the Jacobian of the reading rows without the priors, have rank n; an entry with a proper prior whose posterior spread equals its prior spread is unobserved. Removing a set of rows subtracts their outer products from $\boldsymbol{\Lambda}$, and adding one adds its own, so what a meter costs or buys is known before it fails or is installed.

Five conclusions carry beyond the two halls. The curvature that the fit already forms is the posterior precision, and every spread, correlation and direction of this chapter is linear algebra on it, with no further fit. The balances act through the projection onto the null space of the physics rows, $\mathbf{J} = \mathbf{J}_\mathrm{e}\mathbf{N}$: the full state with its tolerance σ_p reaches the same five-entry covariance as $\sigma_p \to 0$, and the tolerance adds only two large singular values, the directions that unbalance a hall, and a cost in conditioning. The spread of an entry with the others unknown exceeds its spread with them fixed by the factor $1/\sqrt{1-\varrho_i^2}$. Observability is a property of the design and is known before any value is read: the rank of the design columns and the ratio of prior to posterior spread answer whether the sensors can see an entry at all, and removing a row from the Jacobian answers what a failed meter costs. And the Gaussian band is an approximation with known failure conditions, nonlinearity across the band, bounds, few rows and non-Gaussian rows, each with an exact route beside it, while poor conditioning spoils the computation rather than the approximation; an unobservable direction held by a log-scale prior is where the band fails worst, and the rank test finds it first.

The computation as an algorithm. The calculations of this chapter need the rows, their kinds and the MAP state, and they use only linear algebra on the Jacobian.

1. *Inputs.* The normalized rows $r_k(\mathbf{z})$ with their kind (physical, reading or prior), the MAP state $\hat{\mathbf{z}}$ from the fit, and the prior standard deviation $\sigma_{0,i}$ of each entry (infinite for a flat prior).
2. *Jacobian.* Evaluate $\mathbf{J}$ at $\hat{\mathbf{z}}$, analytically as in equation (19.4) or by central differences, as $\mathbf{J}_\mathrm{e}\mathbf{N}$ when the plant's balances are physics rows with a null-space basis $\mathbf{N}$, and split it into the design columns $\mathbf{J}_\mathrm{d}$ and the prior rows.
3. *Curvature and covariance.* Form $\boldsymbol{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$, check that it is positive definite by its Cholesky factorization, and obtain $\boldsymbol{\Sigma}$ by triangular solves with the factor; for non-Gaussian rows use the exact Hessian $\mathbf{H}$ instead.
4. *Spreads.* Report $\sqrt{\Sigma_{ii}}$, the conditional spreads $\Lambda_{ii}^{-1/2}$, ϱ_i^2 by equation (19.19), and the correlations.
5. *Directions.* Compute the singular value decomposition of $\mathbf{J}$; report the s_j, the spreads $1/s_j$, the right singular vectors and $\mathrm{cond}(\mathbf{J}) = s_1/s_n$, with the units of the entries stated.
6. *Observability.* Compute the singular values of $\mathbf{J}_\mathrm{d}$; count as zero those below a tolerance relative to the largest, $10^{-8}s_1$ in double precision; report the rank, the null vectors and the ratios $\sigma_{0,i}/\sqrt{\Sigma_{ii}}$, flagging every ratio near one.
7. *Removed and added rows.* For each reading row, or set of rows, delete it and repeat steps 3 to 6 at the same $\hat{\mathbf{z}}$; report the rank, the spreads, the entries at their prior and the lost directions, and the angle between each lost direction and the weak directions of the full design; add each candidate row by equation (19.31) and report the width it removes.

8. *Checks.* Add the residual term $\mathbf{S}$ of equation (18.15) and compare the spreads; evaluate the objective along the weakest and strongest directions against $\frac{1}{2}c^2$; and, when either check or a condition of Table 19.9 warns, compute the exact posterior by draws or by the one-dimensional integrals that exist.

Table 19.12 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 19.12: The terms of the Hessian, the covariance and identifiability, and where each is defined.

Term	Expression	Units	Meaning	Defined in
Reduced Jacobian	$\mathbf{J} = \mathbf{J}_{\mathrm{e}}\mathbf{N}$, $\mathbf{N}$ the null-space basis of the physics rows	varies	rows along the balances	19.1.3
Design columns	$\mathbf{J}_{\mathrm{d}}$, reading rows only	varies	what the readings can say	19.1.6
Conditional variance	$1/\Lambda_{ii}$	varies	others held fixed	19.3.1
Marginal variance	$\Sigma_{ii} = 1/(\Lambda_{ii} - \boldsymbol{\lambda}_{ri}^{\mathsf{T}}\boldsymbol{\Lambda}_{rr}^{-1}\boldsymbol{\lambda}_{ri})$	varies	Schur complement	19.3.1
Squared multiple correlation	$\varrho_i^2 = 1 - 1/(\Lambda_{ii}\Sigma_{ii})$	–	variance shared with the rest	19.3.1
Correlation	$\Sigma_{ij}/\sqrt{\Sigma_{ii}\Sigma_{jj}}$	–	combinations the readings fix	19.3.1
Singular value decomposition	$\mathbf{J} = \mathbf{F}\,\mathrm{diag}(s_j)\mathbf{V}^{\mathsf{T}}$, $\boldsymbol{\Sigma} = \mathbf{V}\,\mathrm{diag}(s_j^{-2})\mathbf{V}^{\mathsf{T}}$	varies	directions of the curvature	19.3.1
Spread along a direction	$1/s_j$ along $\mathbf{v}_j$	varies	weakest is $\mathbf{v}_n$	19.3.1
Condition number	$\mathrm{cond}(\mathbf{J}) = s_1/s_n$; $\mathrm{cond}(\boldsymbol{\Lambda}) = \mathrm{cond}(\mathbf{J})^2$	–	depends on units	19.3.1
Unobservable direction	$\mathbf{J}_{\mathrm{d}}\mathbf{b} = \mathbf{0}$	–	held by priors only	19.3.2
Prior-to-posterior ratio	$\sigma_{0,i}/\sqrt{\Sigma_{ii}} \geq 1$	–	one means unobserved	19.3.2
Removed rows	$\boldsymbol{\Lambda}_{-a} = \boldsymbol{\Lambda} - \sum_{k\in a}\mathbf{j}_k\mathbf{j}_k^{\mathsf{T}}$	varies	what a meter held	19.3.3
Exact tail without the meters	one-dimensional integral over g_i, equation (19.30)	–	load held by a log-scale prior	19.3.4
Laplace approximation	$\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}},\ \mathbf{H}^{-1})$	–	local Gaussian posterior	19.2.1
Normalizing integral	$(2\pi)^{n/2}(\det \mathbf{H})^{-1/2}$	–	exists iff $\mathbf{H}$ positive definite	19.2.1
Exact Hessian	$\mathbf{J}^{\mathsf{T}}\mathbf{J} + \sum_k r_k\nabla^2 r_k$	varies	curvature with residual term	19.2.1
Information matrix	$\boldsymbol{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J} = \sum_k \mathbf{j}_k\mathbf{j}_k^{\mathsf{T}}$	varies	Gauss–Newton curvature; precisions add	19.2.1
Conditions of failure	Table 19.9	–	when $\boldsymbol{\Sigma} \neq \boldsymbol{\Lambda}^{-1}$	19.4.5
Conditioning	error up to $\mathrm{cond}(\boldsymbol{\Lambda}) \times 1.1 \times 10^{-16}$	–	rows within two decades	19.4.4
Preposterior of a row	$\boldsymbol{\Sigma}_a = (\boldsymbol{\Sigma}^{-1} + \mathbf{j}_a\mathbf{j}_a^{\mathsf{T}})^{-1}$	varies	what a new row buys	19.6.1

19.8 Exercises

Exercise 19.1. *Formulation.* Two entries z_1 and z_2, in kelvin, have flat priors. One reading measures $z_1 + z_2$ with $\sigma = 0.5\,\mathrm{K}$ and a second measures z_2 with $\sigma = 1\,\mathrm{K}$. (a) Write the normalized

Jacobian, the information matrix and the covariance. (b) Compute the marginal and conditional standard deviations of each entry and their correlation. (c) A different design has the Jacobian rows $(1,\ 2)$, $(-2,\ -4)$ and $(0.5,\ 1)$. State its rank, the unidentified direction and a row that would restore full rank. ((a) $\mathbf{J} = \begin{bmatrix} 2 & 2 \\ 0 & 1 \end{bmatrix}$ in K^{-1}, $\mathbf{\Lambda} = \begin{bmatrix} 4 & 4 \\ 4 & 5 \end{bmatrix}$, $\det \mathbf{\Lambda} = 4$, $\mathbf{\Sigma} = \begin{bmatrix} 1.25 & -1 \\ -1 & 1 \end{bmatrix}$ K^2. (b) Marginal 1.118 and 1.000 K, conditional $1/\sqrt{4} = 0.500$ and $1/\sqrt{5} = 0.447$ K, correlation $-1/\sqrt{1.25} = -0.894$: the first reading fixes the sum, so the entries trade. (c) Every row is a multiple of $(1,\ 2)$: rank one; the direction $(2,\ -1)/\sqrt{5}$ changes no reading; any row not proportional to $(1,\ 2)$, such as a reading of z_1 alone, restores rank two.)

Exercise 19.2. *Derivation.* Partition the state into a block a of k entries and the rest b, with the information matrix in the blocks $\mathbf{\Lambda}_{aa}$, $\mathbf{\Lambda}_{ab}$, $\mathbf{\Lambda}_{bb}$. (a) Derive the marginal covariance of block a and the conditional mean and covariance of a given b. (b) Check (a) on Exercise 19.1. ((a) $\boldsymbol{\delta}^{\mathsf{T}}\mathbf{\Lambda}\boldsymbol{\delta} = \boldsymbol{\delta}_a^{\mathsf{T}}\mathbf{\Lambda}_{aa}\boldsymbol{\delta}_a + 2\boldsymbol{\delta}_a^{\mathsf{T}}\mathbf{\Lambda}_{ab}\boldsymbol{\delta}_b + \boldsymbol{\delta}_b^{\mathsf{T}}\mathbf{\Lambda}_{bb}\boldsymbol{\delta}_b$; at fixed $\boldsymbol{\delta}_b$, $\mathrm{Cov}(\mathbf{z}_a \mid \mathbf{z}_b) = \mathbf{\Lambda}_{aa}^{-1}$ and $\mathbb{E}[\mathbf{z}_a \mid \mathbf{z}_b] = \hat{\mathbf{z}}_a - \mathbf{\Lambda}_{aa}^{-1}\mathbf{\Lambda}_{ab}\boldsymbol{\delta}_b$; completing the square in $\boldsymbol{\delta}_b$ and integrating leaves $\exp[-\frac{1}{2}\boldsymbol{\delta}_a^{\mathsf{T}}(\mathbf{\Lambda}_{aa} - \mathbf{\Lambda}_{ab}\mathbf{\Lambda}_{bb}^{-1}\mathbf{\Lambda}_{ba})\boldsymbol{\delta}_a]$, so $\mathbf{\Sigma}_{aa} = (\mathbf{\Lambda}_{aa} - \mathbf{\Lambda}_{ab}\mathbf{\Lambda}_{bb}^{-1}\mathbf{\Lambda}_{ba})^{-1}$. (b) $\Sigma_{11} = 1/(4 - 4^2/5) = 1/0.8 = 1.25\,\mathrm{K}^2$, and $\mathrm{Var}(z_1 \mid z_2) = 1/4$.)

Exercise 19.3. *Derivation.* (a) Show that adding one row with normalized Jacobian row $\mathbf{a}$ never increases any posterior variance, and state when a variance is unchanged. (b) Deduce inequality (19.25). (c) Add to Exercise 19.1 a prior row $z_1 \sim \mathcal{N}(0,\ 1^2)$ and compute the new spreads. ((a) $\mathbf{\Lambda}' = \mathbf{\Lambda} + \mathbf{a}\mathbf{a}^{\mathsf{T}}$; the matrix-inversion lemma gives $\mathbf{\Sigma}' = \mathbf{\Sigma} - \mathbf{\Sigma}\mathbf{a}\mathbf{a}^{\mathsf{T}}\mathbf{\Sigma}/(1 + \mathbf{a}^{\mathsf{T}}\mathbf{\Sigma}\mathbf{a})$, so $\Sigma'_{ii} = \Sigma_{ii} - (\mathbf{\Sigma}\mathbf{a})_i^2/(1 + \mathbf{a}^{\mathsf{T}}\mathbf{\Sigma}\mathbf{a}) \le \Sigma_{ii}$, with equality iff $(\mathbf{\Sigma}\mathbf{a})_i = 0$: the new row is uncorrelated with entry i under the posterior. (b) The posterior is the prior with the reading rows added one at a time, each step non-increasing. (c) $\mathbf{a} = (1,\ 0)$, $\mathbf{\Sigma}\mathbf{a} = (1.25,\ -1)$, $\mathbf{a}^{\mathsf{T}}\mathbf{\Sigma}\mathbf{a} = 1.25$: $\Sigma'_{11} = 1.25 - 1.5625/2.25 = 0.556$ and $\Sigma'_{22} = 1 - 1/2.25 = 0.556\,\mathrm{K}^2$; directly, $\mathbf{\Lambda}' = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$ and $\mathbf{\Sigma}' = \frac{1}{9}\begin{bmatrix} 5 & -4 \\ -4 & 5 \end{bmatrix}$.)

Exercise 19.4. *Derivation.* (a) From equation (19.21) derive that the spread along the right singular vector $\mathbf{v}_j$ is $1/s_j$ and that the objective rises by one half at the distance $1/s_j$ along it. (b) Show that $\Sigma_{ii} = \sum_j V_{ij}^2/s_j^2$, and evaluate it for the load of hall 2 from Table 19.5. (c) Which direction carries most of that variance, and what does that say about the meter that would reduce it? ((a) $\mathbf{v}_j^{\mathsf{T}}\mathbf{\Sigma}\mathbf{v}_j = \mathbf{v}_j^{\mathsf{T}}\mathbf{V}\,\mathrm{diag}(s^{-2})\mathbf{V}^{\mathsf{T}}\mathbf{v}_j = \mathbf{e}_j^{\mathsf{T}}\,\mathrm{diag}(s^{-2})\mathbf{e}_j = s_j^{-2}$; along $t\mathbf{v}_j$ the quadratic form is $t^2\mathbf{v}_j^{\mathsf{T}}\mathbf{\Lambda}\mathbf{v}_j = s_j^2 t^2$, so $\frac{1}{2}s_j^2 t^2 = \frac{1}{2}$ at $t = 1/s_j$. (b) $\Sigma_{ii} = \mathbf{e}_i^{\mathsf{T}}\mathbf{V}\,\mathrm{diag}(s^{-2})\mathbf{V}^{\mathsf{T}}\mathbf{e}_i = \sum_j V_{ij}^2 s_j^{-2}$; for Q_2, $0.414^2 \times 1.106^2 + 0.910^2 \times 2.530^2 = 0.209 + 5.30 = 5.51\,\mathrm{kW}^2$, the other three terms below 10^{-3}, against 5.505 in Table 19.3. (c) $\mathbf{v}_5$, 96 %: a reading whose Jacobian row lies along the split of the load, such as a load meter on hall 2.)

Exercise 19.5. *Application.* Two halls with the conductances of this chapter pinned, $G_1 = 5.49$ and $G_2 = 4.055\,\mathrm{kW/K}$, share one temperature sensor in their common return header, which reads the mean of the two hall temperatures for equal return flows, with $\sigma = 0.5\,\mathrm{K}$, and the supply header has its sensor, $\sigma = 0.3\,\mathrm{K}$. There are no load meters and the loads have flat priors; the state is $\mathbf{z} = (Q_1, Q_2, T_s)^{\mathsf{T}}$, the projection of Chapter 17 along the two hall balances, so that the return sensor's row compares its reading with $T_r(\mathbf{z}) = T_s + \frac{1}{2}(Q_1/G_1 + Q_2/G_2)$. (a) Write the design columns, their rank, the observable combinations and the unobservable direction. (b) Which single meter restores full rank: a load meter on hall 1 or on hall 2 ($\sigma = 1.5\,\mathrm{kW}$), or the chiller heat meter on $Q_1 + Q_2$ ($\sigma = 2\,\mathrm{kW}$)? Give the spreads of the loads in each case. (c) What changes if the two conductances are equal? ((a) Rows $(0,\ 0,\ 3.333)$ and $(1/(2G_1\sigma),\ 1/(2G_2\sigma),\ 2) = (0.182,\ 0.247,\ 2)$: rank two of three. Observable: T_s and the mean rise $\frac{1}{2}(Q_1/G_1 + Q_2/G_2)$; unobservable: $(G_1,\ -G_2,\ 0) \propto (0.804,\ -0.594,\ 0)$, load moved from hall 2 to hall 1 in the ratio of the conductances, which leaves the mixed temperature

unchanged. (b) Each candidate gives rank three, but not equally: a hall-1 meter gives spreads 1.50 and 4.86 kW, a hall-2 meter 6.72 and 1.50 kW, and the chiller meter 19.6 and 19.0 kW with correlation -0.995 and smallest singular value 0.037, because its row $(1,\ 1)$ is nearly parallel to the return sensor's $(1/G_1,\ 1/G_2) \propto (1,\ 1.35)$. Either load meter restores full rank usefully; the chiller meter only formally. (c) With $G_1 = G_2$ the chiller meter's row is exactly parallel to the return sensor's on the loads, the rank stays two, and only a hall meter helps.)

Exercise 19.6. *Application.* Two identical halls have nameplate priors $Q_i \sim \mathcal{N}(28,\ 5^2)$ kW, and the only reading on their loads is the chiller heat meter on $Q_1 + Q_2$ with $\sigma_P = 2$ kW. (a) Compute the posterior covariance, the marginal spreads and the ratio of prior to posterior spread. (b) Compute the spreads of $Q_1 + Q_2$ and $Q_1 - Q_2$ and compare them with their prior spreads. (c) Is either load observed? State what the rank test and the shape of the posterior say, against the marginals. ((a) Rows $(0.5,\ 0.5)$, $(0.2,\ 0)$, $(0,\ 0.2)$; $\mathbf{\Lambda} = \begin{bmatrix} 0.29 & 0.25 \\ 0.25 & 0.29 \end{bmatrix}$, $\det = 0.0216$, $\mathbf{\Sigma} = \begin{bmatrix} 13.43 & -11.57 \\ -11.57 & 13.43 \end{bmatrix}$ kW^2; spreads 3.66 kW each, ratio $5/3.66 = 1.37$, correlation -0.86. (b) $\sqrt{2(13.43 - 11.57)} = 1.93$ kW for the sum, against the prior 7.07 kW; $\sqrt{2(13.43 + 11.57)} = \sqrt{50} = 7.07$ kW for the difference, exactly its prior $\sqrt{2 \times 25}$. (c) Neither load is observed on its own: the design column $(0.5,\ 0.5)$ has rank one and the direction $(1,\ -1)$ is unobservable. Each marginal is small and has shrunk by a factor 1.37, yet the difference is unidentified, its spread returned unchanged; the marginals reflect the sum, which the meter fixes, shared between two entries.)

Chapter 20

Propagating Uncertainty to Engineering Outputs

20.1 Hall 1 at twice its present load

The operator of the two data halls of Chapter 17 plans to raise the heat load of hall 1 from about 31 kW to 60 kW, and the equipment in that hall is rated for at most 25°C. The record is the one fitted in Chapters 18 and 19: five readings taken at the present loads, three handbook priors, and from them a conductance of hall 1's heat-removal path of 5.49 kW/K with the supply water at 12.11°C. The first reading of the record is a line of arithmetic. Twice the load through the same path doubles the temperature rise, $60/5.49 = 10.93$ K above the supply, so the hall would run at 23.0°C, two kelvin under the limit, and the load can be moved.

The engineer who must sign for the move asks what the two kelvin rest on. The conductance was inferred at half the scenario load from one temperature sensor and two meters, and at 60 kW its uncertainty is multiplied by the temperature rise; the supply temperature under the hall is uncertain too, and the same hall reading ties it to the conductance. The arithmetic does not say how likely the hall is to pass 25°C, which uncertain quantity sets that probability, how well the other outputs the operator watches are known, or which added sensor would narrow the answer most. The question of this chapter is the engineer's,

whether hall 1 at 60 kW*, predicted at* 23.0°C *with two kelvin to spare under its* 25°C *rating, is safe to sign for, and how sure that answer is.*

Jargon: derived quantity, scenario node, delta method, variance budget, cross term, credible interval, threshold probability, pushforward, second-order correction, importance reweighting, effective sample size, preposterior, matrix-inversion lemma, rank-one update

A *derived quantity* is a function $f(\mathbf{z})$ of the state that is not itself an entry of it, such as a temperature at a load not operated or a ratio of two conductances. A *scenario node* is a node added to the plant graph for a question about a condition that has not been observed; it shares entries of the state with the observed nodes and carries no reading. The *delta method* replaces f by its tangent plane at $\hat{\mathbf{z}}$, so that its variance is $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ with $\mathbf{g}$ the gradient of f. The *variance budget* is that quadratic form written out term by term, one term per entry and one *cross term* per pair of correlated entries. A *credible interval* contains the unknown with a stated posterior probability; a *threshold probability* is the posterior probability that the unknown lies on one side of a limit. The *pushforward* of a distribution through a function is the distribution of the function's value when its argument has the original distribution; by draws, it is obtained by evaluating the function at each draw. A *second-order correction* keeps the quadratic term of the Taylor expansion of f. *Importance reweighting* turns draws from one distribution into averages under another by weighting each draw with the ratio of the two densities; the *effective sample size* is the number of equally weighted draws that would give the same precision. The *preposterior* is the posterior that a measurement not yet taken would produce, and the *matrix-inversion lemma*, in its rank-one form also called the Sherman–Morrison formula, gives the inverse of a matrix plus an outer product from the inverse of the matrix, the *rank-one update* of the covariance.

Key idea

Every engineering output of a fitted plant is a function of the state, and under the Laplace approximation its band, its interval and its threshold probability follow from three objects already in hand: the MAP state, the covariance and the gradient of the output. The variance budget names the entry that sets the band, and its cross terms can be negative. The linearization is checked against the full posterior, not only against draws of the Gaussian. A candidate measurement adds one outer product to the information matrix, and its value for an output is the squared correlation between the output and the reading it would return.

20.1.1 Outputs that are not entries of the state

An output that is an entry of the state needs no propagation: its posterior mean is $\hat{z}_i$ and its spread $\sqrt{\Sigma_{ii}}$, read from the diagonal of the covariance. The supply temperature T_s of the loop is such an output. The temperature of hall 1 at its present load is not an entry of the state on which the inference runs: it is an entry of the full state of Chapter 17, carried by a physics row, and on the reduced state of that chapter it is a function of the free entries. The temperature of hall 1 at a load it has never carried is not an entry either. Neither is the ratio of the two conductances, which compares the cooling paths of the two halls, nor the total heat load, which is what the chiller must remove. Each is a function $f(\mathbf{z})$ of several entries, and its uncertainty depends on the variances of those entries and on their covariances. Chapter 4 derived the delta method, equation (4.49), and applied it to a three-entry state whose exact physical rows leave one free direction; Chapter 8 applied it to the logit of an availability, a linear function of two independent entries, equation (8.29). The present plant has five entries, correlated with one another, and outputs that are nonlinear in them, so the general form is needed together with the checks that tell when it holds.

20.1.2 The plant and its record (step 0)

The plant of Chapter 17 has two halls, each with a heat load Q_i in kW and a heat-removal path of conductance G_i in kW/K to a supply header at temperature T_s, served by a chiller. The conductances are carried as log-conductances $g_i = \ln(G_i/G_{\text{ref}})$ with $G_{\text{ref}} = 1\,\text{kW/K}$, so that $G_i = \text{e}^{g_i}$ numerically. The full state of that chapter has eight entries, the five free quantities together with the two hall temperatures and the chiller's heat, and its three balances are physics rows of tolerance σ_p. Chapter 17, Section 17.2.1, projected it onto the null directions of the physics rows, and the inference of Part V runs on the resulting reduced state,

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \end{bmatrix} = \begin{bmatrix} g_1 \\ g_2 \\ Q_1 \\ Q_2 \\ T_s \end{bmatrix}, \tag{20.1}$$

on which each hall temperature is the function $T_i = T_s + Q_i\,\text{e}^{-g_i}$ given by its balance, and the chiller's heat the sum $Q_1 + Q_2$.

The record is that of Chapter 17, Tables 17.1 and 17.2: two hall-temperature readings, $y_{T_1} = 17.655°\text{C}$ and $y_{T_2} = 17.700°\text{C}$, with $\sigma_T = 0.5\,\text{K}$; a supply-temperature reading, 12.10°C with 0.3 K; the chiller's heat meter on $Q_1 + Q_2$, 52.5 kW with 2 kW; an electrical meter on hall 1's load, 31.0 kW with 1.5 kW; the handbook priors $g_i \sim \mathcal{N}(\ln 5,\ 0.3^2)$; and the set-point prior $T_s \sim \mathcal{N}(12,\ 1^2)$. The readings are illustrative, generated from known values with stated errors, as Chapter 17 explains. All five were taken at the present loads. Chapters 18 and 19 fitted the state to this record and measured its spread; their results enter this chapter as its posterior, Section 20.2.2.

20.1.3 The questions (step 1)

The operator asks four questions of the fitted plant, and step 1 writes each as a probability or an expectation over the state, with the unknown on the left of the bar and the evidence $\mathbf{y}$, the five readings and the three priors, on the right. Three of them concern hall 1 at a load of 60 kW, a condition never operated; its temperature there is written T_1^* and the load Q_1^*, a starred quantity being a scenario value as the notation chapter states. Table 20.1 lists the questions: the first column is the question in engineering language; the second the general form of its category in the notation of Chapter 2, for a quantity $f(\mathbf{z})$ derived from the state and a level f^* of it; the third its mathematical form for the plant; the fourth its category; and the last the subsection where it is answered. The first and fourth are estimation of derived quantities. The second, the operator's question about the hall at 60 kW, is a prediction under a threshold, a tail of the posterior of T_1^*. The third is estimation read as an attribution, the budget of a variance. Which sensor to add is not a row: it is the next measurement for these four questions, step 7. The atlas of Appendix A lists the same forms.

20.1.4 The scenario node and the rows (step 2)

The first three questions concern a condition never operated: hall 1 at a load of $Q_1^* = 60\,\text{kW}$, about twice its present load. Step 2 adds a node for it. The scenario node is a derived quantity, not a new entry: it has the same heat-removal path as hall 1, so it shares the entry g_1, and it rejects its heat to the same header at the same supply temperature, so it shares T_s; its load is pinned at the scenario value. Its steady balance is $Q_1^* = G_1\,(T_1^* - T_s)$, and solving for the temperature,

$$T_1^*(\mathbf{z}) = T_s + Q_1^*\,\text{e}^{-g_1}, \qquad Q_1^* = 60\,\text{kW}, \tag{20.2}$$

Table 20.1: The questions asked of the fitted two-hall plant, the general form of their category, their mathematical form for the plant, their category and where each is answered. $\mathbf{y}$ is the five readings and the three priors; $T_1^* = T_s + Q_1^* \, \mathrm{e}^{-g_1}$ is hall 1's temperature at the scenario load $Q_1^* = 60\,\mathrm{kW}$; $\mathbf{g}$ is the gradient of a derived quantity at the MAP state.

Engineering question	General form	Mathematical form	Category	Answered in
What will hall 1's temperature be at 60 kW, and how well is that known?	$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y})$; location and central interval	$p(T_1^* \mid \mathbf{y})$: location $T_1^*(\hat{\mathbf{z}})$, $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$, central 95 % interval	estimation of a derived quantity	20.3.1
Will the hall stay under the limit at 60 kW?	$\Pr(f(\mathbf{z}) \ge f^* \mid \mathbf{y})$	$\Pr(T_1^* < 25^\circ\mathrm{C} \mid \mathbf{y})$	prediction under a threshold	20.3.2
Which entry's uncertainty drives the answer?	$\mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y})$, term by term	the budget of $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$ for T_1^*	estimation (attribution)	20.3.3
What are the conductance ratio, the total load and hall 1's conductance, with bands?	location and central interval	$p(G_1/G_2 \mid \mathbf{y})$, $p(Q_1 + Q_2 \mid \mathbf{y})$, $p(G_1 \mid \mathbf{y})$	estimation of derived quantities	20.3.4

in °C, with the limit $T_{\max} = 25^\circ\mathrm{C}$. The scenario node adds no row: it carries no reading and no prior, so the posterior of the state is unchanged by it, and extending the graph costs nothing. What it adds is a function of the state whose distribution the questions ask about.

The modeling choices. Three modeling choices enter every answer. The path of hall 1 is taken to keep at 60 kW the conductance inferred at the present load of about 31 kW: this is what lets the scenario node share the entry g_1 and makes its temperature a function of the fitted state, and it fails if the conductance of the coil depends on the flow of chilled water and the flow on the load, which no reading taken at the present loads can test (Section 20.6.6). The chiller is taken to hold the supply temperature at the doubled load: this lets the node share T_s, and it fails for a chiller near its capacity, whose supply rises with the load (Section 20.6.6). The posterior of the reduced state is taken as its Laplace Gaussian and the output as its tangent plane at $\hat{\mathbf{z}}$: together these make the output normal, so that its band is one quadratic form; the first fails where the posterior is skewed along g_1, the second where e^{-g_1} curves across the band, and the checks of Section 20.4 test each separately.

Figure 20.1 draws the extended graph. The green boxes are free nodes and carry entries of the state: the two loads and the supply header with T_s. The two blue boxes are the halls, balance nodes whose temperatures are, on the reduced state, functions of the free entries through their balances. The solid green arrows are the physical paths: each load into its hall, each hall's heat to the header along the path whose log-conductance, an entry, labels it, and the header's heat to the chiller, a white box whose duty is, on the reduced state, the sum of the loads. The dashed gray arrows leave the graph for the five readings. The dashed blue box on the left is the scenario node, a derived quantity fed by a gray pinned load of 60 kW and joined to the header by a dashed path that carries the same g_1; it has no reading arrow.

The evidence enters as the eight rows of the reduced state, equation (17.15): the five readings, the hall readings through the balances of their halls, and the three priors. Which row acts on which entry is a property of the graph, and it fixes where the reduced Jacobian $\mathbf{J} = \mathbf{J}_\mathrm{e}\mathbf{N}$ of equation (17.18) can be nonzero, marked $*$; the scenario output's gradient has its own pattern,

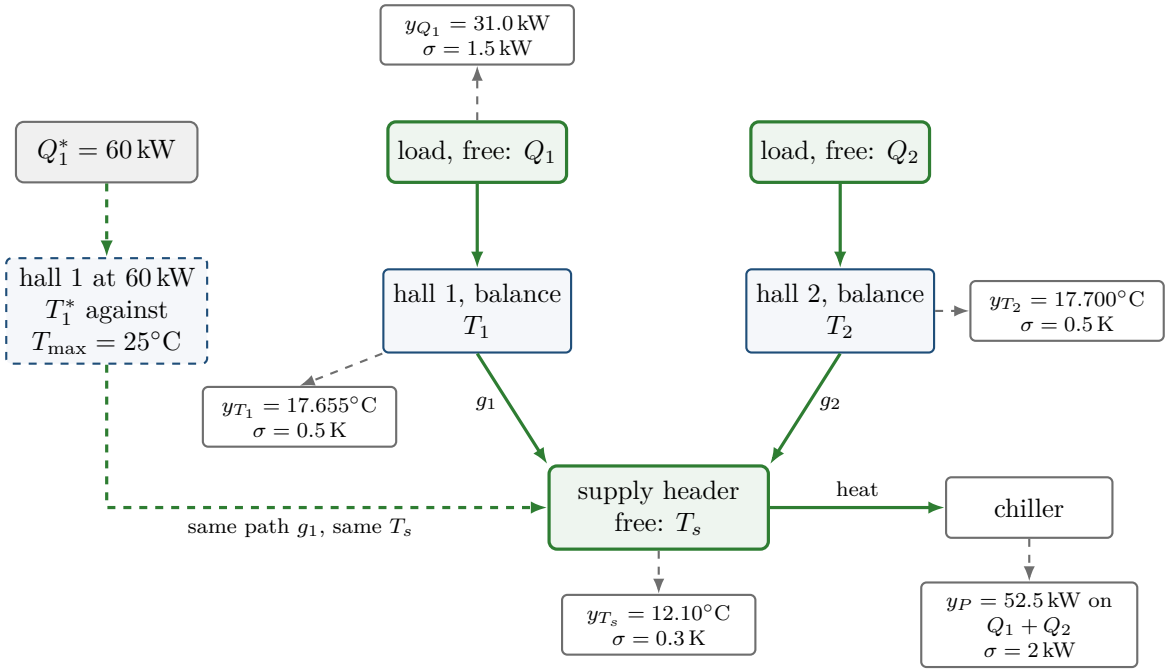


Figure 20.1: The two-hall plant of Chapter 17 extended for the questions of this chapter. Free (green): the loads Q_1, Q_2, the supply temperature T_s at the header, and the log-conductances g_1, g_2 of the two heat-removal paths, which label the solid arrows. Balance nodes (blue): the two halls, whose temperatures are, on the reduced state of Chapter 17, functions of the free entries. Pinned (gray): the scenario load $Q_1^* = 60\,\mathrm{kW}$; the chiller (white) is a sink whose duty is the sum of the loads. Observed (dashed gray arrows): the two hall temperatures, the supply temperature, the chiller's heat meter on $Q_1 + Q_2$ and the electrical meter on Q_1; the priors on g_1, g_2 and T_s are rows not drawn. Dashed box and path: the scenario node, hall 1 at 60 kW, a derived quantity sharing g_1 and T_s and carrying no reading; its output is $T_1^* = T_s + 60\,\mathrm{e}^{-g_1}$ against the limit $T_{\max} = 25°\mathrm{C}$.

set by the two entries the scenario node shares:

$$\mathbf{J} = \begin{bmatrix} * & 0 & * & 0 & * \\ 0 & * & 0 & * & * \\ 0 & 0 & 0 & 0 & * \\ 0 & 0 & * & * & 0 \\ 0 & 0 & * & 0 & 0 \\ * & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * \end{bmatrix} \begin{matrix} \text{hall-1 reading} \\ \text{hall-2 reading} \\ \text{supply reading} \\ \text{heat meter} \\ \text{load meter} \\ \text{prior on } g_1 \\ \text{prior on } g_2 \\ \text{prior on } T_s \end{matrix} \qquad \mathbf{g}^{\mathsf{T}} = \begin{bmatrix} * & 0 & 0 & 0 & * \end{bmatrix}, \tag{20.3}$$

with the columns in the order $(g_1, g_2, Q_1, Q_2, T_s)$ of equation (20.1). The columns of g_1 and T_s, which the scenario output depends on, meet in one row only, the hall-1 reading: it is the one row that sees both, and it sees Q_1 with them. The question of the hall at 60 kW adds no row to this pattern; it adds only the gradient.

20.2 The posterior of the state (step 3)

Every question of Table 20.1 asks for a function of the reduced state, and four objects of the fit recur in the specialized forms. The *MAP state* $\hat{\mathbf{z}}$ is the minimum of the objective that the eight rows put on the state; the *information matrix* $\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$ is the curvature of that objective at the minimum, with $\mathbf{J}$ the Jacobian of the normalized rows; the *Laplace covariance* $\mathbf{\Sigma} = \mathbf{\Lambda}^{-1}$ makes the posterior of the state approximately normal, $\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}, \mathbf{\Sigma})$; and a derived quantity f is carried by its *gradient* $\mathbf{g} = \partial f / \partial \mathbf{z}$ at $\hat{\mathbf{z}}$, with the variance $\mathbf{g}^{\mathsf{T}}\mathbf{\Sigma}\mathbf{g}$. All four questions are read from one posterior: the posterior of the reduced state that the eight rows of step 2 give, approximately the Laplace Gaussian. This section derives it and computes it for the plant; Section 20.3 reads every answer from it.

20.2.1 The posterior from the eight rows

The objects every question uses are those of the fit of Chapters 18 and 19, on the reduced state of Chapter 17. The state is the five-entry vector of equation (20.1), and the rows are the eight rows of equation (17.15), displayed again here,

$$\mathbf{r}(\mathbf{z}) = \begin{bmatrix} (T_s + Q_1\,\mathrm{e}^{-g_1} - y_{T_1})/\sigma_T \\ (T_s + Q_2\,\mathrm{e}^{-g_2} - y_{T_2})/\sigma_T \\ (T_s - y_{T_s})/\sigma_s \\ (Q_1 + Q_2 - y_P)/\sigma_P \\ (Q_1 - y_{Q_1})/\sigma_Q \\ (g_1 - g^0)/\sigma_g \\ (g_2 - g^0)/\sigma_g \\ (T_s - T_s^0)/\sigma_0 \end{bmatrix},$$

each divided by its standard deviation, with σ_s, σ_P and σ_Q those of the supply sensor, the heat meter and the load meter, $g^0 = \ln 5$ and $\sigma_g = 0.3$ the handbook prior and $T_s^0 = 12°\mathrm{C}$, $\sigma_0 = 1\,\mathrm{K}$ the set-point prior. The first row is the hall-1 reading row on the reduced state, equation (17.16). The objective is $\mathcal{L}(\mathbf{z}) = \frac{1}{2}\,\mathbf{r}^{\mathsf{T}}\mathbf{r}$, equation (18.9), and the posterior is $p(\mathbf{z} \mid \mathbf{y}) = \mathrm{e}^{-\mathcal{L}(\mathbf{z})}/c$ with c a constant. The Jacobian of the rows is the projection of Chapter 17, $\mathbf{J} = \mathbf{J}_{\mathrm{e}}\mathbf{N}$, equation (17.18), an 8×5 matrix with the pattern (20.3); its first line, the hall-1 reading, is

$$\frac{1}{\sigma_T}\,(-Q_1\mathrm{e}^{-g_1},\ 0,\ \mathrm{e}^{-g_1},\ 0,\ 1).$$

The fit solves the normal equations of Chapter 3,

$$(\mathbf{J}^\mathsf{T}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^\mathsf{T}\mathbf{r}, \qquad \mathbf{z}_{j+1} = \mathbf{z}_j + \Delta\mathbf{z},$$

from the prior point, as Chapter 18 iterated them to $\hat{\mathbf{z}}$. With every row normalized the information matrix is $\mathbf{\Lambda} = \mathbf{J}^\mathsf{T}\mathbf{J}$; with unnormalized rows and weights $w_k = 1/\sigma_k^2$ it is $\mathbf{J}^\mathsf{T}\mathbf{W}\mathbf{J}$, the same matrix. The two hall reading rows are nonlinear on the reduced state, through the products $Q_i e^{-g_i}$, so the exact curvature of $\mathcal{L}$ adds the term $\sum_k r_k\,\partial^2 r_k/\partial\mathbf{z}\,\partial\mathbf{z}^\mathsf{T}$ over those rows; Chapter 18 evaluated it at $\hat{\mathbf{z}}$ and kept the Gauss–Newton curvature. The Laplace covariance is

$$\mathbf{\Sigma} = \mathbf{\Lambda}^{-1} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1} \quad \text{at } \hat{\mathbf{z}},$$

and by equation (17.19) it is the free-entry block of the covariance of the full state with its physics rows at tolerance σ_p, in the limit $\sigma_p \to 0$; Section 20.4.5 checks that limit on the scenario output. The fit returns each entry $\hat{z}_i$ with $\sqrt{\Sigma_{ii}}$ and each correlation $\mathrm{corr}(z_i, z_j)$; every derived quantity of the chapter is read from $\hat{\mathbf{z}}$ and $\mathbf{\Sigma}$ through its gradient, and no new fit is needed for any question: a band costs one quadratic form and a threshold probability one evaluation of Φ.

20.2.2 The posterior of the two halls

The plant's inputs are the five readings and the three priors of the record, Section 20.1.2. The Gauss–Newton iterations of Chapter 18 end at the MAP state of Table 18.4: $G_1 = 5.490\,\mathrm{kW/K}$ ($g_1 = 1.7030$), $G_2 = 4.055\,\mathrm{kW/K}$ ($g_2 = 1.3999$), $Q_1 = 30.69\,\mathrm{kW}$, $Q_2 = 22.23\,\mathrm{kW}$ and $T_s = 12.111°\mathrm{C}$. The posterior spreads $\sqrt{\Sigma_{ii}}$ of Chapter 19, Table 19.3, are 0.1067, 0.1370, 1.452 kW, 2.346 kW and 0.2801 K, and four correlations matter below: $\mathrm{corr}(g_1, T_s) = 0.434$, positive as Section 20.3.3 reads it off the hall-1 row; $\mathrm{corr}(g_1, g_2) = -0.029$; $\mathrm{corr}(Q_1, Q_2) = -0.572$; $\mathrm{corr}(g_2, Q_2) = 0.691$. The entries of $\mathbf{\Sigma}$ that the hall-1 scenario uses are

$$\Sigma_{11} = 0.011382, \qquad \Sigma_{15} = 0.012974\,\mathrm{K}, \qquad \Sigma_{55} = 0.078433\,\mathrm{K}^2, \tag{20.4}$$

the variance of g_1, its covariance with T_s and the variance of T_s.

20.2.3 The Laplace pushforward

Every answer of step 4 is computed from $\hat{\mathbf{z}}$ and $\mathbf{\Sigma}$ through the tangent plane of its output at $\hat{\mathbf{z}}$, the delta method. Two routes go beyond the tangent plane and give the values set beside those answers. The first, this subsection, removes the tangent plane of the output while the Gaussian posterior is kept, in three forms: the second-order correction to the mean, the pushforward of an output of two entries as one integral, and draws. The second, the next subsection, also removes the Gaussian and reaches the full posterior.

The second-order mean. Write the deviation from the MAP state as $\mathbf{d} = \mathbf{z} - \hat{\mathbf{z}}$, so that $\mathbf{d} \mid \mathbf{y} \approx \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$. The first-order expansion $f(\mathbf{z}) \approx f(\hat{\mathbf{z}}) + \mathbf{g}^\mathsf{T}\mathbf{d}$, with $\mathbf{g}$ the gradient of f at $\hat{\mathbf{z}}$, gives the mean $f(\hat{\mathbf{z}})$. The next term of the Taylor expansion is

$$f(\mathbf{z}) \approx f(\hat{\mathbf{z}}) + \mathbf{g}^\mathsf{T}\mathbf{d} + \tfrac{1}{2}\,\mathbf{d}^\mathsf{T}\,\nabla^2 f(\hat{\mathbf{z}})\,\mathbf{d}, \tag{20.5}$$

with $\nabla^2 f(\hat{\mathbf{z}})$ the matrix of second derivatives of f at $\hat{\mathbf{z}}$. Take the posterior expectation. The linear term vanishes because $\mathbb{E}[\mathbf{d} \mid \mathbf{y}] = \mathbf{0}$ under the Gaussian centered at $\hat{\mathbf{z}}$. The quadratic form is a scalar, equal to its own trace, and the trace is invariant under a cyclic permutation:

$$\mathbf{d}^\mathsf{T}\,\nabla^2 f\,\mathbf{d} = \mathrm{tr}\big(\mathbf{d}^\mathsf{T}\,\nabla^2 f\,\mathbf{d}\big) = \mathrm{tr}\big(\nabla^2 f\,\mathbf{d}\mathbf{d}^\mathsf{T}\big).$$

The trace is linear, so the expectation passes inside it, and $\mathbb{E}[\mathbf{d}\mathbf{d}^\mathsf{T} \mid \mathbf{y}] = \mathbf{\Sigma}$:

$$\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}] \approx f(\hat{\mathbf{z}}) + \tfrac{1}{2}\,\mathrm{tr}\big(\mathbf{\Sigma}\,\nabla^2 f(\hat{\mathbf{z}})\big). \tag{20.6}$$

The second term is the bias of the plug-in value $f(\hat{\mathbf{z}})$ as an estimate of the posterior mean of f. It is positive for a convex output and negative for a concave one, and it is the matrix form of the leading Jensen gap of Chapter 9, equation (9.22), where the output was the survival function $\mathrm{e}^{-\lambda t}$ and the state the single rate λ. Its size relative to $\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}$ says how far the posterior of the output departs from a normal centered at $f(\hat{\mathbf{z}})$; a ratio above a few percent calls for the routes below. The derivation assumed that $\hat{\mathbf{z}}$ is the posterior mean, which holds for the Laplace Gaussian but not, in general, for the full posterior; the correction accounts for the curvature of the output and for nothing else. The scenario temperature has one nonzero second derivative, $\partial^2 T_1^*/\partial g_1^2 = Q_1^* \mathrm{e}^{-\hat{g}_1}$, the magnitude of its first derivative in g_1, the rise of the hall at the scenario load.

The pushforward of an output of two entries. The scenario output (20.2) depends on two entries only, $z_1 = g_1$ and $z_5 = T_s$. Under the Gaussian posterior the marginal of any subset of entries is Gaussian with the corresponding block of $\boldsymbol{\Sigma}$, so the other three entries integrate out without calculation, and (g_1, T_s) has mean $(\hat{g}_1, \hat{T}_s)$ and the block of Σ_{11}, Σ_{15} and Σ_{55}. Conditional on $g_1 = x$, the supply temperature is normal, with the conditional mean and variance of a bivariate normal (the Schur complement of Chapters 3 and 19):

$$T_s \mid g_1 = x,\ \mathbf{y} \sim \mathcal{N}\big(m(x),\ \Sigma_{5|1}\big), \qquad m(x) = \hat{T}_s + \frac{\Sigma_{15}}{\Sigma_{11}}\,(x - \hat{g}_1), \qquad \Sigma_{5|1} = \Sigma_{55} - \frac{\Sigma_{15}^2}{\Sigma_{11}}. \tag{20.7}$$

By the law of total probability over g_1, the event $T_1^* < T_{\max}$ at $g_1 = x$ is the event $T_s < T_{\max} - Q_1^* \mathrm{e}^{-x}$, whose conditional probability is a normal tail:

$$\Pr(T_1^* < T_{\max} \mid \mathbf{y}) = \int_{-\infty}^{\infty} \Phi\left(\frac{T_{\max} - Q_1^*\,\mathrm{e}^{-x} - m(x)}{\sqrt{\Sigma_{5|1}}}\right) p(g_1 = x \mid \mathbf{y})\,\mathrm{d}x, \tag{20.8}$$

with $p(g_1 = x \mid \mathbf{y})$ the normal density of mean $\hat{g}_1$ and variance Σ_{11}. The integrand is smooth and the Gaussian weight decays fast, so a quadrature over $\hat{g}_1 \pm 10\sqrt{\Sigma_{11}}$ evaluates it to any required accuracy. Equation (20.8) is the Laplace pushforward of the output: it removes the linearization of the output and keeps the Gaussian posterior, and it is exact only to the extent that the Gaussian is. Its quantiles follow by solving $\Pr(T_1^* < t \mid \mathbf{y}) = q$ for t, its density is its derivative in $T_{\max}$, and its mean has a closed form: since $\mathbb{E}[\mathrm{e}^{-g_1}] = \mathrm{e}^{-\hat{g}_1 + \Sigma_{11}/2}$ for a normal g_1, by its moment-generating function at -1,

$$\mathbb{E}[T_1^* \mid \mathbf{y}] = \hat{T}_s + Q_1^*\,\mathrm{e}^{-\hat{g}_1}\,\mathrm{e}^{\Sigma_{11}/2} \qquad \text{(Laplace posterior)}. \tag{20.9}$$

The pushforward by draws. The procedure of Chapter 4, Section 4.4.7, applies to any output and any number of entries: draw $\mathbf{z}^{(j)} = \hat{\mathbf{z}} + \mathbf{L}\mathbf{e}^{(j)}$, $j = 1, \ldots, N$, with $\mathbf{L}$ the Cholesky factor of $\boldsymbol{\Sigma}$ and $\mathbf{e}^{(j)}$ standard normal vectors, evaluate $f(\mathbf{z}^{(j)})$, and read the threshold probability as the fraction $\hat{P}$ of draws below the limit, with standard error $\sqrt{\hat{P}(1-\hat{P})/N}$, and the quantiles as empirical quantiles. Like equation (20.8), this is the Laplace pushforward: it removes the linearization and keeps the Gaussian.

20.2.4 The full posterior by reweighting

Chapter 19 computed the exact posterior of this plant by importance sampling, Section 19.4.1, and found its band of g_1 longer than the Laplace band on the side of large conductances. The same device carries the draws of the Laplace Gaussian to the full posterior. The posterior is $p(\mathbf{z} \mid \mathbf{y}) = \mathrm{e}^{-\mathcal{L}(\mathbf{z})}/c$, with $\mathcal{L}$ the objective of Section 20.2.1, and the Laplace density is $p_{\mathrm{L}}(\mathbf{z}) = \mathrm{e}^{-\frac{1}{2}\mathbf{d}^{\mathsf{T}}\boldsymbol{\Lambda}\mathbf{d}}/c_{\mathrm{L}}$. At a draw $\mathbf{z}^{(j)}$ their ratio is

$$\frac{p(\mathbf{z}^{(j)} \mid \mathbf{y})}{p_{\mathrm{L}}(\mathbf{z}^{(j)})} = \frac{c_{\mathrm{L}}\,\mathrm{e}^{-\mathcal{L}(\hat{\mathbf{z}})}}{c}\,\zeta_j, \qquad \zeta_j = \exp\Big(-\big[\mathcal{L}(\mathbf{z}^{(j)}) - \mathcal{L}(\hat{\mathbf{z}})\big] + \tfrac{1}{2}\,\mathbf{d}^{(j)\mathsf{T}}\boldsymbol{\Lambda}\,\mathbf{d}^{(j)}\Big). \tag{20.10}$$

For a function ψ of the state, the average of ψ times this ratio over the draws estimates $\mathbb{E}[\psi(\mathbf{z}) \mid \mathbf{y}]$; for $\psi = 1$ the same average estimates 1, and dividing the two removes the unknown constant:

$$\mathbb{E}[\psi(\mathbf{z}) \mid \mathbf{y}] \approx \sum_{j=1}^{N} \bar{\zeta}_j \, \psi(\mathbf{z}^{(j)}), \qquad \bar{\zeta}_j = \frac{\zeta_j}{\sum_{l=1}^{N} \zeta_l}. \tag{20.11}$$

This is self-normalized importance reweighting, a standard result; with ψ the indicator of the event it gives the threshold probability under the full posterior, and with the weights accumulated in the order of the values $f(\mathbf{z}^{(j)})$ it gives the quantiles. The weights are computed from the log weights less their maximum, which leaves equation (20.11) unchanged and keeps the exponentials in range. Its precision follows from writing the error of a reweighted probability, with $\mathbf{1}_j$ the indicator of the event at draw j, as

$$\hat{P} - P = \frac{\sum_j \zeta_j \, (\mathbf{1}_j - P)}{\sum_l \zeta_l}.$$

The terms of the numerator are independent across draws and have mean zero, because the expectation of $\zeta(\mathbf{1} - P)$ under the Laplace density is proportional to $\mathbb{E}[\mathbf{1} - P \mid \mathbf{y}] = 0$; its variance is estimated by $\sum_j \zeta_j^2 (\mathbf{1}_j - \hat{P})^2$, and the denominator, a sum of many positive terms, is treated as fixed. Dividing,

$$\mathrm{se}(\hat{P}) \approx \Big(\sum_j \bar{\zeta}_j^2 \, (\mathbf{1}_j - \hat{P})^2 \Big)^{1/2}, \qquad N_{\mathrm{eff}} = \Big(\sum_j \bar{\zeta}_j^2 \Big)^{-1}, \tag{20.12}$$

the second being the effective sample size, the N for which $\sqrt{\hat{P}(1-\hat{P})/N}$ gives about the same standard error: it equals N when all weights are equal, that is, when the Laplace Gaussian is the posterior, and falls as the two densities part. The reweighting needs the objective at each draw and nothing else. It fails when a few draws carry most of the weight; when N_{eff} falls below about a tenth of N, a draw distribution with heavier tails, as Chapter 19 used, or a chain that samples the posterior directly is needed.

The pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 20.2.2 is what step 4 uses: every answer of the chapter is computed from it through the gradient of its output. The full posterior, by the reweighting (20.11), gives the value set beside each answer, with the gap, and the Laplace pushforward (20.8) shows how much of the gap the tangent plane alone accounts for.

20.3 The answers (step 4)

The questions of Table 20.1 are estimation of derived quantities (the first and the fourth, and the third read as an attribution) and a prediction under a threshold (the second). Each is answered from the posterior of Section 20.2: its engineering question, the general form of its category, the form it takes for the plant, the quantities of the posterior it uses, and the plant's numbers. The answers are read by the delta method, the three objects $\hat{\mathbf{z}}$, $\boldsymbol{\Sigma}$ and $\mathbf{g}$, and that is the value the chapter reports; the full posterior of Section 20.2.4 gives the value set beside it, with the gap, and the Laplace pushforward of Section 20.2.3 the value in between where it matters. The questions name the scenario load $Q_1^* = 60\,\mathrm{kW}$ and the limit $T_{\max} = 25°\mathrm{C}$.

20.3.1 The scenario temperature and its band

Engineering question. What will hall 1's temperature be at 60 kW, and how well is that known? The operator asks for the temperature the hall would settle at under the doubled load, and for the band that the record leaves on it.

General form. The estimation form of Chapter 2, equations (2.9) to (2.11), for a quantity $f(\mathbf{z})$ derived from the state: its posterior distribution, summarized by a location and a spread, with the central 95 % credible interval

$$\Pr(f(\mathbf{z}) \le f_L \mid \mathbf{y}) = 0.025, \qquad \Pr(f(\mathbf{z}) \le f_U \mid \mathbf{y}) = 0.975.$$

The location of the estimation form is the posterior mean $\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]$ or the median.

Specialized form. The question is equations (2.9) to (2.11) with the state $\mathbf{z}$ the reduced state (20.1), the derived quantity $f(\mathbf{z}) = T_1^*(\mathbf{z})$, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 20.2.2. For hall 1 at 60 kW the derived quantity is the scenario temperature, $f(\mathbf{z}) = T_1^*(\mathbf{z})$ of equation (20.2). Under the Laplace Gaussian of Chapter 19, $\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$, write the deviation from the MAP state as $\mathbf{d} = \mathbf{z} - \hat{\mathbf{z}}$, so that $\mathbf{d} \mid \mathbf{y} \approx \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$. The first-order Taylor expansion of f about $\hat{\mathbf{z}}$ is

$$f(\mathbf{z}) \approx f(\hat{\mathbf{z}}) + \mathbf{g}^{\mathsf{T}}\mathbf{d}, \qquad \mathbf{g} = \nabla f(\hat{\mathbf{z}}), \tag{20.13}$$

with $[\mathbf{g}]_i = \partial f / \partial z_i$ at $\hat{\mathbf{z}}$. (The entries of the gradient are written $[\mathbf{g}]_i$ in this chapter because g_1 and g_2 are the log-conductances.) For the scenario temperature the gradient has the pattern of equation (20.3),

$$\mathbf{g} = \left.\frac{\partial T_1^*}{\partial \mathbf{z}}\right|_{\hat{\mathbf{z}}} = \begin{bmatrix} -Q_1^* \, \mathrm{e}^{-\hat{g}_1} \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \tag{20.14}$$

two nonzero entries, in the columns of g_1 and T_s. The magnitude of the first equals the temperature rise of the hall at the scenario load: the lever through which the uncertainty of the conductance acts grows with the load. The second says that the supply temperature passes into the hall one for one. The same form at the present load, $T_1(\mathbf{z}) = T_s + Q_1 \mathrm{e}^{-g_1}$ with the gradient $(-Q_1 \mathrm{e}^{-\hat{g}_1},\ 0,\ \mathrm{e}^{-\hat{g}_1},\ 0,\ 1)^{\mathsf{T}}$, gives the hall temperature the record was taken at, against which the scenario's band is compared.

Chapter 4 took the variance of the right-hand side of equation (20.13), equation (4.49); for a threshold and an interval the distribution is needed, not only the variance. A linear combination of jointly normal variables is normal. To see it, take the moment-generating function of $X = \mathbf{g}^{\mathsf{T}}\mathbf{d}$: for a real t,

$$\mathbb{E}\big[\mathrm{e}^{tX}\big] = \mathbb{E}\big[\mathrm{e}^{(t\mathbf{g})^{\mathsf{T}}\mathbf{d}}\big] = \exp\!\big(\tfrac{1}{2}\,(t\mathbf{g})^{\mathsf{T}}\boldsymbol{\Sigma}\,(t\mathbf{g})\big) = \exp\!\big(\tfrac{1}{2}\,t^2\,\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}\big),$$

where the second equality is the moment-generating function of the multivariate normal $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$ evaluated at $t\mathbf{g}$. The last expression is the moment-generating function of $\mathcal{N}(0, \mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g})$, and a moment-generating function that exists near zero determines its distribution. Therefore, to first order,

$$f(\mathbf{z}) \mid \mathbf{y} \approx \mathcal{N}\big(f(\hat{\mathbf{z}}),\ \mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}\big), \qquad \mathrm{sd}(f \mid \mathbf{y}) \approx \sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}, \tag{20.15}$$

and the central 95 % credible interval, as in equation (4.53), is

$$\Big[f(\hat{\mathbf{z}}) - 1.960\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}},\ f(\hat{\mathbf{z}}) + 1.960\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}\Big]. \tag{20.16}$$

The location the delta method reports is therefore the plug-in value $f(\hat{\mathbf{z}})$, which is the mean and the median of the normal output (20.15) and equals $\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]$ to first order; the second-order term (20.6) is the difference under the Laplace posterior. The value beside it comes from the full posterior: its mean, median, spread and central interval by the reweighting (20.11), with the Laplace pushforward (20.9) and (20.8) in between. Two approximations stand behind these lines, and they are distinct: the Gaussian posterior of the state, which is the Laplace approximation of Chapter 19, and the tangent plane of the output, equation (20.13). When f is linear the second is exact; when the posterior is Gaussian the first is.

Building blocks. The MAP state $\hat{\mathbf{z}}$ and the covariance $\boldsymbol{\Sigma}$ of the reduced state (Section 20.2.1); the gradient (20.14) at $\hat{\mathbf{z}}$; the quadratic form $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$, which needs only the block of $\boldsymbol{\Sigma}$ in the rows and columns of g_1 and T_s, equation (20.4); and the normal quantile 1.960. For the values beside them: the 200 000 draws of $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ reweighted to the full posterior by equation (20.11), and the Laplace pushforward's mean (20.9) and integral (20.8).

The scenario temperature at 60 kW. At the MAP state the thermal resistance of hall 1's path is $\mathrm{e}^{-\hat{g}_1}/G_{\mathrm{ref}} = 1/(5.490\,\mathrm{kW/K}) = 0.18214\,\mathrm{K/kW}$, and the scenario output (20.2) is

$$T_1^*(\hat{\mathbf{z}}) = 12.111 + 60 \times 0.18214 = 12.111 + 10.928 = 23.040^{\circ}\mathrm{C},$$

a rise of 10.93 K above the supply, twice the rise of 5.59 K the fit gives at the present load of 30.69 kW. The gradient (20.14) is

$$\mathbf{g} = (-10.928\,\mathrm{K},\ 0,\ 0,\ 0,\ 1)^{\mathsf{T}} :$$

a log-conductance higher by 0.01, a conductance higher by about 1 %, lowers the temperature at 60 kW by 0.109 K. Table 20.2 collects the fitted state with the gradient: each row is an entry of the state, with its MAP value, its posterior spread, its gradient entry and the spread $\sqrt{V_i} = |[\mathbf{g}]_i|\sqrt{\Sigma_{ii}}$ that it alone would put on T_1^*.

Table 20.2: The fitted state of the two-hall plant (Chapters 18 and 19) and the gradient of the scenario output $T_1^* = T_s + 60\,\mathrm{e}^{-g_1}$ at the MAP state.

Entry	MAP value	$\sqrt{\Sigma_{ii}}$	$[\mathbf{g}]_i$	$\sqrt{V_i}$ [K]
$z_1 = g_1$ (hall 1)	1.7030	0.1067	−10.928 K	1.166
$z_2 = g_2$ (hall 2)	1.3999	0.1370	0	0
$z_3 = Q_1$	30.69 kW	1.452 kW	0	0
$z_4 = Q_2$	22.23 kW	2.346 kW	0	0
$z_5 = T_s$	12.111°C	0.280 K	1	0.280
T_1^* at 60 kW	23.040°C	1.074 K	–	–

With the covariance block (20.4), the variance of the output, the quadratic form $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ written entry by entry, is

$$\begin{aligned}\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g} &= (-10.928)^2 \times 0.011382 + 1^2 \times 0.078433 + 2 \times (-10.928) \times 1 \times 0.012974 \\ &= 1.3594 + 0.0784 - 0.2836 = 1.1543\,\mathrm{K}^2,\end{aligned}$$

so $\mathrm{sd}(T_1^* \mid \mathbf{y}) = 1.074\,\mathrm{K}$, and the central 95 % interval (20.16) is $23.040 \pm 1.960 \times 1.0744 = 23.040 \pm 2.106$,

$$T_1^* \in [20.93,\ 25.15]^{\circ}\mathrm{C}.$$

The temperature of hall 1 at 60 kW is 23.04°C with a spread of 1.07 K, more than twice the spread of 0.48 K of its temperature at the present load.

The value beside it comes from the full posterior. The 200 000 draws of $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$, from the PCG64 generator seeded with 17, reweighted by equation (20.11) with an effective sample size of 182 933, give the mean 22.98°C, the median 22.95°C, the standard deviation 1.071 K and the central 95 % band [20.97, 25.16]°C: the center lies 0.059 K below $T_1^*(\hat{\mathbf{z}})$, the spread is 0.003 K narrower, and the delta band's ends lie 0.03 and 0.01 K below the full ones. The Laplace pushforward, which keeps the Gaussian and removes only the tangent plane, moves the other way: by equation (20.9) its mean is $23.0399 + 10.928 \times (\mathrm{e}^{0.011382/2} - 1) = 23.102^{\circ}\mathrm{C}$, 0.062 K above the delta center, with the band [21.15, 25.39]°C from equation (20.8). For the operator the gap is immaterial: every route places the hall near 23°C with a spread of about 1.07 K.

20.3.2 The threshold probability

Engineering question. Will the hall stay under the limit at 60 kW? The equipment of hall 1 is rated for 25°C, and the operator asks how probable it is that the hall stays below it once the load is doubled.

General form. The threshold form of Chapter 2, equation (2.12): the posterior probability of the event that the requirement holds,

$$\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) \geq f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

with an upper limit written as $f = -T$ and $f^* = -T_{\max}$.

Specialized form. The question is equation (2.12) with the state $\mathbf{z}$ the reduced state (20.1), $f(\mathbf{z}) = -T_1^*(\mathbf{z})$ and $f^* = -T_{\max} = -25$°C, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 20.2.2. One choice enters here: the reported route is the normal tail of the linearized output. It buys the probability as one evaluation of Φ at a standardized distance, and it fails in the tail first, where the curvature of the output and the shape of the posterior weigh most; the full posterior gives the value set beside it.

For the hall the event is $T_1^* < T_{\max}$, the distribution of T_1^* is the pushforward of the posterior through equation (20.2), and the operator's question is

$$\Pr(T_1^* < T_{\max} \mid \mathbf{y}) = \int \mathbf{1}\{T_s + Q_1^*\, \mathrm{e}^{-g_1} < T_{\max}\}\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \tag{20.17}$$

the posterior mass of the states whose scenario temperature lies below the limit, as Chapter 17 wrote it, equation (17.2). In Chapter 2 the corresponding event depended on one entry and could be rewritten exactly as an event on it; here it depends on two, g_1 and T_s, and the boundary $T_s + 60\,\mathrm{e}^{-g_1} = 25$ is a curve in their plane. Under the normal output (20.15) the event is a half-line, and standardizing, as in equation (4.54),

$$\Pr(f < T_{\max} \mid \mathbf{y}) \approx \Phi\left(\frac{T_{\max} - f(\hat{\mathbf{z}})}{\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}}\right), \tag{20.18}$$

with Φ the standard normal cumulative distribution; the probability of exceeding the limit is its complement.

Building blocks. The output $f(\hat{\mathbf{z}})$ and the quadratic form $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ of the previous question; the limit $T_{\max} = 25$°C; Φ. For the values beside them, the integral (20.17) without the tangent plane: under the full posterior by the reweighted draws (20.11) with their standard error (20.12), and under the Laplace posterior by equation (20.8).

The threshold probability at the limit. The standardized distance from the center to the limit is $(25 - 23.040)/1.0744 = 1.824$, and by equation (20.18)

$$\Pr(T_1^* < 25°\mathrm{C} \mid \mathbf{y}) \approx \Phi(1.824) = 0.966. \tag{20.19}$$

The probability that the hall exceeds the limit at 60 kW is 0.034. The upper end of the interval, 25.15°C, lies above the limit, as it must when the probability of staying below is smaller than 0.975. At the less strict limit of 27.5°C the standardized distance is 4.151 and the probability is $\Phi(4.151) = 0.99998$.

The value beside it comes from the full posterior: the reweighted draws give $\Pr(T_1^* < 25°\mathrm{C} \mid \mathbf{y}) = 0.966 \pm 0.0004$, the delta value within the draws' standard error. The Laplace pushforward,

equation (20.8) with the conditional slope $\Sigma_{15}/\Sigma_{11} = 1.140\,\mathrm{K}$ and spread $\sqrt{\Sigma_{5|1}} = 0.2523\,\mathrm{K}$, gives 0.952, an exceedance of 0.048 against 0.034: removing the tangent plane without the shape of the posterior moves the answer away from the full posterior. The agreement at 25°C is a property of this limit, near 1.8 standard deviations from the center: at 24°C the delta method gives 0.814 against the full posterior's 0.830, and at 27.5°C 0.99998 against 0.9999.

20.3.3 The variance budget

Engineering question. Which entry's uncertainty drives the answer? The engineer asks which of the uncertain quantities, the conductance or the supply temperature, sets the band on the hall at 60 kW, and so which one a measurement should aim at.

General form. The variance of a derived quantity, the second of equations (2.10),

$$\mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y}) = \int \left(f(\mathbf{z}) - \mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]\right)^2 p(\mathbf{z} \mid \mathbf{y})\,\mathrm{d}\mathbf{z},$$

read term by term over the entries it comes from.

Specialized form. The question is the second of equations (2.10) with the state $\mathbf{z}$ the reduced state (20.1), $f(\mathbf{z}) = T_1^*(\mathbf{z})$, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 20.2.2, its variance read term by term. One choice enters here: the budget is written in the coordinates in which the state carries the entries, the log-conductance and the supply temperature. It buys one term per entry and one per pair, each with a sign that can be read; it fails as a ranking of measurements, since a reparameterization of the state changes the terms and leaves the total, and the ranking that decides a measurement is the preposterior of Section 20.6.1, which does not depend on the coordinates.

Under the delta method the variance is $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$. Writing the quadratic form entry by entry, and collecting the two equal off-diagonal products $[\mathbf{g}]_i[\mathbf{g}]_j\Sigma_{ij}$ and $[\mathbf{g}]_j[\mathbf{g}]_i\Sigma_{ji}$,

$$\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g} = \sum_i [\mathbf{g}]_i^2\,\Sigma_{ii} + 2\sum_{i<j} [\mathbf{g}]_i\,[\mathbf{g}]_j\,\Sigma_{ij}. \tag{20.20}$$

This is the variance budget of the output, the general form of equation (4.50). A band built from the marginal spreads $\sqrt{\Sigma_{ii}}$ alone keeps the first sum and drops the second; it is wrong by as much as the cross terms weigh, in either direction, so the object kept from the fit is the covariance, not its diagonal. Let $\mathrm{corr}(z_i, z_j) = \Sigma_{ij}/\sqrt{\Sigma_{ii}\Sigma_{jj}}$ be the posterior correlation of entries i and j, and define

$$V_i = [\mathbf{g}]_i^2\,\Sigma_{ii}, \tag{20.21}$$

the diagonal term of entry i, the variance that entry alone puts on the output; $\sqrt{V_i} = |[\mathbf{g}]_i|\sqrt{\Sigma_{ii}}$ is its spread. The cross term of the pair (i, j) is

$$2[\mathbf{g}]_i[\mathbf{g}]_j\Sigma_{ij} = 2\,\mathrm{corr}(z_i, z_j)\,\mathrm{sgn}([\mathbf{g}]_i[\mathbf{g}]_j)\,\sqrt{V_iV_j}, \tag{20.22}$$

with sgn the sign. The cross term is negative when the two entries are positively correlated and enter the output with opposite signs, or negatively correlated and enter with the same sign. In either case the readings have fixed a combination of the two entries that moves the output less than either entry does alone.

A bound on a cross term. Since $|\,\mathrm{corr}(z_i, z_j)| \le 1$ and, for any two non-negative numbers, $2ab \le a^2 + b^2$,

$$|2[\mathbf{g}]_i[\mathbf{g}]_j\Sigma_{ij}| = 2|\,\mathrm{corr}(z_i, z_j)|\,\sqrt{V_iV_j} \le |\,\mathrm{corr}(z_i, z_j)|\,(V_i + V_j). \tag{20.23}$$

A cross term is at most the correlation times the two diagonal terms it joins. For an output of two entries, with $D = V_1 + V_2$ the diagonal sum and c the absolute correlation of the two entries, the total therefore lies between $(1 - c)\,D$ and $(1 + c)\,D$, and when the cross term is negative the diagonal terms together amount to as much as $1/(1 - c)$ of the total. A budget's shares, each term divided by $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$, need not lie between zero and one: a negative cross term pushes the diagonal shares above one.

An attribution per entry. A second way of writing the same total assigns each cross term half to each of its two entries:

$$\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g} = \sum_i [\mathbf{g}]_i\,(\boldsymbol{\Sigma}\mathbf{g})_i, \tag{20.24}$$

which follows from $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g} = \sum_i [\mathbf{g}]_i \sum_j \Sigma_{ij}[\mathbf{g}]_j$. Its terms also sum to the total and can also be negative: an entry whose correlation with the others works against its own diagonal term receives a negative share. The two forms answer different questions. Equation (20.20) shows each entry's own spread and the correction for the entries moving together; equation (20.24) assigns the whole variance to the entries.

For the scenario temperature, with the gradient (20.14), the budget has three terms, $V_1 = [\mathbf{g}]_1^2\Sigma_{11}$ for the conductance, $V_5 = \Sigma_{55}$ for the supply temperature and the cross term $2[\mathbf{g}]_1\Sigma_{15}$. The sign of the cross term is read off the objects. By the pattern (20.3) the hall-1 reading row is the only row whose line of $\mathbf{J}$ has entries in both the g_1 and the T_s columns, $-Q_1\mathrm{e}^{-g_1}/\sigma_T$ and $1/\sigma_T$, of opposite signs: it fixes the sum $T_s + Q_1\mathrm{e}^{-g_1}$, so that a warmer supply goes with a smaller rise and a larger conductance. That row alone puts a negative entry into $\boldsymbol{\Lambda}_{15}$, which the inverse carries into the positive covariance Σ_{15} that Section 20.2.2 reads. The scenario output takes the two entries with opposite signs, $[\mathbf{g}]_1 < 0$ and $[\mathbf{g}]_5 = +1$, so by equation (20.22) its cross term is negative. The physics behind it is the extrapolation itself. Writing the present-load temperature $T_1 = T_s + Q_1\mathrm{e}^{-g_1}$, a function of the state, and writing the scenario output (20.2) through it by the common factor e^{-g_1},

$$T_1^* = T_s + \frac{Q_1^*}{Q_1}\,(T_1 - T_s) = \frac{Q_1^*}{Q_1}\,T_1 - \Big(\frac{Q_1^*}{Q_1} - 1\Big)T_s\,: \tag{20.25}$$

held at the observed temperature of the hall, a warmer supply lowers the scenario temperature, because the rise that the reading attributed to the path is smaller, and it is the rise that doubles. In the coordinates (g_1, T_s) of the state that single negative effect of the supply appears split, as a positive diagonal term V_5 and a larger negative cross term, and the attribution (20.24) can give T_s a negative share.

Building blocks. The gradient (20.14); the entries Σ_{11}, Σ_{15} and Σ_{55} of the covariance (Section 20.2.1); the product $\boldsymbol{\Sigma}\mathbf{g}$ for the attribution.

The variance budget of the scenario temperature. Table 20.3 itemizes the variance of T_1^* by equation (20.20): each row is a term, with its expression, its value in K^2 and its share of the total, and the last column gives the per-entry attribution of equation (20.24). Figure 20.2 draws the same terms as signed bars: the horizontal axis lists the three terms and the total, the vertical axis is variance in K^2, bars above the zero line add variance and the bar below it removes variance.

The conductance of hall 1 sets the band: its diagonal term is 118 % of the variance, the supply temperature's 7 %, and the cross term −25 %. With $\mathrm{corr}(z_1, z_5) = 0.434$ the bound (20.23) lets the two diagonal shares reach $1/(1 - 0.434) = 177$ % of the total. The attribution (20.24) gives 105.5 % to the conductance and −5.5 % to the supply temperature. At $Q_1 = 30.69\,\mathrm{kW}$ the re-expression (20.25) reads $T_1^* = 1.955\,T_1 - 0.955\,T_s$. Without the cross term the spread would be $\sqrt{1.3594 + 0.0784} = 1.199\,\mathrm{K}$, a band 12 % too wide. The budget's terms have no counterpart

Table 20.3: The variance budget of the scenario output T_1^* at 60 kW, equation (20.20), and the per-entry attribution of equation (20.24), which splits the cross term equally between g_1 and T_s.

Term	Expression	Value [K^2]	Share	Attribution
Conductance of hall 1	$V_1 = [\mathbf{g}]_1^2\,\Sigma_{11}$	1.3594	117.8 %	g_1: 105.5 %
Supply temperature	$V_5 = [\mathbf{g}]_5^2\,\Sigma_{55}$	0.0784	6.8 %	T_s: −5.5 %
Cross term	$2[\mathbf{g}]_1[\mathbf{g}]_5\,\Sigma_{15}$	−0.2836	−24.6 %	(split)
Total	$\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$	1.1543	100 %	100 %

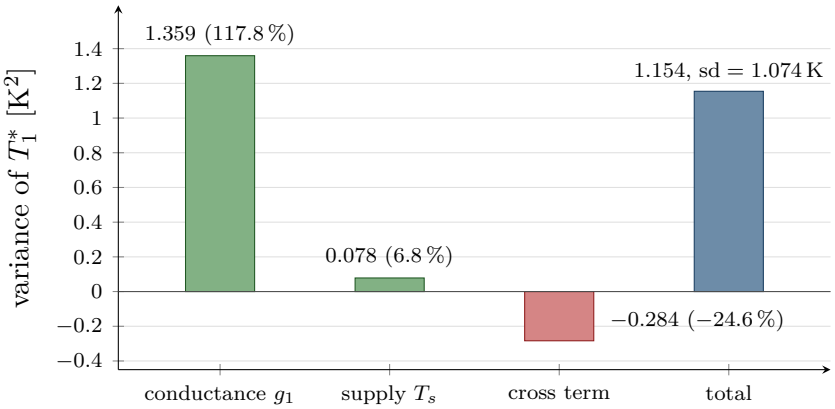


Figure 20.2: The variance budget of hall 1's temperature at 60 kW, equation (20.20), as signed bars: the diagonal terms $[\mathbf{g}]_1^2\Sigma_{11}$ and $[\mathbf{g}]_5^2\Sigma_{55}$ (green), the cross term $2[\mathbf{g}]_1[\mathbf{g}]_5\Sigma_{15}$ (red, negative) and the total $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$ (blue), with $[\mathbf{g}]_1 = -10.928$ K, $[\mathbf{g}]_5 = 1$ and the covariance block (20.4). Labels: value in K^2 and share of the total.

under the full posterior, whose variance is not a quadratic form; the total they split, 1.074 K, stands beside the full posterior's 1.071 K of the first question.

20.3.4 The conductance ratio, the total load and the conductance

Engineering question. What are the conductance ratio, the total load and hall 1's conductance, with bands? The ratio compares the cooling paths of the two halls, the total load is what the chiller must remove, and the conductance of hall 1 is the path the scenario depends on.

General form. The estimation form again, equations (2.10) and (2.11), for three derived quantities, each with its location and its central interval.

Specialized form. The questions are equations (2.10) and (2.11) with the state $\mathbf{z}$ the reduced state (20.1), the derived quantity $f(\mathbf{z})$ in turn G_1/G_2, $Q_1 + Q_2$ and G_1, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 20.2.2. One choice enters here: the ratio and the conductance are asked on their own scale, the one the operator reads, while the state carries them through their logarithms. This buys answers in the operator's units; it costs the symmetry of the delta band, which, centered on $f(\hat{\mathbf{z}})$, can reach where a positive quantity's posterior does not lie; the full posterior gives the band set beside it.

The three outputs and their gradients at $\hat{\mathbf{z}}$ are

$$\frac{G_1}{G_2} = \mathrm{e}^{g_1 - g_2}, \quad \mathbf{g} = \frac{\hat{G}_1}{\hat{G}_2}\begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}; \qquad Q_1 + Q_2, \quad \mathbf{g} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}; \qquad G_1 = \mathrm{e}^{g_1}, \quad \mathbf{g} = \hat{G}_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{20.26}$$

with $\hat{G}_i = \mathrm{e}^{\hat{g}_i}$, so that by equation (20.20) the ratio has the variance $(\hat{G}_1/\hat{G}_2)^2(\Sigma_{11} + \Sigma_{22} - 2\Sigma_{12})$, the total load $\Sigma_{33} + \Sigma_{44} + 2\Sigma_{34}$ and the conductance $\hat{G}_1^2\Sigma_{11}$. By the pattern (20.3) the columns of g_1 and g_2 share no row: the two log-conductances meet only through T_s and the loads, so Σ_{12} is small and the ratio's variance is nearly the sum of the two. The total load is read by one row only, the heat meter, which reads exactly the combination asked about; by the scalar case of the rank-one update (4.63), with the Jacobian held at $\hat{\mathbf{z}}$, its precision adds to what the other rows say of the sum (the load meter on Q_1, the hall-2 reading through the conductance prior on Q_2), so the sum is known better than the heat meter alone reads it, and the two loads, fixed in sum better than singly, are negatively correlated. The total load is linear in the state, so the delta method is exact for it under the Laplace posterior. The other two are monotone functions of the linear combinations $g_1 - g_2$ and g_1, which are normal under the Laplace posterior; their Laplace pushforward is therefore log-normal in closed form, the exponential of the normal interval of the combination. The value beside each delta band comes from the full posterior, the weighted quantiles of the reweighted draws (20.11).

Building blocks. The gradients (20.26); the entries of $\boldsymbol{\Sigma}$ in the rows and columns of (g_1, g_2) and of (Q_1, Q_2); and, for the bands beside them, the reweighted draws of the first question and, for the ratio and G_1, the exponential of the normal interval of $g_1 - g_2$ and of g_1.

The conductance ratio. $G_1/G_2 = \mathrm{e}^{0.30304} = 1.354$ at $\hat{\mathbf{z}}$, with the gradient $1.354 \times (1, -1, 0, 0, 0)^\mathsf{T}$ of equation (20.26), so

$$\begin{aligned} \mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g} &= 1.354^2\,(\Sigma_{11} + \Sigma_{22} - 2\Sigma_{12}) \\ &= 1.354^2 \times (0.0113824 + 0.0187648 + 0.0008347) = 1.354^2 \times 0.0309819, \end{aligned}$$

a spread of $1.354 \times 0.17602 = 0.238$ and a delta interval $[0.887,\ 1.821]$; the covariance term $-2\Sigma_{12} = 0.000835$ is 3 % of the sum. The value beside it comes from the full posterior, the interval $[0.962,\ 1.929]$: the delta band sits 0.075 too low at its lower end and 0.108 too low at its upper, because a symmetric band on a ratio ignores the skew of the exponential. The log-normal pushforward, $\exp(0.30304 \pm 1.960 \times 0.17602)$, gives $[0.959,\ 1.912]$, within 0.02 of the full posterior.

The total load. $Q_1 + Q_2$ has the constant gradient $(0, 0, 1, 1, 0)^\mathsf{T}$:

$$\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g} = \Sigma_{33} + \Sigma_{44} + 2\Sigma_{34} = 2.1085 + 5.5048 - 3.8980 = 3.7153\,\mathrm{kW}^2,$$

and $Q_1 + Q_2 = 52.92 \pm 1.93\,\mathrm{kW}$, interval $[49.14,\ 56.70]\,\mathrm{kW}$, below the heat meter's 2 kW. With the Jacobian held at $\hat{\mathbf{z}}$ and the heat-meter row removed the sum would have a spread of 7.22 kW, and $1/7.22^2 + 1/2^2 = 0.0192 + 0.2500 = 0.2692\,\mathrm{kW}^{-2}$, the precision of a spread of 1.93 kW. The value beside it comes from the full posterior, $[49.18,\ 56.69]\,\mathrm{kW}$, within 0.04 kW at either end: the sum is linear in the state, and the delta method is exact for it under the Laplace posterior.

The conductance of hall 1. $G_1 = \mathrm{e}^{g_1}$ has the gradient $(5.490, 0, 0, 0, 0)^{\mathsf{T}}$ and the delta band 5.490 ± 0.586, interval $[4.342,\ 6.638]$ kW/K. The value beside it comes from the full posterior, median 5.546 and interval $[4.544,\ 6.938]$ kW/K: the delta band lies 0.20 and 0.30 kW/K to the left at its two ends, because the posterior of g_1 is itself skewed toward large conductances and the exponential adds its own; the log-normal pushforward, $[4.454,\ 6.767]$ kW/K, removes about half of the gap.

20.4 The checks (step 5)

The delta answers rest on two approximations, the Gaussian posterior of the reduced state and the tangent plane of each output, and on two inputs, the limit of exact balances and the priors. Each is set beside a calculation that does without it.

20.4.1 The Laplace pushforward by draws and by one integral

$N = 200\,000$ draws of $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$, from the PCG64 pseudo-random generator seeded with 17 (every set of draws in this chapter comes from that generator and seed), were pushed through equation (20.2). Their sample mean is 23.103°C, their standard deviation 1.085 K, their median 23.044°C and their central 95 % band $[21.16,\ 25.41]$°C; their skewness is 0.33. The fraction below 25°C is 0.952, with standard error $\sqrt{0.952 \times 0.048 / 200\,000} = 0.0005$, against the delta method's 0.966.

The second-order correction. With the one nonzero second derivative, 10.928 K, equation (20.6) gives

$$\tfrac{1}{2}\,\mathrm{tr}\big(\boldsymbol{\Sigma}\,\nabla^2 T_1^*\big) = \tfrac{1}{2} \times 0.011382 \times 10.928 = 0.0622\,\mathrm{K},$$

while by equation (20.9) the mean of the Laplace pushforward exceeds $T_1^*(\hat{\mathbf{z}})$ by the rise times $\mathrm{e}^{\Sigma_{11}/2} - 1$, which is $10.928 \times 0.005707 = 0.0624$ K, a mean of $23.0399 + 0.0624 = 23.102$°C; the draws give 23.103°C. The shift is 6 % of the spread.

The one-dimensional integral. Equation (20.8), with the conditional slope $\Sigma_{15}/\Sigma_{11} = 1.140$ K and the conditional spread $\sqrt{\Sigma_{5|1}} = \sqrt{0.078433 - 0.012974^2/0.011382} = 0.2523$ K from equation (20.7), evaluated by adaptive quadrature, gives 0.952; its quantiles are 21.15, 23.04 and 25.39°C at 2.5, 50 and 97.5 %, its standard deviation 1.084 K and its skewness 0.33. The draws reproduce it within their standard error. Under the Laplace posterior, then, the linearization understates the probability of exceeding the limit: 0.034 by the tangent plane against 0.048 by the Laplace pushforward. The cause is the convexity of e^{-g_1}: a normal g_1 pushed through it gives a right-skewed temperature, whose upper tail is heavier than the normal's.

20.4.2 The full posterior by reweighting

The Laplace pushforward tests the tangent plane, and only the tangent plane: it assumes the Gaussian posterior that the delta method also assumes. Chapter 19, Section 19.4.1, computed the exact posterior of this plant by importance sampling from a multivariate t distribution and found it close to the Laplace Gaussian in its spreads but not in its shape along g_1: the exact mean of g_1 lies 0.014 above the MAP value, and the upper tail of the Laplace band of g_1 holds 4.0 % of the exact posterior instead of 2.5 %. The same 200 000 draws used above, reweighted by equation (20.11) with the objective evaluated at each of them, carry that shape to the scenario output. The log weights $\ln \zeta_j$ range from -5.9 to 3.4, and the effective sample size (20.12) is 182 933, so the weights are close to equal and the estimates precise. Under the full posterior the

scenario temperature has mean 22.98°C, standard deviation 1.071 K, median 22.95°C, central 95 % band [20.97, 25.16]°C and skewness 0.17, and

$$\Pr(T_1^* < 25°\mathrm{C} \mid \mathbf{y}) = 0.966 \pm 0.0004.$$

A second run, with 200 000 draws from a multivariate Student t with five degrees of freedom and 1.5 times the Laplace spread ($N_{\mathrm{eff}} = 85\,968$), gives 0.966 again.

Table 20.4 sets the three routes side by side, the second in its two computations. The columns are the delta method, the Laplace pushforward by the integral and by the draws, and the full posterior; the rows are the summaries of T_1^*, followed by the probability of staying below limits from 23.5 to 27.5°C, each labeled with its standardized distance $(T_{\max} - 23.040)/1.0744$. One row of the arithmetic, for 24°C: $(24 - 23.040)/1.0744 = 0.894$ and $\Phi(0.894) = 0.814$. Figure 20.3 draws the densities. The horizontal axis is the scenario temperature and the vertical axis its density in K^{-1}; the solid blue curve is the delta-method normal, the dashed red curve the density of the Laplace pushforward, the derivative of equation (20.8) with respect to $T_{\max}$, and the green dotted curve the full posterior, a kernel estimate from the reweighted draws. Panel (b) enlarges the upper tail; the shaded area under the green curve to the right of the limit is the probability of exceeding it under the full posterior.

Table 20.4: The check of the scenario output T_1^* at 60 kW by three routes: the delta method, the Laplace pushforward (one-dimensional integral (20.8) and 200 000 draws), and the full posterior (the same draws reweighted, $N_{\mathrm{eff}} = 182\,933$). In brackets after each limit, its standardized distance from the delta center. Standard errors of the draw probabilities: at most 0.0011 in either column.

Summary	Delta method	Laplace pushforward, integral	Laplace pushforward, draws	Full posterior
mean [°C]	23.040	23.102	23.103	22.981
standard deviation [K]	1.074	1.084	1.085	1.071
median [°C]	23.040	23.043	23.044	22.951
95 % band [°C]	[20.93, 25.15]	[21.15, 25.39]	[21.16, 25.41]	[20.97, 25.16]
skewness [–]	0	0.33	0.33	0.17
$\Pr(T_1^* < 23.5°\mathrm{C})$ [0.43]	0.666	0.661	0.660	0.693
$\Pr(T_1^* < 24.0°\mathrm{C})$ [0.89]	0.814	0.802	0.802	0.830
$\Pr(T_1^* < 24.5°\mathrm{C})$ [1.36]	0.913	0.897	0.896	0.918
$\Pr(T_1^* < 25.0°\mathrm{C})$ [1.82]	0.966	0.952	0.952	0.966
$\Pr(T_1^* < 25.5°\mathrm{C})$ [2.29]	0.989	0.979	0.979	0.987
$\Pr(T_1^* < 26.0°\mathrm{C})$ [2.76]	0.997	0.992	0.992	0.996
$\Pr(T_1^* < 27.5°\mathrm{C})$ [4.15]	1.0000	0.9997	0.9997	0.9999

Why the two corrections cancel. In panel (a) the red curve leans to the right of the blue one and the green curve does not: the Laplace pushforward is skewed, and the full posterior is nearly symmetric. Panel (b) shows the consequence in the tail, where the red curve lies above the other two. The reason is the hall-1 reading row. For a given load and supply temperature it is linear in the thermal resistance e^{-g_1}, not in its logarithm g_1, and the readings dominate the conductance prior (Chapter 19 found the prior spread 2.8 times the posterior spread), so the full posterior is nearly symmetric in the resistance. In the logarithm it is then skewed to the right, the long side of large conductances that Chapter 19 found: under the full posterior g_1 has skewness 0.17 and mean 1.716, above its MAP value of 1.703, and the resistance has skewness 0.16. The Laplace approximation puts a symmetric Gaussian on g_1, which makes the resistance log-normal with skewness $(\mathrm{e}^{\Sigma_{11}} + 2)\sqrt{\mathrm{e}^{\Sigma_{11}} - 1} = 0.322$. The scenario output is linear

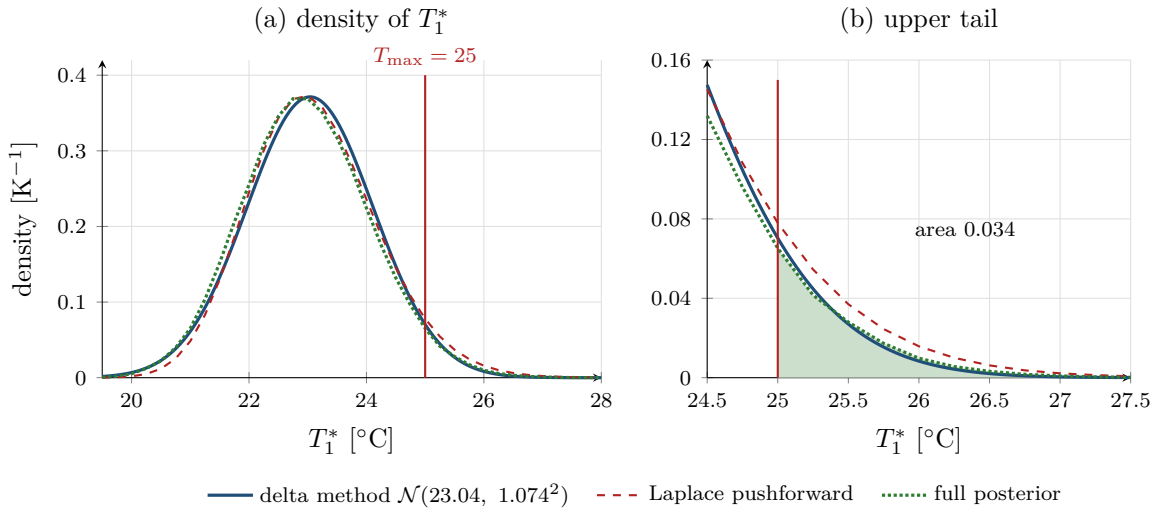


Figure 20.3: Density of hall 1's temperature at 60 kW by three routes. Blue: the delta-method normal $\mathcal{N}(23.040, 1.0744^2)$, equation (20.15). Red dashed: the Laplace pushforward, the derivative of equation (20.8) in $T_{\max}$, evaluated on a grid of 0.25 K. Green dotted: the full posterior, a Gaussian kernel estimate (bandwidth 0.12 K) of the 200 000 draws reweighted by equation (20.11). (b) The upper tail enlarged; shaded, $\Pr(T_1^* > 25°\text{C} \mid \mathbf{y}) = 0.034$ under the full posterior, against 0.048 under the red curve.

in the resistance, so it inherits the posterior's shape in the resistance: nearly symmetric under the full posterior, skewed under the Laplace one. The delta method, which linearizes the output at $\hat{\mathbf{z}}$, does not depend to first order on whether the posterior is approximated in g_1 or in the resistance, and it ends close to the full posterior, though not on it (Section 20.4.3). The Laplace pushforward corrects the output for its curvature in g_1 and keeps a Gaussian whose error in g_1 has the opposite sign; one correction made without the other moves the answer away from the truth. The second-order mean correction does the same: +0.062 K under the Laplace posterior, while the full posterior's mean lies 0.059 K below $T_1^*(\hat{\mathbf{z}})$.

> **Principle 20.1.** *A correction for the curvature of an output is worth making only together with the correction for the shape of the posterior. Made alone, by the Laplace pushforward, it can move the answer away from the full posterior. The check that decides is the one against the full posterior, which reweighting provides at the cost of one evaluation of the objective per draw.*

20.4.3 The error of the delta method across limits

The lower rows of Table 20.4 give the error of the delta method against the full posterior at each limit. It is −0.027 at 23.5°C, 0.43 standard deviations above the delta center, and −0.026 at 0.5 standard deviations; −0.016 at 24°C (0.89); −0.005 at 24.5°C (1.36); it changes sign near 1.8 standard deviations, where at the question's 25°C (1.82) the two agree within the draws' standard error of 0.0004; beyond, the delta method is high, by 0.0016 at 25.5°C (2.29, standard error 0.0002) and 0.0011 at 26°C (2.76, standard error 0.0001). Two features of the full posterior produce this pattern. Its center lies below the delta center: the delta center, $T_1^*(\hat{\mathbf{z}})$, is 0.05 standard deviations above the full posterior's mean, so near the center the delta method puts too little probability below a limit. Its residual skewness of 0.17 thickens its upper tail, so far from the center the delta method puts too much. The two effects cancel near 1.8 standard deviations, and the question's limit happens to lie there. The agreement of equation (20.19)

with the full posterior is therefore a property of this limit, not of the delta method at 60 kW; at the limits nearer the center an operator might also ask about, the delta method understates the probability of staying below by up to 0.03, and the Laplace pushforward understates it at every limit, by up to 0.032.

The error depends on the standardized distance of the limit, not on the load. At 90 kW, where the band is 1.53 times wider, the same comparison gives −0.017 at 0.9 standard deviations, against −0.016 at 60 kW, and at 1.824 standard deviations it agrees within 0.0005, close to the draws' standard error of 0.0004, as at 60 kW: the scenario output is a fixed temperature plus the load times the resistance, the load scales the resistance's contribution and its spread together, and the shape of the standardized posterior hardly changes (Exercise 20.5).

20.4.4 The other outputs against the full posterior

Table 20.5 sets the three further outputs beside their other routes. Each row is an output, with its value at $\hat{\mathbf{z}}$, its delta spread and interval from Section 20.3.4, its Laplace pushforward interval, in closed form because each output is a monotone function of a normal combination of entries, and its interval under the full posterior from the same 200 000 reweighted draws. Figure 20.4 draws the same intervals in three panels, one per output, each with its own horizontal axis in the output's units; within a panel, the three horizontal bars are the delta interval (top), the Laplace pushforward (middle) and the full-posterior interval (bottom), and each dot marks the center of its bar, $f(\hat{\mathbf{z}})$ or the median.

Table 20.5: Three further derived quantities: value at the MAP state, delta-method spread and 95 % interval, the Laplace pushforward interval (log-normal for the ratio and for G_1; equal to the delta interval for the linear total), and the interval under the full posterior from 200 000 reweighted draws.

Output	$f(\hat{\mathbf{z}})$	$\sqrt{\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}}$	Delta interval	Laplace pushforward	Full posterior
G_1/G_2 [–]	1.354	0.238	[0.887, 1.821]	[0.959, 1.912]	[0.962, 1.929]
$Q_1 + Q_2$ [kW]	52.92	1.93	[49.14, 56.70]	[49.14, 56.70]	[49.18, 56.69]
G_1 [kW/K]	5.490	0.586	[4.342, 6.638]	[4.454, 6.767]	[4.544, 6.938]

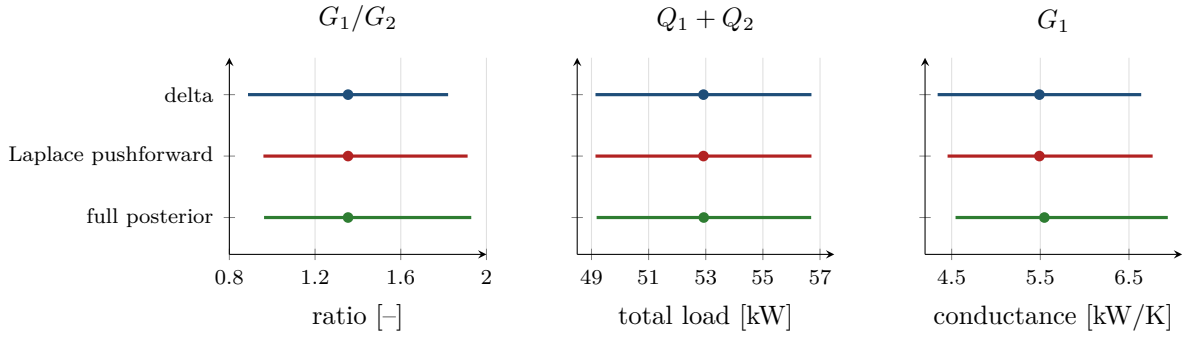


Figure 20.4: Central 95 % intervals of three derived quantities (Table 20.5). Blue, top: the delta method, $f(\hat{\mathbf{z}}) \pm 1.960\sqrt{\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}}$, dot at $f(\hat{\mathbf{z}})$. Red, middle: the Laplace pushforward, the exponential of the normal interval of $g_1 - g_2$ or g_1 (for the linear total, the delta interval), dot at the median. Green, bottom: the full posterior from 200 000 draws reweighted by equation (20.11), dot at the weighted median.

The total load is linear and its three bars coincide: the delta method and the Laplace pushforward are the same for it, and the full posterior agrees within 0.04 kW. For the two

exponential outputs the delta interval is symmetric about $f(\hat{\mathbf{z}})$ and lies to the left of the other two; for the ratio, whose log-scale spread of 0.176 is the larger, the delta interval reaches down to 0.89 where the Laplace pushforward and the full posterior stop at 0.96. Here, unlike the scenario temperature, the Laplace pushforward is the closer of the two Gaussian routes to the full posterior. The full posterior of G_1 lies to the right of the Laplace one, with median 5.55 against 5.49 kW/K: the posterior of g_1 is itself skewed to the right, the effect that Section 20.4.2 traced to the hall-1 reading row.

20.4.5 The limit of exact balances

The reduced state rests on the limit $\sigma_p \to 0$ of the physics rows, equation (17.19). Table 20.6 checks it on the scenario output: each of the first three columns is the full state of Chapter 17, with its eleven rows and tolerances σ_p in kelvin and $\sigma_{p,P}$ in kilowatts of the same number, refitted and with its covariance inverted at its own MAP state; the last column is the reduced state. The rows are the spread of g_1, its covariance with T_s, the scenario temperature at the MAP state, its delta spread and its threshold probability.

Table 20.6: The scenario output on the full state with physics rows of tolerance σ_p, against the reduced state of Chapter 17; delta method, gradient $(-60\,\mathrm{e}^{-\hat{g}_1},\ 0,\ 0,\ 0,\ 1)$ on the free entries.

Quantity	$\sigma_p = 0.1$	0.01	0.001	reduced state
$\sqrt{\Sigma_{11}}$ of g_1 [–]	0.1078	0.1067	0.1067	0.1067
Σ_{15} [K]	0.012930	0.012973	0.012974	0.012974
$T_1^*(\hat{\mathbf{z}})$ [°C]	23.043	23.040	23.040	23.040
$\sqrt{\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}}$ [K]	1.0883	1.0745	1.0744	1.0744
$\Pr(T_1^* < 25°\mathrm{C} \mid \mathbf{y})$ [–]	0.9639	0.9659	0.9660	0.9660

By $\sigma_p = 0.01$ K, the tolerance of Chapter 17, the spread of the scenario temperature has stopped moving in its third decimal, and at 0.001 K every printed number equals the reduced state's. A balance allowed to fail by a tenth of a kelvin widens the band by 1.3 % and lowers the threshold probability by 0.002, less than the errors of the approximations above.

20.4.6 Sensitivity to the priors and to the meters

Table 20.7 refits the reduced state under five alternatives to the handbook priors and repeats the delta answers for the hall at 60 kW. Each row is a prior: the first the one used; the next four move the conductance prior, wider or narrower than the handbook's 35 % or centered on 4 or 6 kW/K instead of 5; the last widens the set-point prior. The columns give the fitted conductance, the scenario temperature, its spread, the threshold probability and the conductance term's share of the budget.

The threshold probability moves between 0.945 and 0.978 over the conductance priors, an effect as large as the delta method's error nearer the center, and the set-point prior moves nothing. The verdict on the budget holds under every prior: the conductance term carries 116 to 119 % of the variance. A narrower conductance prior pulls the fitted conductance toward the handbook's 5 kW/K, raising the scenario temperature by 0.31 K while it narrows the band, so that the probability falls; the prior's center moves the answer more than its width.

The meters weigh more than any prior. With the Jacobian held at $\hat{\mathbf{z}}$, as in Chapter 19's removal of meters, the spread of T_1^* would be 2.13 K without the electrical meter on hall 1's load, and the probability of staying below 25°C 0.821; without that meter and the chiller's heat meter it would be 3.29 K, and the probability 0.724. The meters fix the load at which hall 1's temperature rise was observed, and through it the conductance; without them the conductance falls back toward its prior, and the scenario inherits the prior's spread multiplied by a lever

Table 20.7: Sensitivity of the hall at 60 kW to the priors of Table 17.2; reduced state refitted under each, delta method. The share is $V_1/\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$.

Prior	$\hat{G}_1$ [kW/K]	$T_1^*(\hat{\mathbf{z}})$ [°C]	sd [K]	Pr(<25°C) [–]	share [–]
handbook: $g_i \sim \mathcal{N}(\ln 5,\ 0.3^2)$, $T_s \sim \mathcal{N}(12,\ 1)$	5.490	23.040	1.074	0.966	117.8 %
conductance prior widened, 0.6	5.549	22.910	1.116	0.969	118.3 %
conductance prior narrowed, 0.15	5.349	23.354	0.944	0.959	115.7 %
conductance prior centered on 4 kW/K	5.349	23.268	1.082	0.945	117.7 %
conductance prior centered on 6 kW/K	5.618	22.843	1.068	0.978	117.9 %
set-point prior widened, 2 K	5.497	23.034	1.075	0.966	118.8 %

of 10.9 K. At the 27.5°C of Chapter 2 the same two cases give 0.982 and 0.912, and the full record 0.99998; that chapter gave 0.717 at 27.5°C for the single-path hall, equation (2.37), with a narrower band of 0.46 K but a mean only 0.27 K below the limit, because that hall rejected its heat to an ambient of 18°C; hall 1 rejects it to a supply at 12°C, and its mean lies 4.5 K, or 4.2 standard deviations, below the same limit.

20.5 The report (step 6)

The two kelvin of margin are not a certainty. Hall 1 at 60 kW would run at 23.0°C with a spread of 1.07 K, so the limit lies only 1.8 standard deviations above it: the probability of exceeding 25°C is 0.034, about one chance in thirty, and the upper end of the 95 % band lies above the limit. The delta method's 0.966 agrees with the full posterior at this limit only because two errors cancel there, while the Laplace pushforward, the more careful-looking route, gives 0.952. Each answer is reported from the posterior of the state by the delta method, with the full posterior's value beside it and the gap named where the Gaussian band departs from it; the answers are given in the units in which the questions were asked.

20.5.1 The answers and their bands

Table 20.8 is the report for the four questions of Table 20.1, in their order: one row per question, with the answer from the posterior of the state by the delta method in the second column, the full posterior's value beside it in the third, and in the fourth the gap and what the checks of Section 20.4 add.

The report line. At a load of 60 kW, with the conductance and the supply temperature it has now, hall 1 is expected to run at 23.0°C, with a 95 % band from 20.9 to 25.2°C (21.0 to 25.2°C under the full posterior); the probability that it stays below 25°C is 0.97 on both routes, and below 27.5°C it is above 0.9999. The band is set by the conductance of hall 1's heat-removal path, inferred from one temperature sensor at about half the scenario load; the supply temperature contributes little, and less than its own spread suggests, because the reading that fixes the conductance also ties it to the supply. Every band and probability stands beside its full-posterior value, and the gap is largest for the conductance ratio and the conductance of hall 1, whose delta bands lie to the left of the full ones. The assessment assumes that the path's conductance and the supply temperature do not change with the load.

Table 20.8: The answers to the questions asked of the fitted two-hall plant. Bands are central 95 % posterior intervals; "full" denotes the full posterior by reweighted draws, "delta" the delta method, whose location is the plug-in value $f(\hat{\mathbf{z}})$.

Question	Answer (state route)	Exact (full posterior)	Gap and checks
What will hall 1's temperature be at 60 kW, and how well is that known?	plug-in $T_1^*(\hat{\mathbf{z}}) = 23.04°\mathrm{C}$, spread 1.074 K; delta [20.93, 25.15]	mean 22.98°C, median 22.95°C, spread 1.07 K; [20.97, 25.16]	delta center 0.06 K high; band ends 0.03 and 0.01 K below the full ones
Will the hall stay under the limit at 60 kW?	$\Pr(T_1^* < 25°\mathrm{C} \mid \mathbf{y}) =$ 0.966; exceedance 0.034	0.966	within the standard error, 0.0004; Laplace pushforward 0.952; 0.945 to 0.978 over the conductance priors; at the other limits of Table 20.4 the delta method is low by up to 0.027 nearer the center and high by up to 0.0016 beyond; at 27.5°C, full 0.9999
Which entry's uncertainty drives the answer?	hall 1's conductance: diagonal term 118 %, supply 7 %, cross term −25 %	– (no term-by-term counterpart)	conductance term 116 to 119 % under every prior
What are the conductance ratio, the total load and hall 1's conductance, with bands?	plug-in $f(\hat{\mathbf{z}})$ 1.354, 52.92 kW, 5.490 kW/K; delta [0.89, 1.82], [49.14, 56.70] kW, [4.34, 6.64] kW/K	medians 1.354, 52.93 kW, 5.55 kW/K; [0.96, 1.93], [49.2, 56.7] kW, [4.54, 6.94] kW/K	the delta bands of the ratio and of G_1 are shifted left and fail at both ends, more at the upper (the skew of the exponential); the delta band of the total is within 0.04 kW

20.6 The next measurement (step 7)

The report leaves the margin at 60 kW at one chance in thirty of exceedance, with a band set by hall 1's conductance. The candidate measurements are rows that could be added to the plant before the load is moved: a second hall-1 temperature sensor, a better supply sensor, or a load meter on hall 2.

20.6.1 The preposterior of a candidate row

A design question about a measurement has the preposterior form of Chapter 2, equation (2.16), and for a Gaussian posterior and a linear Gaussian row it needs only the preposterior covariance (2.17). A candidate measurement a would add one row to the objective,

$$r_a(\mathbf{z}) = \frac{\hat{y}_a(\mathbf{z}) - \tilde{y}_a}{\sigma_a},$$

with $\hat{y}_a(\mathbf{z})$ the reading the model predicts at the state $\mathbf{z}$, $\tilde{y}_a$ the value the instrument would return and σ_a its standard deviation. Let $\mathbf{j}_a = \nabla r_a(\hat{\mathbf{z}})$ be the column gradient of the normalized row at the MAP state, so that $\mathbf{j}_a^{\mathsf{T}}$ is the line the measurement would add to $\mathbf{J}$ and $\mathbf{J}_a = \mathbf{j}_a^{\mathsf{T}}$ in the notation of equation (2.17). Its term $\frac{1}{2}r_a^2$ adds to the Gauss–Newton curvature the outer product $\mathbf{j}_a\mathbf{j}_a^{\mathsf{T}}$, which does not depend on $\tilde{y}_a$. The information matrix after the measurement is therefore known before it is taken,

$$\mathbf{\Lambda} + \mathbf{\Lambda}_a, \qquad \mathbf{\Lambda}_a = \mathbf{j}_a\mathbf{j}_a^{\mathsf{T}}, \qquad \mathbf{\Sigma}_a = \left(\mathbf{\Sigma}^{-1} + \mathbf{\Lambda}_a\right)^{-1}, \tag{20.27}$$

the preposterior covariance of Chapter 2, used in Chapter 8 on the logit of an availability, equation (8.59), where $\mathbf{\Lambda}_a$ was diagonal. Here it is a full outer product, and the inverse has a closed form.

The matrix-inversion lemma. For the covariance $\mathbf{\Sigma}$ and the candidate's gradient $\mathbf{j}_a$, the lemma in its rank-one form states

$$\mathbf{\Sigma}_a = \left(\mathbf{\Sigma}^{-1} + \mathbf{j}_a\mathbf{j}_a^{\mathsf{T}}\right)^{-1} = \mathbf{\Sigma} - \frac{\mathbf{\Sigma}\mathbf{j}_a\,\mathbf{j}_a^{\mathsf{T}}\mathbf{\Sigma}}{1 + \mathbf{j}_a^{\mathsf{T}}\mathbf{\Sigma}\mathbf{j}_a}. \tag{20.28}$$

The denominator $s_a = 1 + \mathbf{j}_a^{\mathsf{T}}\mathbf{\Sigma}\mathbf{j}_a$ is a scalar, and it is positive because $\mathbf{\Sigma}$ is positive definite. To verify the lemma, multiply the right-hand side by $\mathbf{\Sigma}^{-1} + \mathbf{j}_a\mathbf{j}_a^{\mathsf{T}}$:

$$\begin{aligned}\left(\mathbf{\Sigma}^{-1} + \mathbf{j}_a\mathbf{j}_a^{\mathsf{T}}\right)\left(\mathbf{\Sigma} - \tfrac{1}{s_a}\mathbf{\Sigma}\mathbf{j}_a\mathbf{j}_a^{\mathsf{T}}\mathbf{\Sigma}\right) &= \mathbf{I} + \mathbf{j}_a\mathbf{j}_a^{\mathsf{T}}\mathbf{\Sigma} - \tfrac{1}{s_a}\mathbf{j}_a\mathbf{j}_a^{\mathsf{T}}\mathbf{\Sigma} - \tfrac{1}{s_a}\mathbf{j}_a\left(\mathbf{j}_a^{\mathsf{T}}\mathbf{\Sigma}\mathbf{j}_a\right)\mathbf{j}_a^{\mathsf{T}}\mathbf{\Sigma} \\ &= \mathbf{I} + \mathbf{j}_a\mathbf{j}_a^{\mathsf{T}}\mathbf{\Sigma}\left(1 - \frac{1 + \mathbf{j}_a^{\mathsf{T}}\mathbf{\Sigma}\mathbf{j}_a}{s_a}\right) = \mathbf{I},\end{aligned}$$

where the second line collects the three terms in $\mathbf{j}_a\mathbf{j}_a^{\mathsf{T}}\mathbf{\Sigma}$, the scalar $\mathbf{j}_a^{\mathsf{T}}\mathbf{\Sigma}\mathbf{j}_a$ commuting with everything. Equation (20.28) is the covariance half of the rank-one update of Chapter 4, equation (4.63), for any number of entries: with a reading of $\mathbf{h}^{\mathsf{T}}\mathbf{z}$ of standard deviation σ, $\mathbf{j}_a = \mathbf{h}/\sigma$, the subtracted term is $S\,\mathbf{K}\mathbf{K}^{\mathsf{T}}$ with $S = \mathbf{h}^{\mathsf{T}}\mathbf{\Sigma}\mathbf{h} + \sigma^2$ the predictive variance of the reading and $\mathbf{K} = \mathbf{\Sigma}\mathbf{h}/S$ the gain vector, one entry per entry of the state. It is a correction by the outer product of the single vector $\mathbf{\Sigma}\mathbf{j}_a$, at the cost of one product of $\mathbf{\Sigma}$ with a vector and no inversion.

The width removed from an output. Sandwiching equation (20.28) between $\mathbf{g}^\mathsf{T}$ and $\mathbf{g}$,

$$\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}_a\mathbf{g} = \mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g} - \frac{\left(\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{j}_a\right)^2}{1+\mathbf{j}_a^\mathsf{T}\boldsymbol{\Sigma}\mathbf{j}_a}, \tag{20.29}$$

so that $\mathrm{sd}(f \mid \mathbf{y}, a) = \sqrt{\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}_a\mathbf{g}}$. The two quantities in the fraction have a meaning. Under the linearized model the normalized reading would be $\mathbf{j}_a^\mathsf{T}\mathbf{d}$ plus a unit-variance error, so its predictive variance is $1+\mathbf{j}_a^\mathsf{T}\boldsymbol{\Sigma}\mathbf{j}_a$, the posterior predictive variance of Chapter 4, equation (4.61), and its covariance with the output is $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{j}_a$. The variance removed is the squared covariance divided by the predictive variance, so the fraction of the output's variance a candidate removes is the squared correlation between the output and the reading it would return. Three consequences follow. A measurement never increases the variance of any output, since the removed term is nonnegative. A measurement uncorrelated with the output under the posterior, $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{j}_a = 0$, removes nothing from it, however precise it is. And the same candidate ranks differently for different outputs, because the correlation depends on $\mathbf{g}$. The width removed, $1-\sqrt{\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}_a\mathbf{g}/\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}}$, is what step 7 ranks. The preposterior is exact for a linear-Gaussian model and uses the Jacobian at $\hat{\mathbf{z}}$ otherwise: a reading that moved $\hat{\mathbf{z}}$ far would leave the Jacobian at the old point stale, and the fit would be repeated with the reading included.

20.6.2 Three candidate rows

Three measurements could be added to the plant.

- *A second hall-1 temperature sensor*, of the same quality as the first, $\sigma = 0.5\,\mathrm{K}$, reading $T_s + Q_1\mathrm{e}^{-g_1}$. Its normalized row has the gradient of the hall-1 row of Chapter 17, equation (17.16), at $\hat{\mathbf{z}}$:

$$\mathbf{j}_a = \frac{1}{0.5}\left(-Q_1\mathrm{e}^{-g_1},\ 0,\ \mathrm{e}^{-g_1},\ 0,\ 1\right)^\mathsf{T} = (-11.179,\ 0,\ 0.364,\ 0,\ 2)^\mathsf{T},$$

  in units of the inverse of each entry's units; by the pattern (20.3) it is the one candidate whose line meets both columns of the scenario's gradient.
- *An added supply-temperature sensor* of $\sigma = 0.1\,\mathrm{K}$, three times better than the present one, installed beside it: $\mathbf{j}_a = (0,0,0,0,10)^\mathsf{T}\,\mathrm{K}^{-1}$.
- *An electrical meter on hall 2's load*, of $\sigma = 1.5\,\mathrm{kW}$, the twin of hall 1's, with the gradient $\mathbf{j}_a = (0,0,0,0.667,0)^\mathsf{T}\,\mathrm{kW}^{-1}$.

One candidate worked. For the second hall-1 sensor, the product of the covariance with the candidate's gradient is $\boldsymbol{\Sigma}\mathbf{j}_a = (-0.0781,\ -0.0004,\ 0.0632,\ -0.0624,\ 0.0129)^\mathsf{T}$, in the units of the entries; then $\mathbf{j}_a^\mathsf{T}\boldsymbol{\Sigma}\mathbf{j}_a = 0.9223$ and, with the gradient of T_1^*, $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{j}_a = (-10.928)(-0.0781) + (1)(0.0129) = 0.8668\,\mathrm{K}$. By equation (20.29),

$$\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}_a\mathbf{g} = 1.1543 - \frac{0.8668^2}{1+0.9223} = 1.1543 - \frac{0.7514}{1.9223} = 1.1543 - 0.3909 = 0.7634\,\mathrm{K}^2,$$

so $\mathrm{sd}(T_1^* \mid \mathbf{y}, a) = 0.874\,\mathrm{K}$: the band on T_1^* narrows from 1.074 to 0.874 K, a width removed of 18.7 %. The fraction of the variance removed, $0.3909/1.1543 = 0.339$, is the squared correlation between T_1^* and the second sensor's reading under the predictive distribution.

Table 20.9 lists the three candidates. The columns give the nonzero entries of each gradient, the two scalars of equation (20.29), the preposterior spread of T_1^* and the width removed from it, and the same for the load of hall 2, Q_2, whose gradient is $(0,0,0,1,0)^\mathsf{T}$. Figure 20.5 draws the spreads: the horizontal axis lists the present state and the three candidates, and for each the blue bar is the spread of T_1^* in K and the gray bar that of Q_2 in kW.

Table 20.9: The preposterior of three candidate rows, equations (20.27) and (20.29): the spread of hall 1's temperature at 60 kW and of hall 2's load after each, and the width each removes. $\mathbf{g}$ is the gradient of T_1^*.

Candidate	Nonzero entries of $\mathbf{j}_a$	$\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{j}_a$ [K]	$\mathbf{j}_a^\mathsf{T}\boldsymbol{\Sigma}\mathbf{j}_a$	sd(T_1^*) [K]	Width removed	sd(Q_2) [kW]	Width removed
present	–	–	–	1.074	–	2.346	–
second hall-1 sensor, 0.5 K	−11.18, 0.364, 2 on g_1, Q_1, T_s	0.867	0.922	0.874	18.7 %	2.346	0.0 %
added supply sensor, 0.1 K	10 on T_s	−0.633	7.843	1.053	2.0 %	2.345	0.1 %
hall-2 load meter, 1.5 kW	0.667 on Q_2	0.439	2.447	1.048	2.5 %	1.264	46.1 %

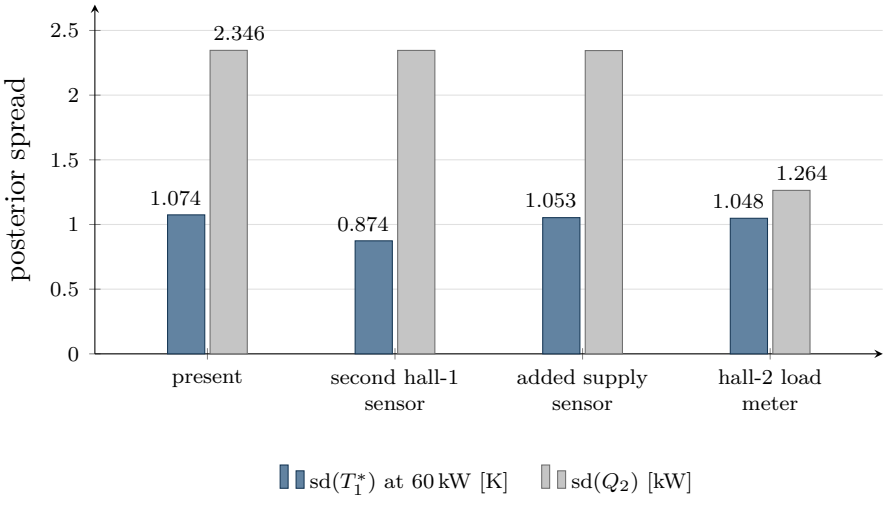


Figure 20.5: Preposterior spreads of hall 1's temperature at 60 kW (blue, K) and of hall 2's load (gray, kW) for the present state and three candidate rows, equation (20.29) (Table 20.9).

20.6.3 The ranking and its dependence on the output

For the hall at 60 kW the second hall-1 sensor ranks first by a wide margin: it removes 18.7 % of the width, against 2.0 and 2.5 % for the other two. It does so because it reads the combination that carries the budget's largest term: its correlation with T_1^* comes through g_1, whose diagonal term is 118 % of the variance. The supply sensor, five times more precise than the hall sensor in standard deviation and twenty-five times in precision, removes almost nothing, and no supply sensor could do much better: with T_s known exactly, the variance of T_1^* would be the conditional variance of the conductance term, $[\mathbf{g}]_1^2 \Sigma_{11}[1 - \mathrm{corr}(z_1, z_5)^2] = 1.3594 \times (1 - 0.434^2) = 1.1031\,\mathrm{K}^2$, a spread of 1.050 K. Replacing the present supply sensor by the better one, instead of adding it, gives 1.053 K, the same to three digits. The hall-2 load meter helps T_1^* only through the chiller meter, by sharpening Q_1 as the complement of Q_2 in the measured sum.

For hall 2's load the ranking is reversed. The hall-2 load meter takes $\mathrm{sd}(Q_2)$ from 2.35 to 1.26 kW, a width removed of 46 %, and the two temperature sensors remove nothing from it. The weakest direction that Chapter 19 found in the readings, the split of the total load between the halls, is the direction this meter reads. Which measurement is worth most depends on the output asked about; a ranking without an output is not defined. Further hall-1 sensors would continue to narrow the band on T_1^*, to 0.791 K with two added sensors (three in all) and 0.661 K with eight added (nine in all), more slowly with each, as the uncertainty of the load at which hall 1 was observed takes over.

20.6.4 The recommendation

If the margin at 60 kW is the decision, the next measurement is a second temperature sensor in hall 1, of the quality of the first, installed before the load is moved. It would narrow the band on T_1^* by about a fifth, from 1.074 to 0.874 K, and, if the center does not move, raise the probability of staying below 25°C to 0.988. Its reading enters as a ninth row of the reduced state, and the method returns to step 3 with it; the preposterior was computed with the Jacobian at the present $\hat{\mathbf{z}}$, and the refit says where the center actually goes. A better supply sensor is not worth installing for this question. If the split of the loads between the halls is the question instead, the hall-2 load meter is the measurement. The general treatment of the next measurement, by its expected information gain and by its value for a decision, is the subject of a later chapter on sensor placement and the value of information, and of the companion volume *Reliability Engineering*.

20.6.5 What the analysis showed

The first reading, 23.0°C at 60 kW and two kelvin to spare, said the load could be moved; the analysis sharply qualifies it. The two kelvin are 1.8 standard deviations, one chance in thirty of passing the limit, because the conductance of hall 1 was measured at half the scenario load and the doubled load multiplies its uncertainty by a lever of 10.9 K. The supply temperature, which the operator would naturally suspect, weighs less than its own spread: the hall-1 reading ties it to the conductance, and their cross term takes a quarter of the variance away. The careful-looking correction for the curvature of the output, made alone, moves the probability the wrong way, and only the check against the full posterior shows that the plain delta method is right at this limit, by a cancellation that holds near 25°C and not nearer the center. What the engineer carries back to the site is therefore specific: report the margin as a probability, not as two kelvin; buy a second temperature sensor in hall 1, not a better supply sensor; and confirm, before the move, that the coil's conductance and the chiller's supply temperature hold at the doubled load, which no reading of this record can show.

20.6.6 Limits of the model

The probability of equation (20.19) is conditional on the steady-state model and on the two assumptions of the scenario node.

The conductance at the scenario load. The conductance of hall 1's path was inferred at a load of 31 kW and is used at 60 kW. If the path's conductance depends on the flow of chilled water through it, and the flow on the load, the scenario node needs its own conductance tied to g_1 by a row, and its spread grows.

The supply temperature at the scenario load. The supply temperature is assumed held by the chiller at the doubled load; a chiller near its capacity lets T_s rise with the load, and a row for the chiller's capacity would be needed.

What the checks cannot see. Neither assumption can be tested by the readings of this plant, which were all taken at the present loads, about 31 kW in hall 1 and 22 kW in hall 2, and the checks of Section 20.4 cannot detect their failure: they test how well the formulas evaluate the posterior of the model, not whether the model holds at a load it has not seen. The steady balance also says nothing of the transient while the load is raised, during which the hall can pass its steady temperature; a storage term in the balance, as Chapter 1 wrote it, would be needed for that question.

The checks themselves hold only for the record and the question they were run on, as Chapter 19's conditions for trusting $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}^{-1}$ did, and four changes would spoil them.

A limit nearer the center. The delta method's error follows the standardized distance of the limit, not the load (Section 20.4.3); a question at a stricter limit is answered from the full posterior, not from the delta band.

Fewer rows behind the conductance. Without the load meters the posterior of g_1 is dominated by its prior, which is Gaussian in g_1; the hall-1 reading no longer makes the posterior nearly symmetric in the resistance, so the Laplace pushforward becomes the better route and the agreement that favors the delta method at 25°C is lost. With the band on T_1^* widened to 2.1 to 3.3 K (Section 20.4.6), the curvature of the output across it is larger, and the reweighting has to be repeated before either Gaussian route is reported.

A bounded quantity near its bound. A temperature limit close to the supply temperature, or an output such as a probability, has a posterior that a normal cannot represent; such outputs are asked about on a scale on which they are unbounded, as Chapter 8 did with the logit.

A reading in error. The checks test the approximations, not the readings. A biased hall-1 sensor would move the whole posterior and leave every check of Section 20.4 satisfied; the residuals and the χ^2 test of Chapter 18 are the guard, and Chapter 22 frees a bias term per sensor.

20.7 What to retain

Definition 20.1 (Propagation of the posterior to a derived quantity)**.** Let the posterior of the state be approximated by $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ and let $f(\mathbf{z})$ be a derived quantity with gradient $\mathbf{g}$ at $\hat{\mathbf{z}}$. To first order $f \mid \mathbf{y} \sim \mathcal{N}(f(\hat{\mathbf{z}}),\ \mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g})$, so that $\Pr(f < T_{\max} \mid \mathbf{y}) \approx \Phi\big((T_{\max} - f(\hat{\mathbf{z}}))/\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}\big)$ and the central 95 % interval is $f(\hat{\mathbf{z}}) \pm 1.960\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}$. The variance budget $\sum_i [\mathbf{g}]_i^2 \Sigma_{ii} + 2\sum_{i<j} [\mathbf{g}]_i [\mathbf{g}]_j \Sigma_{ij}$

has cross terms of either sign, each bounded by the correlation times the two diagonal terms it joins. The second-order mean is $f(\hat{\mathbf{z}}) + \frac{1}{2}\operatorname{tr}(\boldsymbol{\Sigma}\nabla^2 f)$ under the Gaussian posterior. Each answer is reported from $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$ by the delta method, and beside it stand the Laplace pushforward, draws of the Gaussian pushed through f, which tests the linearization, and the same draws reweighted with $\exp(-[\mathcal{L}(\mathbf{z}) - \mathcal{L}(\hat{\mathbf{z}})] + \frac{1}{2}\mathbf{d}^{\mathsf{T}}\boldsymbol{\Lambda}\mathbf{d})$, which give the full posterior. A candidate row with normalized gradient $\mathbf{j}_a$ gives $\boldsymbol{\Sigma}_a = \boldsymbol{\Sigma} - \boldsymbol{\Sigma}\mathbf{j}_a\mathbf{j}_a^{\mathsf{T}}\boldsymbol{\Sigma}/(1 + \mathbf{j}_a^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{j}_a)$ and removes $(\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{j}_a)^2/(1 + \mathbf{j}_a^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{j}_a)$ from the variance of f.

Three conclusions carry beyond the two halls. An engineering output is asked about through its own function of the state, and its uncertainty lives in the full covariance: the cross terms of the budget, which the marginal spreads omit, can remove a quarter of the variance, as they do for hall 1 at 60 kW; its shares can exceed one or fall below zero and change with the coordinates of the state, so the budget explains a band and does not rank measurements. A check of the Gaussian band is complete only when it reaches the full posterior, because a correction for the curvature of the output made without the correction for the shape of the posterior can move the answer the wrong way; and the error that remains depends on how far the limit lies from the center in standard deviations, so an agreement at one limit says little about another. And the next measurement is ranked for a stated output by the squared correlation between that output and the reading it would return, so a second sensor on the hall whose conductance sets the band beats a precise sensor on the supply, while a load meter beats both for the split of the loads.

Table 20.10 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 20.10: The terms of propagation to engineering outputs and where each is defined.

Term	Expression	Units	Meaning	Defined in
Scenario output	$T_1^* = T_s + Q_1^* \mathrm{e}^{-g_1}$	°C	hall 1 at a pinned load	20.1.4
Linearized output	$f(\hat{\mathbf{z}}) + \mathbf{g}^{\mathsf{T}}(\mathbf{z} - \hat{\mathbf{z}})$	varies	tangent plane at $\hat{\mathbf{z}}$	20.3.1
Delta method	$f \mid \mathbf{y} \approx \mathcal{N}(f(\hat{\mathbf{z}}),\ \mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g})$	varies	normal output	20.3.1
Credible interval	$f(\hat{\mathbf{z}}) \pm 1.960\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}$	varies	central 95 %	20.3.1
Threshold probability	$\Phi\big((T_{\max} - f(\hat{\mathbf{z}}))/\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}\big)$	–	normal tail	20.3.2
Variance budget	$\sum_i [\mathbf{g}]_i^2 \Sigma_{ii} + 2\sum_{i<j} [\mathbf{g}]_i [\mathbf{g}]_j \Sigma_{ij}$	varies	terms and cross terms	20.3.3
Diagonal term	$V_i = [\mathbf{g}]_i^2\, \Sigma_{ii}$	varies	entry i alone	20.3.3
Cross-term bound	$\lvert 2[\mathbf{g}]_i [\mathbf{g}]_j \Sigma_{ij} \rvert \le$ $\lvert \operatorname{corr}(z_i, z_j) \rvert (V_i + V_j)$	varies	at most the correlation	20.3.3
Per-entry attribution	$[\mathbf{g}]_i (\boldsymbol{\Sigma}\mathbf{g})_i$	varies	cross terms split	20.3.3
Second-order mean	$f(\hat{\mathbf{z}}) + \frac{1}{2}\operatorname{tr}(\boldsymbol{\Sigma}\nabla^2 f)$	varies	curvature bias	20.2.3
Conditional normal	$m(x)$, $\Sigma_{5\mid 1} = \Sigma_{55} - \Sigma_{15}^2/\Sigma_{11}$	K, K^2	T_s given $g_1 = x$	20.2.3
Laplace pushforward tail	$\int \Phi\big((T_{\max} - Q_1^* \mathrm{e}^{-x} - m(x))/\sqrt{\Sigma_{5\mid 1}}\big) p(x)\, \mathrm{d}x$	–	Gaussian posterior, output not linearized	20.2.3
Importance weight	$\bar{\zeta}_j \propto \mathrm{e}^{-[\mathcal{L}(\mathbf{z}^{(j)}) - \mathcal{L}(\hat{\mathbf{z}})] + \frac{1}{2}\mathbf{d}^{(j)\mathsf{T}}\boldsymbol{\Lambda}\mathbf{d}^{(j)}}$	–	to the full posterior	20.2.4
Effective sample size	$N_{\mathrm{eff}} = (\sum_j \bar{\zeta}_j^2)^{-1}$	–	precision of weights	20.2.4
Reweighted standard error	$(\sum_j \bar{\zeta}_j^2 (\mathbf{1}_j - \hat{P})^2)^{1/2}$	–	of a probability	20.2.4
Candidate information	$\boldsymbol{\Lambda}_a = \mathbf{j}_a \mathbf{j}_a^{\mathsf{T}}$	varies	one added row	20.6.1
Rank-one preposterior	$\boldsymbol{\Sigma} - \boldsymbol{\Sigma}\mathbf{j}_a\mathbf{j}_a^{\mathsf{T}}\boldsymbol{\Sigma}/(1 + \mathbf{j}_a^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{j}_a)$	varies	matrix-inversion lemma	20.6.1
Variance removed	$(\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{j}_a)^2/(1 + \mathbf{j}_a^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{j}_a)$	varies	squared correlation	20.6.1
Width removed	$1 - \sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}_a\mathbf{g}/\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}$	–	ranks candidates	20.6.1

This chapter closes Part V of the book: the state of a physical plant has been formulated, fitted, measured for its spread and now read for its engineering outputs. The next part turns from what the fitted state predicts to what it reveals: the temperature where no sensor is installed, the reading that is lying, the fault that explains a change, and the drawing of the plant that the evidence prefers.

20.8 Exercises

Exercise 20.1. *Formulation.* A hall's posterior on its log-conductance and supply temperature is normal with $\hat{g} = \ln 5$, $\hat{T}_s = 12$°C, spreads $\sqrt{\Sigma_{11}} = 0.10$ and $\sqrt{\Sigma_{22}} = 0.30$ K and correlation 0.5. (a) Write the question whether the hall stays below 21°C at a load of 40 kW as a probability over the state, and derive its temperature and gradient there. (b) Compute the variance budget, the spread, the threshold probability and the 95 % interval. (c) Repeat the spread and the probability without the cross term. ((a) $\Pr(T_s + 40\,\mathrm{e}^{-g} < 21 \mid \mathbf{y})$; $T^* = 12 + 40/5 = 20$°C, $\mathbf{g} = (-40/5,\ 1) = (-8\,\mathrm{K},\ 1)$. (b) $64 \times 0.01 + 0.09 + 2(-8)(1)(0.5 \times 0.1 \times 0.3) = 0.64 + 0.09 - 0.24 = 0.49\,\mathrm{K}^2$, spread 0.70 K; $\Phi(1/0.7) = \Phi(1.429) = 0.923$; [18.63, 21.37]°C. (c) $\sqrt{0.73} = 0.854$ K and $\Phi(1.170) = 0.879$: omitting the negative cross term widens the band by 22 % and understates the probability by 0.044.)

Exercise 20.2. *Derivation.* (a) Derive the second-order correction $\frac{1}{2}\operatorname{tr}(\boldsymbol{\Sigma}\nabla^2 f)$ to the posterior mean of a derived quantity, and evaluate it for hall 1 at 60 kW. (b) Compare it with the mean of the Laplace pushforward and with the full posterior's mean, 22.98°C, and explain the difference in sign. ((a) $\mathbb{E}[\frac{1}{2}\mathbf{d}^\mathsf{T}\nabla^2 f\,\mathbf{d}] = \frac{1}{2}\mathbb{E}[\operatorname{tr}(\nabla^2 f\,\mathbf{d}\mathbf{d}^\mathsf{T})] = \frac{1}{2}\operatorname{tr}(\nabla^2 f\,\boldsymbol{\Sigma})$ and the linear term has zero mean; with one nonzero second derivative, 10.928 K, it is $\frac{1}{2}\times 0.011382\times 10.928 = 0.062$ K. (b) Laplace pushforward, by equation (20.9): $10.928\,(\mathrm{e}^{\Sigma_{11}/2}-1) = 0.062$ K, the same to three digits. The full posterior's mean is 0.059 K below $T_1^*(\hat{\mathbf{z}})$ because the full posterior of g_1 is skewed to the right, its mean 1.716 above the MAP value 1.703; the correction assumes $\hat{\mathbf{z}}$ is the posterior mean, which the Gaussian makes true and the full posterior does not.)

Exercise 20.3. *Derivation.* (a) Derive the rank-one form of the preposterior covariance from $\boldsymbol{\Sigma}_a = (\boldsymbol{\Sigma}^{-1} + \mathbf{j}_a\mathbf{j}_a^\mathsf{T})^{-1}$ by the matrix-inversion lemma, and show that it is the covariance update of equation (4.63) for the three-entry hall of Chapter 4: the state $\mathbf{z} = (\rho, T_1, T_2)^\mathsf{T}$ with the limit covariance $\boldsymbol{\Sigma}$ of that chapter, and a second reading of the test-load temperature T_1 of standard deviation $\sigma_T = 0.5$ K, whose normalized row $(T_1 - y_{T,2})/\sigma_T$ has the Jacobian line $\mathbf{j}_a^\mathsf{T} = (0,\ 1/\sigma_T,\ 0)$. Evaluate the three diagonal entries of $\boldsymbol{\Sigma}_a$, given the second column of $\boldsymbol{\Sigma}$, $(0.0057692,\ 0.17308,\ 0.34615)^\mathsf{T}$, and $\Sigma_{11} = 1.9231\times 10^{-4}\,(\mathrm{K/kW})^2$, $\Sigma_{33} = 0.69231\,\mathrm{K}^2$. (b) Show that the variance of any output is not increased by the candidate, and that the fraction it removes is a squared correlation. ((a) Multiplying $\boldsymbol{\Sigma}^{-1}+\mathbf{j}_a\mathbf{j}_a^\mathsf{T}$ by $\boldsymbol{\Sigma}-\boldsymbol{\Sigma}\mathbf{j}_a\mathbf{j}_a^\mathsf{T}\boldsymbol{\Sigma}/s_a$, $s_a = 1+\mathbf{j}_a^\mathsf{T}\boldsymbol{\Sigma}\mathbf{j}_a$, gives $\mathbf{I}+\mathbf{j}_a\mathbf{j}_a^\mathsf{T}\boldsymbol{\Sigma}\,(1-s_a/s_a) = \mathbf{I}$, which is equation (20.28). With $\mathbf{j}_a = \mathbf{h}/\sigma_T$, $\mathbf{h} = (0,\ 1,\ 0)^\mathsf{T}$, the subtracted term is $\boldsymbol{\Sigma}\mathbf{h}\mathbf{h}^\mathsf{T}\boldsymbol{\Sigma}/(\sigma_T^2+\Sigma_{22}) = S\,\mathbf{K}\mathbf{K}^\mathsf{T}$ with $S = \Sigma_{22}+\sigma_T^2 = 0.17308+0.25 = 0.42308\,\mathrm{K}^2$ and the three-entry gain vector $\mathbf{K} = \boldsymbol{\Sigma}\mathbf{h}/S = (0.013636\,\mathrm{kW}^{-1},\ 0.40909,\ 0.81818)^\mathsf{T}$, so $\boldsymbol{\Sigma}_a = \boldsymbol{\Sigma}-S\,\mathbf{K}\mathbf{K}^\mathsf{T}$, equation (4.63). Its diagonal is $1.9231\times 10^{-4} - 0.42308\times 0.013636^2 = 1.1364\times 10^{-4}\,(\mathrm{K/kW})^2$, $0.17308 - 0.42308\times 0.40909^2 = 0.10227\,\mathrm{K}^2$ and $0.69231 - 0.42308\times 0.81818^2 = 0.40909\,\mathrm{K}^2$, the values of the precision form of that chapter. (b) $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}_a\mathbf{g} = \mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g} - (\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{j}_a)^2/(1+\mathbf{j}_a^\mathsf{T}\boldsymbol{\Sigma}\mathbf{j}_a) \le \mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$, since the subtracted term is a square over a positive number; the removed fraction is $\operatorname{Cov}(f,\tilde{y}_a)^2/(\operatorname{Var} f\,\operatorname{Var}\tilde{y}_a)$, with $\tilde{y}_a = \mathbf{j}_a^\mathsf{T}\mathbf{d} + \epsilon_a$ the normalized predicted reading and ϵ_a its unit-variance error.)

Exercise 20.4. *Derivation.* (a) Show that the cross term of two entries in a variance budget is negative exactly when their correlation and the product of their gradient entries have opposite signs, and that its size is at most the correlation times the sum of the two diagonal terms. (b) For an output of two entries, derive the largest share of the total that the two diagonal terms together can reach, and evaluate it for hall 1 at 60 kW, where $\operatorname{corr}(z_1, z_5) = 0.434$. ((a) $2[\mathbf{g}]_i[\mathbf{g}]_j\Sigma_{ij} = 2\operatorname{corr}(z_i,z_j)\operatorname{sgn}([\mathbf{g}]_i[\mathbf{g}]_j)\sqrt{V_iV_j}$, whose sign is that of $\operatorname{corr}(z_i,z_j)[\mathbf{g}]_i[\mathbf{g}]_j$; by $2ab \le a^2+b^2$ its size is at most $|\operatorname{corr}(z_i,z_j)|(V_i+V_j)$. (b) With D the diagonal sum and c the absolute correlation of the two entries, $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g} \ge (1-c)D$, so $D/\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g} \le 1/(1-c) = 1/0.566 = 1.77$; for the hall the cross term is $-0.2836 = -2\times 0.434\times 1.166\times 0.280\,\mathrm{K}^2$, equal to $2|\operatorname{corr}(z_1,z_5)|\sqrt{V_1V_5}$ in size, as it must be for two entries, and the diagonal share is $1.4378/1.1543 = 1.246$.)

Exercise 20.5. *Application.* Repeat the scenario at $Q_1^* = 90$ kW. (a) Compute $T_1^*(\hat{\mathbf{z}})$, the gradient entry, the budget and the delta band. (b) A limit of 31.5°C lies about as many standard deviations above the center at 90 kW as 25°C does at 60 kW. The same 200 000 draws give, by the Laplace pushforward, mean 28.60°C, band $[25.61,\ 32.13]$°C and $\Pr(T_1^* < 31.5\,^\circ\mathrm{C}) = 0.951$, and under the full posterior mean 28.40°C, band $[25.32,\ 31.74]$°C and probability 0.966. Compare with the delta method, and state what sets the delta method's error. ((a) $12.111 + 90/5.4903 = 12.111 + 16.393 = 28.504$°C; $[\mathbf{g}]_1 = -16.39$ K; $3.0587 + 0.0784 - 0.4253 = 2.7118\,\mathrm{K}^2$, spread 1.647 K, 1.53 times the 60 kW value, as the lever; band $[25.28,\ 31.73]$°C; second-order term 0.093 K. (b) $(31.5 - 28.504)/1.647 = 1.819$ and $\Phi(1.819) = 0.966$, within 0.0005 of the full posterior's, as at 25°C and 60 kW, while the Laplace pushforward is again low by 0.015. The delta band matches the full posterior's within 0.04 K at both ends. The error is set by the limit's standardized distance, not by the load: at 30°C, 0.91 standard deviations above the center,

the delta method gives 0.818 against the full posterior's 0.835, an error of -0.017 against the -0.016 it shows at 24°C, 0.89 standard deviations, for 60 kW; the load scales the resistance's contribution and its spread together and leaves the standardized shape almost unchanged.)

Exercise 20.6. *Application.* Hall 2 at the same scenario load has $T_2^* = T_s + 60\,\mathrm{e}^{-g_2}$, with $T_2^*(\hat{\mathbf{z}}) = 26.91$°C and gradient $(0, -14.80, 0, 0, 1)^{\mathsf{T}}$. (a) Compute its spread and its probability of staying below 25°C. (b) Rank the second hall-1 sensor and the hall-2 load meter of Table 20.9 for T_1^* and for T_2^*, given that the preposterior spreads of T_2^* are 1.958 and 1.554 K, and explain the flip. ((a) $14.80^2 \times 0.018765 + 0.078433 + 2(-14.80)(0.011989) = 4.1085 + 0.0784 - 0.3548 = 3.8321\,\mathrm{K}^2$, spread 1.958 K; $\Phi((25 - 26.908)/1.958) = \Phi(-0.975) = 0.165$: hall 2, with the smaller conductance, would probably exceed the limit. (b) For T_1^* the hall-1 sensor wins, 0.874 against 1.048 K; for T_2^* the load meter wins, 1.554 against 1.958 K, the hall-1 sensor removing nothing. The load meter fixes Q_2, and with it, through the hall-2 reading and the correlation 0.691 of g_2 with Q_2, the conductance that sets T_2^*; the hall-1 sensor is uncorrelated with g_2.)

Part VI

Diagnosis and Competing Hypotheses

Chapter 21

Virtual Sensors and Unobserved States

21.1 Quantities where no sensor is installed

The temperature sensor of hall 2 has been taken out of service for calibration, and the operators of the two halls still have to say how warm hall 2 is. The shift log offers the obvious answer: the sensor read 17.70°C when it came out, nothing in the plant has been changed since, so hall 2 is entered at 17.70°C with the sensor's error of 0.5 K. The engineer who signs the log is less sure. Nothing now reads hall 2, and whatever number is written while the sensor is away rests on the other four instruments and on the physics that ties them to the hall. A second place has the same difficulty permanently: the return header, where the water from both halls mixes on its way back to the chiller, has never had a sensor, and the chiller limits the temperature of the water it receives to 18°C.

The log's entry answers where hall 2 was, not how well the rest of the plant knows where it is, and it says nothing about whether the last reading was right. The question of this chapter is

how the distribution of a quantity that is not directly measured is derived from the fitted state; what removing a reading from the fit does to the state and its covariance; how a reading held out of the fit is predicted and used as a check; and what a real sensor at the same place would add, so that a virtual and a real sensor are compared on one scale.

It is answered on the plant of Part V with the method of the earlier chapters. What changes is the quantity asked for: it is not an entry of the reduced state and it has no reading row, either because its row has been removed from the fit or because it never had one.

Jargon: virtual sensor, soft sensor, unmeasured quantity, return header, heat-capacity rate, loop balance, held-out reading, leave-one-out residual, rank-one downdate, critical measurement, predictive distribution of a reading, leverage

A *virtual sensor*, called a *soft sensor* in process control, is a quantity of a plant computed from a model and the readings of other instruments instead of being read by its own; an *unmeasured quantity* is any quantity of the plant without a reading row. The *return header* is the pipe in which the water leaving the cooling coils mixes before it re-enters the chiller. The *heat-capacity rate* $\dot{m}c_p$ of the loop, in kW/K, is the mass flow of water times its specific heat, the heat the flowing water carries per kelvin of temperature rise, and the *loop balance* equates the heat the water gains between the supply and the return headers with the heat the halls deliver. A *held-out reading* is a reading removed from a fit so that the fit can predict it; the *leave-one-out residual* is the difference between the held-out reading and that prediction. The *rank-one downdate* removes one outer product from an information matrix and corrects the covariance accordingly. A *critical measurement* is a reading whose removal makes some direction of the state unobservable; its redundancy number is zero. The *predictive distribution of a reading* is the distribution of a reading not yet seen, given the evidence in hand. The *leverage* of a row is the diagonal entry K_{kk} of the hat matrix of Chapter 18, the complement of its redundancy number.

Key idea

A virtual sensor is a derived quantity of the fitted state, read with the same delta method, budget and checks as a scenario output; what is new is the fit with a reading removed and the comparison with a sensor that could be installed. Removing a row adds to the covariance one outer product divided by the row's redundancy number, so a reading that the rest of the plant barely checks leaves behind it a large virtual spread, set by whatever prior remains. A quantity that several readings fix in combination can be known better than an instrument would read it. In the linear model the fit gives a reading the weight K_{kk} and the virtual sensor of the same quantity the weight $1 - K_{kk}$, the held-out check of a reading is its normalized residual, and a real sensor at a virtual sensor's place adds its precision to the virtual one for its own quantity and removes from another output only what that output shares with it.

21.1.1 From a scenario output to a quantity of the present plant

The scenario node of Chapter 20 was a derived quantity of the state that carried no row: hall 1 at 60 kW, sharing the conductance g_1 and the supply temperature T_s with the observed graph. The quantities of this chapter are of the same mathematical kind and of a different engineering kind. They are quantities of the plant in its present condition, so that, unlike a scenario value, each one could in principle be measured now by installing an instrument, and the comparison of the virtual answer with the answer such an instrument would give is part of the question. Two cases occur in practice. A quantity can have had a sensor that has failed or been removed, so that its reading row is gone from the fit and the quantity becomes a derived quantity of a fit with one row fewer; the hall temperatures are such quantities, functions of the reduced state of Chapter 17 through the hall balances and, without their readings, predicted by them. A quantity can also be one that the plant has never measured, a node of the graph whose value is fixed by a physical relation and the state but which no reading row has ever touched; the return-header temperature is such a quantity. In both cases the answer is the posterior of a function of the state, and when the function is an entry of the state, an unmeasured entry, its posterior is the marginal on the diagonal of the covariance.

The condition-monitoring chapter of the companion volume *Reliability Engineering* uses

virtual sensors to watch quantities that are expensive or impossible to instrument, and to cross-check the instruments that exist. This chapter derives what such a sensor's spread is made of, and when it is narrower or wider than the error of the instrument it stands in for.

21.1.2 A virtual sensor and what stands behind its spread

A virtual sensor's spread is computed from the same covariance $\mathbf{\Sigma}$ as every other output of the fit, and it is only as good as the rows that inform the entries the quantity depends on. Where those rows are several independent readings tied together by the physics, the spread can be smaller than the error of an instrument the operator would install. Where the only row left behind an entry is a handbook prior, the spread is the prior's spread pushed through the physics, and the virtual sensor reports the handbook rather than the plant.

> **Principle 21.1.** *A virtual sensor is the posterior of a derived quantity of the fitted state, with the variance* $\mathbf{g}^{\mathsf{T}}\mathbf{\Sigma}\mathbf{g}$ *and the checks of any other derived quantity. Its spread measures what the remaining rows say about the quantity, which may be much or nothing beyond a prior.*

21.1.3 The plant graph and the return header (step 0)

The plant of Chapter 17 has two halls, each with a heat load Q_i in kW rejected through a heat-removal path of conductance $G_i = G_{\text{ref}}\mathrm{e}^{g_i}$ in kW/K to a supply header at temperature T_s, with $G_{\text{ref}} = 1\,\text{kW/K}$, so that e^{g_i} is the conductance in kW/K. The state is the reduced state of that chapter, the five free entries,

$$\mathbf{z} = (z_1, z_2, z_3, z_4, z_5) = (g_1,\ g_2,\ Q_1,\ Q_2,\ T_s), \tag{21.1}$$

and the hall temperatures are functions of it through the balances, equation (17.4):

$$T_i(\mathbf{z}) = T_s + Q_i\,\mathrm{e}^{-g_i}, \qquad i = 1, 2, \tag{21.2}$$

in °C, with G_{ref} absorbed.

The water leaves the chiller at the supply temperature T_s, passes through the coils of the two halls, where it takes up their loads, and mixes in the return header before it re-enters the chiller. Let $\dot{m}$ be the mass flow of water in the loop, in kg/s, c_p its specific heat, in kJ/(kg K), and T_r the temperature in the return header, in °C. In steady state the water stores no heat, and if the heat gained by the pipes and the pump work are neglected, as in the header balance (17.5), the heat the water carries away from the halls is its heat-capacity rate times its temperature rise:

$$\dot{m}c_p\,(T_r - T_s) = Q_1 + Q_2. \tag{21.3}$$

Since $\dot{m}c_p > 0$, dividing equation (21.3) by it and adding T_s,

$$T_r(\mathbf{z}) = T_s + \frac{Q_1 + Q_2}{\dot{m}c_p}, \qquad \dot{m}c_p = 10\,\text{kW/K}. \tag{21.4}$$

The heat-capacity rate is pinned. The loop's flow meter reads the flow, about 2.39 kg/s, and the specific heat of water near 15°C is 4.19 kJ/(kg K); their product is 10.0 kW/K, known to about 1 to 2 % (an illustrative accuracy) when the flow meter is in calibration. The return temperature is linear in the state, a fixed combination of T_s and the sum of the loads, and it has no row: no reading in the plant depends on it, so adding the return header to the graph leaves the posterior of the state unchanged, as the scenario node of Chapter 20 did.

Figure 21.1 draws the graph of this chapter. The green boxes are free nodes carrying entries of the state: the two loads at the top and the supply header with T_s in the middle. The blue

boxes are the halls, balance nodes whose temperatures are functions of the state; the solid green arrows from the halls to the supply header are the heat-removal paths, labeled by their log-conductances. The dashed blue box at the bottom is the return header, a balance node with no reading, fed by the loop's flow from the supply header and joined by a gray arrow to the pinned heat-capacity rate on its left; the white box on its right is the chiller, which returns the water to the supply header. The dashed gray arrows leave the graph for the readings. The hall-2 reading on the right is drawn struck out, because its row is removed from the fit while the sensor is away; hall 2's box is drawn dashed for the same reason.

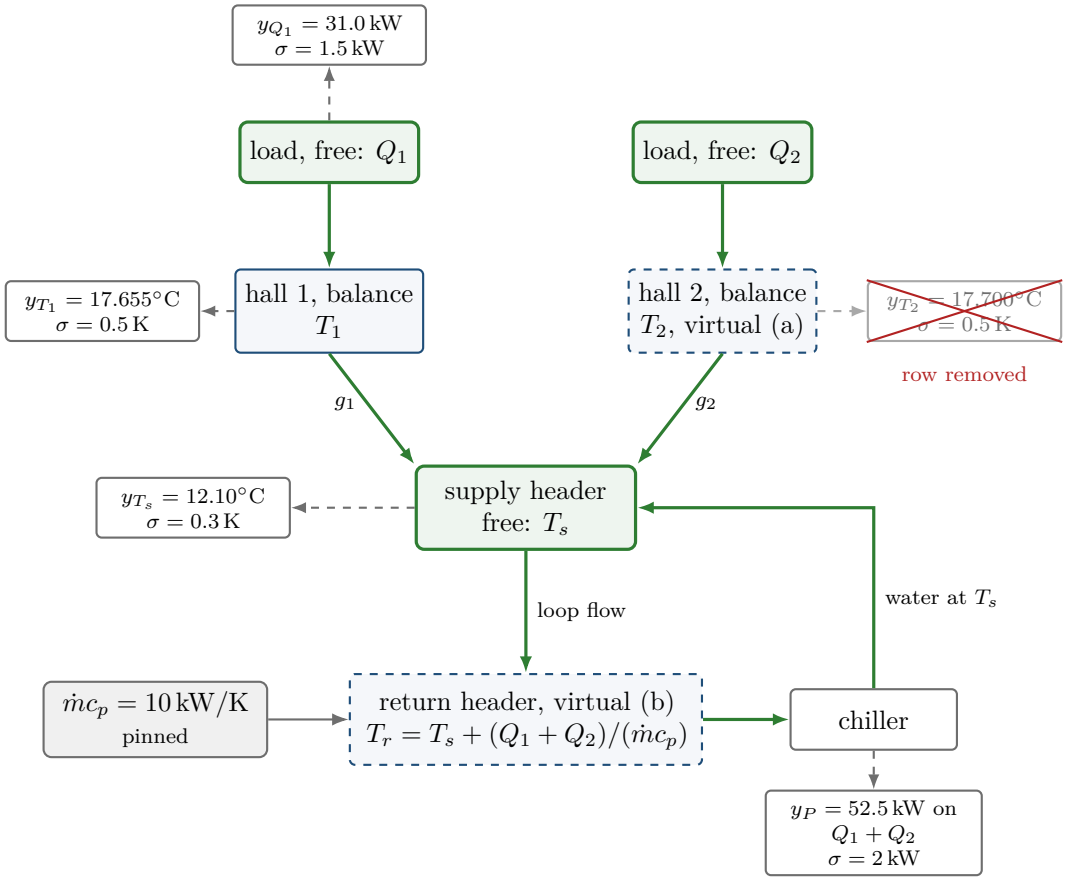


Figure 21.1: The two-hall plant with its two virtual sensors. Free (green): the loads Q_1, Q_2 and the supply temperature T_s; the log-conductances g_1, g_2 label the heat-removal paths. Balance nodes (blue): the halls, whose temperatures are functions of the reduced state; hall 2, dashed, is virtual sensor (a), its reading row removed (struck out). The return header, dashed, is virtual sensor (b): a balance node with no reading, fixed by the loop balance (21.4) with the pinned heat-capacity rate $\dot{m}c_p = 10\,\mathrm{kW/K}$ (gray). Observed (dashed gray arrows): hall 1's temperature, the supply temperature, the chiller's heat meter on $Q_1 + Q_2$ and the electrical meter on Q_1. The priors on g_1, g_2 and T_s are rows not drawn.

21.1.4 The record and the two bodies of evidence

The record is that of Chapter 17: the five readings of Table 17.1, the two hall sensors ($\sigma_T = 0.5\,\mathrm{K}$), the supply sensor ($0.3\,\mathrm{K}$), the chiller's heat meter on $Q_1 + Q_2$ ($2\,\mathrm{kW}$) and the electrical meter on hall 1's load ($1.5\,\mathrm{kW}$), with the three priors of Table 17.2, the handbook priors $g_i \sim \mathcal{N}(\ln 5,\ 0.3^2)$ and the set-point prior $T_s \sim \mathcal{N}(12°\mathrm{C},\ 1\,\mathrm{K}^2)$. The readings are illustrative, generated from known values with stated errors. Two bodies of evidence occur. $\mathbf{y}$ is the five readings with the three priors, as in Part V; $\mathbf{y}_-$ is the same evidence without the hall-2 reading $y_{T_2} = 17.70°\mathrm{C}$, the

evidence of the plant while that sensor is out of service. The return header's question is asked of $\mathbf{y}$, since it adds no row; hall 2's of $\mathbf{y}_-$, with y_{T_2} kept aside as the reading to be predicted.

21.1.5 The questions (step 1)

The operators ask six questions of the plant, and step 1 writes each as a probability or an expectation over the state, with the unknown on the left of the bar and the evidence on the right. Table 21.1 lists them: the first column is the question in engineering language; the second the general form of its category in the notation of Chapter 2; the third its mathematical form for the plant; the fourth its category; and the last the subsection where it is answered. The first is the plant owner's question, in the owner's words, with hall 2 as its case. The first two ask for the distribution of a derived quantity, the third and fourth for the split of its variance, by entry of the state and by row of the evidence, the fourth comparing the posteriors under $\mathbf{y}$ and $\mathbf{y}_-$ through the observability rule of Chapter 19. The fifth asks whether the held-out reading agrees with what the rest of the plant predicts for it, a tail of the predictive distribution of a reading, and the sixth is a threshold on the return temperature. Whether a real sensor at either place would improve the answers asks nothing new of the present evidence; it is the next measurement for these six, Section 21.6. The atlas of Appendix A lists the first and the sixth.

Table 21.1: The questions asked of the two-hall plant, the general form of their category, their mathematical form, their category and where each is answered. $\mathbf{y}$ is the five readings and the three priors; $\mathbf{y}_-$ the same without the hall-2 reading; T_2 and T_r the functions (21.2) and (21.4) of the state; $\mathbf{g}$ the gradient of a virtual quantity; $T_{\max} = 18°\mathrm{C}$ the chiller's limit on its return water.

Engineering question	General form	Mathematical form	Category	Answered in
What is the temperature where there is no sensor? (hall 2 with its sensor out)	$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y})$	$\Pr(T_2(\mathbf{z}) \le t \mid \mathbf{y}_-)$, $T_2 = T_s + Q_2 \mathrm{e}^{-g_2}$	estimation of a derived quantity	21.3.1
What is the return-header temperature, which has never been measured?	$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y})$	$\Pr(T_r(\mathbf{z}) \le t \mid \mathbf{y})$, $T_r = T_s + (Q_1 + Q_2)/(\dot{m} c_p)$	estimation of a derived quantity	21.3.1
Which entries carry each virtual reading's uncertainty?	$\mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y})$	$\mathbf{g}^{\mathsf{T}} \boldsymbol{\Sigma} \mathbf{g}$ entry by entry, for T_2 under $\mathbf{y}_-$ and T_r under $\mathbf{y}$	estimation (variance split by entry)	21.3.2
How much did the removed sensor know, and what does the virtual sensor lose?	$\mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y})$ under two evidences	$\mathrm{Var}(T_2 \mid \mathbf{y})$ against $\mathrm{Var}(T_2 \mid \mathbf{y}_-)$; the spreads of $\mathbf{z}$ under both	estimation (two posteriors compared)	21.3.2
Does the sensor's last reading agree with the rest of the plant?	$p(y_{\text{new}} \mid \mathbf{y}) = \int p(y_{\text{new}} \mid \mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}$, a tail of it	$p(y_{T_2} \mid \mathbf{y}_-)$ and $\Pr(Y_{T_2} \ge 17.70°\mathrm{C} \mid \mathbf{y}_-)$	prediction under a threshold (a predictive tail)	21.3.3
Is the return temperature below the chiller's limit of 18°C?	$\Pr(f(\mathbf{z}) \ge f^* \mid \mathbf{y})$	$\Pr(T_r(\mathbf{z}) < T_{\max} \mid \mathbf{y})$	prediction under a threshold	21.3.4

21.1.6 The evidence as rows (step 2)

Step 2 attaches the evidence to the graph as the eight rows of the reduced state, equation (17.15): the five reading rows and the three prior rows. The hall-2 reading row is row 2,

$$r_2(\mathbf{z}) = \frac{T_s + Q_2\,\mathrm{e}^{-g_2} - y_{T_2}}{\sigma_T}, \qquad y_{T_2} = 17.70^{\circ}\mathrm{C}, \tag{21.5}$$

and it is the only reading row in which g_2 appears. For the return header nothing is added: the node carries no row. For hall 2 with its sensor out, the evidence loses that row, and the objective becomes

$$\mathcal{L}_-(\mathbf{z}) = \mathcal{L}(\mathbf{z}) - \tfrac{1}{2}\,r_2(\mathbf{z})^2, \tag{21.6}$$

with $\mathcal{L}$ the objective of equation (18.9); the posterior becomes $p(\mathbf{z} \mid \mathbf{y}_-) \propto \mathrm{e}^{-\mathcal{L}_-(\mathbf{z})}$, with its own MAP state $\hat{\mathbf{z}}_-$ and covariance $\mathbf{\Sigma}_-$. Seven rows then act on five entries, and the redundancy falls from three to two. Removing the row is not the same as removing the hall: the balance (21.2) still holds, and T_2 is still a function of the state, now one that no reading checks.

The modeling choices. Three modeling choices enter every answer. The state is the reduced state of Chapter 17: its Jacobian is the projection $\mathbf{J} = \mathbf{J}_{\mathrm{e}}\mathbf{N}$ of the reading and prior rows on the null directions of the physics rows, equation (17.18), and its covariance is the free-entry marginal of the full state's in the limit $\sigma_p \to 0$, equation (17.19); this is what makes T_2 a function of five entries instead of an entry of eight, and it fails if a hall balance misses by a physical amount, which would then be a discrepancy with its own variance (the limit is checked in Section 21.4.2). The sensor's absence removes its row and nothing else: the other readings describe the same plant, so the objective loses one term, equation (21.6); this fails if the outage coincides with a change of the plant, such as a fouling coil, which no remaining row would see (Section 21.6.6). The loop balance is taken with the pump work and the pipe gains neglected and $\dot{m}c_p$ pinned at 10 kW/K: this makes T_r a linear function of the state with no new entry, and it fails if the flow meter is out of calibration, an error that no interval below contains (Section 21.6.6).

Which row acts on which entry is the pattern of the reduced Jacobian of Table 17.5, a property of the graph; with $*$ for an entry that can be nonzero,

$$\mathbf{J} = \begin{bmatrix} * & 0 & * & 0 & * \\ 0 & * & 0 & * & * \\ 0 & 0 & 0 & 0 & * \\ 0 & 0 & * & * & 0 \\ 0 & 0 & * & 0 & 0 \\ * & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * \end{bmatrix} \begin{matrix} \text{hall-1 reading} \\ \text{hall-2 reading (removed under } \mathbf{y}_-) \\ \text{supply reading} \\ \text{heat meter} \\ \text{load meter} \\ \text{prior on } g_1 \\ \text{prior on } g_2 \\ \text{prior on } T_s \end{matrix}, \tag{21.7}$$

$$\mathbf{g}_{T_2}^{\mathsf{T}} = (0,\ *,\ 0,\ *,\ *), \qquad \mathbf{g}_{T_r}^{\mathsf{T}} = (0,\ 0,\ *,\ *,\ *),$$

with columns g_1, g_2, Q_1, Q_2, T_s. The two virtual quantities are not rows; their gradients, on the second line, have the patterns of the functions (21.2) and (21.4). The column of g_2 holds two rows, the hall-2 reading and the prior on g_2; with the reading removed it holds the prior alone. The column of Q_2 then holds only the heat meter, which reads $Q_1 + Q_2$. The gradient of T_r shares its columns with the heat meter, the load meter and the supply sensor and none with the conductances.

21.2 The posterior of the state (step 3)

In the book's notation $p(\mathbf{z} \mid \mathbf{y})$ is the posterior density of the state, $f(\mathbf{z})$ a quantity derived from it and f^* a value or a requirement level of that quantity.

Six objects recur in the specialized forms. The *MAP state* $\hat{\mathbf{z}}$ of the five-reading fit of Chapter 18 is the minimum of the objective that the eight rows put on the state; the *information matrix* $\mathbf{\Lambda} = \mathbf{J}^\mathsf{T}\mathbf{J}$ is the curvature of that objective at the minimum, with $\mathbf{J}$ the 8×5 Jacobian of the normalized rows; the *Laplace covariance* $\mathbf{\Sigma} = \mathbf{\Lambda}^{-1}$ of Chapter 19 makes the posterior approximately normal, $\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}, \mathbf{\Sigma})$. The fit without the hall-2 row has its own MAP state $\hat{\mathbf{z}}_-$ and covariance $\mathbf{\Sigma}_-$. A virtual quantity f is carried by its *gradient* $\mathbf{g} = \partial f/\partial \mathbf{z}$ at the MAP state of the evidence in use, with the variance $\mathbf{g}^\mathsf{T}\mathbf{\Sigma}\mathbf{g}$, and a reading row k by its normalized gradient $\mathbf{j}_k = \nabla r_k(\hat{\mathbf{z}})$, the k-th row of $\mathbf{J}$. All six questions are read from two posteriors of the reduced state, with and without the hall-2 row, that the rows of step 2 give, each approximately a Laplace Gaussian. This section derives them and computes them for the plant; Section 21.3 reads every answer from them.

21.2.1 The rows and the posterior without the hall-2 row

The questions use two posteriors, with and without the hall-2 row. Both rest on the eight rows of the reduced state, equation (17.15), each a Gaussian row normalized by its standard deviation, whose terms $\mathcal{L}_k = \frac{1}{2}r_k^2$ sum to the objective (18.9). The fit is the Gauss–Newton iteration of Chapter 18, the normal equations $(\mathbf{J}^\mathsf{T}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^\mathsf{T}\mathbf{r}$ of equation (17.23) repeated from the prior point to convergence, and the covariance at the minimum is $\mathbf{\Sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}$ of Chapter 19. With every row normalized the information matrix is $\mathbf{\Lambda} = \mathbf{J}^\mathsf{T}\mathbf{J}$; with unnormalized rows and weights $w_k = 1/\sigma_k^2$ it is $\mathbf{J}^\mathsf{T}\mathbf{W}\mathbf{J}$, $\mathbf{W} = \mathrm{diag}(w_k)$. The fit without the hall-2 row is the same iteration on $\mathcal{L}_-$ of equation (21.6), with seven rows; what removing the row does to the covariance and to the state can also be read from the five-reading fit alone.

Removing a row from the information. Let row k have the normalized form $r_k(\mathbf{z})$, with unit weight, and gradient $\mathbf{j}_k = \nabla r_k(\hat{\mathbf{z}})$. The Gauss–Newton information is a sum of outer products of rows, equation (19.11), so removing row k while the Jacobian is held at $\hat{\mathbf{z}}$ removes one of them:

$$\tilde{\mathbf{\Lambda}}_- = \mathbf{\Lambda} - \mathbf{j}_k\mathbf{j}_k^\mathsf{T}, \qquad \tilde{\mathbf{\Sigma}}_- = \tilde{\mathbf{\Lambda}}_-^{-1}. \tag{21.8}$$

The tilde marks the downdated covariance, computed at the fixed linearization point $\hat{\mathbf{z}}$; for linear rows it equals the refit's $\mathbf{\Sigma}_-$, and for nonlinear rows the two differ by the move of the point. When $K_{kk} < 1$, with the leverage $K_{kk} = \mathbf{j}_k^\mathsf{T}\mathbf{\Sigma}\mathbf{j}_k$ of the row,

$$\tilde{\mathbf{\Sigma}}_- = \mathbf{\Sigma} + \frac{\mathbf{\Sigma}\mathbf{j}_k\,\mathbf{j}_k^\mathsf{T}\mathbf{\Sigma}}{1 - K_{kk}}. \tag{21.9}$$

To verify it, multiply the right-hand side by $\tilde{\mathbf{\Lambda}}_- = \mathbf{\Sigma}^{-1} - \mathbf{j}_k\mathbf{j}_k^\mathsf{T}$ of equation (21.8):

$$\begin{aligned}(\mathbf{\Sigma}^{-1} - \mathbf{j}_k\mathbf{j}_k^\mathsf{T})\left(\mathbf{\Sigma} + \tfrac{1}{1-K_{kk}}\,\mathbf{\Sigma}\mathbf{j}_k\mathbf{j}_k^\mathsf{T}\mathbf{\Sigma}\right) &= \mathbf{I} - \mathbf{j}_k\mathbf{j}_k^\mathsf{T}\mathbf{\Sigma} + \tfrac{1}{1-K_{kk}}\,\mathbf{j}_k\mathbf{j}_k^\mathsf{T}\mathbf{\Sigma} - \tfrac{1}{1-K_{kk}}\,\mathbf{j}_k\left(\mathbf{j}_k^\mathsf{T}\mathbf{\Sigma}\mathbf{j}_k\right)\mathbf{j}_k^\mathsf{T}\mathbf{\Sigma} \\ &= \mathbf{I} + \mathbf{j}_k\mathbf{j}_k^\mathsf{T}\mathbf{\Sigma}\left(-1 + \frac{1 - K_{kk}}{1 - K_{kk}}\right) = \mathbf{I},\end{aligned}$$

where the second line collects the three terms in $\mathbf{j}_k\mathbf{j}_k^\mathsf{T}\mathbf{\Sigma}$ and uses $\mathbf{j}_k^\mathsf{T}\mathbf{\Sigma}\mathbf{j}_k = K_{kk}$. Equation (21.9) is the matrix-inversion lemma (20.28) with the sign of the outer product reversed: the update divides by $1 + \mathbf{j}_a^\mathsf{T}\mathbf{\Sigma}\mathbf{j}_a$, the downdate by $1 - K_{kk}$, the redundancy number of the row removed. Multiplying it by $\mathbf{j}_k$ gives

$$\tilde{\mathbf{\Sigma}}_-\mathbf{j}_k = \mathbf{\Sigma}\mathbf{j}_k + \frac{\mathbf{\Sigma}\mathbf{j}_k\,K_{kk}}{1 - K_{kk}} = \frac{\mathbf{\Sigma}\mathbf{j}_k}{1 - K_{kk}}, \qquad \mathbf{j}_k^\mathsf{T}\tilde{\mathbf{\Sigma}}_-\mathbf{j}_k = \frac{K_{kk}}{1 - K_{kk}}, \tag{21.10}$$

the second relation being, for linear rows, V_-/σ_k^2 of Section 21.3.2 read on the reading's own scale.

Removing a row never reduces a variance. For any derived quantity with gradient $\mathbf{g}$, sandwiching equation (21.9),

$$\mathbf{g}^{\mathsf{T}}\tilde{\boldsymbol{\Sigma}}_{-}\mathbf{g} = \mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g} + \frac{\left(\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{j}_k\right)^2}{1 - K_{kk}} \;\geq\; \mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}, \tag{21.11}$$

since the added term is a square over a positive number. With $\mathbf{g}$ a unit vector this holds for every entry: at a fixed linearization point, no spread shrinks when a row is removed, the mirror of Exercise 19.3, where no spread grows when a row is added. The widening is large when the row's redundancy number is small: the denominator $1 - K_{kk}$ is the fraction of the row's error variance that the other rows can check, and a row that the others barely check has carried information that no other row duplicates.

A critical measurement. When $K_{kk} = 1$, the redundancy number is zero and equation (21.9) fails. Then $\mathbf{j}_k^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{j}_k = 1$, and the vector $\mathbf{b} = \boldsymbol{\Sigma}\mathbf{j}_k$ satisfies $\tilde{\boldsymbol{\Lambda}}_{-}\mathbf{b} = \mathbf{j}_k - \mathbf{j}_k(\mathbf{j}_k^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{j}_k) = \mathbf{0}$, so $\tilde{\boldsymbol{\Lambda}}_{-}$ is singular: the row was a critical measurement, the only information on the direction $\mathbf{b}$, and without it that direction is unobservable in the sense of Chapter 19. Whether a removal leaves $\tilde{\boldsymbol{\Lambda}}_{-}$ singular depends on the rows that remain. For the hall-2 row it does not: by the pattern (21.7) the prior on g_2, the heat meter on Q_2 and the supply sensor and its prior on T_s still hold every direction the row touched.

The state without the row. The MAP state moves when a row is removed, and the move can be computed without a new fit. At $\hat{\mathbf{z}}$ the gradient of the full objective is zero, $\mathbf{J}^{\mathsf{T}}\mathbf{r}(\hat{\mathbf{z}}) = \mathbf{0}$, so by equation (21.6) the gradient of the reduced objective there is the removed row's term with its sign reversed,

$$\nabla\mathcal{L}_{-}(\hat{\mathbf{z}}) = \mathbf{J}^{\mathsf{T}}\mathbf{r}(\hat{\mathbf{z}}) - r_k(\hat{\mathbf{z}})\,\mathbf{j}_k = -\,r_k(\hat{\mathbf{z}})\,\mathbf{j}_k.$$

One Gauss–Newton step from $\hat{\mathbf{z}}$ on the reduced objective, with its curvature $\tilde{\boldsymbol{\Lambda}}_{-}$ and the relation (21.10), gives

$$\hat{\mathbf{z}}_{-} \approx \hat{\mathbf{z}} - \tilde{\boldsymbol{\Sigma}}_{-}\,\nabla\mathcal{L}_{-}(\hat{\mathbf{z}}) = \hat{\mathbf{z}} + \tilde{\boldsymbol{\Sigma}}_{-}\mathbf{j}_k\,r_k(\hat{\mathbf{z}}) = \hat{\mathbf{z}} + \frac{\boldsymbol{\Sigma}\mathbf{j}_k\,r_k(\hat{\mathbf{z}})}{1 - K_{kk}}. \tag{21.12}$$

For linear rows the reduced objective is quadratic and the step is exact, the approximation an equality; for nonlinear rows it is the first step of a refit started at $\hat{\mathbf{z}}$. Like the downdate of the covariance, the move is large when the redundancy number is small, and multiplying it by $\mathbf{j}_k^{\mathsf{T}}$ returns the leave-one-out residual (21.23), $r_k(\hat{\mathbf{z}}) + K_{kk}r_k(\hat{\mathbf{z}})/(1 - K_{kk})$.

21.2.2 The posteriors with and without the hall-2 row

The plant's inputs are the five readings of Table 17.1 and the three priors of Table 17.2, the hall-2 row removed for the evidence $\mathbf{y}_{-}$. The five-reading fit is that of Chapter 18, with the covariance of Table 19.3. The Gauss–Newton iteration on the reduced objective (21.6), from the prior point, converges to $\hat{\mathbf{z}}_{-}$ with $2\mathcal{L}_{-}(\hat{\mathbf{z}}_{-}) = 0.123$. Table 21.2 compares the two fits. Each row is an entry of the state; the columns give the MAP value and posterior standard deviation of the five-reading fit and of the fit without the hall-2 row, and in the last column the ratio of the prior spread to the posterior spread without the row, equation (19.25).

Without the row, g_2 is at its prior, mean $\ln 5$ and spread 0.300, uncorrelated with every other entry. The sum of the loads is the heat meter's reading, $\hat{Q}_1 + \hat{Q}_2 = 52.50\,\mathrm{kW}$ with variance $\sigma_P^2 = 4\,\mathrm{kW}^2$, so that $\mathrm{Var}(Q_2 \mid \mathbf{y}_{-}) = 4 + 2.199 = 6.199\,\mathrm{kW}^2$ and $\mathrm{Cov}(Q_1, Q_2 \mid \mathbf{y}_{-}) = -2.199\,\mathrm{kW}^2$. The entries of hall 1 and the supply temperature move by at most a sixth of their spreads.

Every answer of step 4 is computed from these two posteriors of the state, $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ and $(\hat{\mathbf{z}}_{-}, \boldsymbol{\Sigma}_{-})$. The full posterior, $\mathrm{e}^{-\mathcal{L}}$ without the Gaussian approximation, gives the value set beside an answer where the chapter has one: 200 000 draws of the Laplace normal, seeded with 21 and reweighted to it by equation (20.11).

Table 21.2: The fit of the two halls with and without the hall-2 reading row: MAP value and posterior standard deviation for the five-reading fit and for the refit with the row removed; last column, the ratio of prior to posterior spread without the row (flat priors on the loads). Loads in kW, T_s in °C and K.

	Five readings		Without row 2		Prior over posterior, without
Entry	MAP	sd	MAP	sd	[–]
g_1 [–]	1.7030	0.1067	1.7043	0.1061	2.83
g_2 [–]	1.3999	0.1370	1.6094	0.3000	1.00
Q_1 [kW]	30.69	1.452	30.92	1.483	∞
Q_2 [kW]	22.23	2.346	21.58	2.490	∞
T_s [°C]	12.111	0.2801	12.076	0.2837	3.52

21.3 The answers (step 4)

The six questions of Table 21.1 fall into two categories of Chapter 2 and take four of its general forms: estimation, as the distribution of a derived quantity (the first two) and as its variance split by entry and by row (the third and fourth); and prediction under a threshold, as a tail of the predictive distribution of a reading (the fifth) and as a tail of the posterior (the sixth). Each is answered from the posteriors of Section 21.2: its engineering question, the general form of its category, the form it takes for the plant, the quantities of the posteriors it uses, and the plant's numbers. The questions name the pinned heat-capacity rate $\dot{m}c_p = 10\,\text{kW/K}$ and the chiller's limit $T_{\max} = 18°\text{C}$. The hall-2 virtual reading is reported by the Laplace pushforward, because it depends nonlinearly on an entry that only its prior informs; the return temperature, which is linear in the state, by the delta method, exact under the Laplace posterior; the held-out check and the split of the precision between the reading and the virtual sensor by the refit, at $\hat{\mathbf{z}}_-$ with $\boldsymbol{\Sigma}_-$. Each reported value is the state route; where the full posterior gives a value, it is set beside the reported one in the same answer, with the gap.

21.3.1 The two virtual readings

Engineering questions. What is the temperature where there is no sensor, hall 2 with its sensor out? What is the return-header temperature, which has never been measured? Both ask how warm a place of the present plant is and how well the remaining rows know it.

General form. An estimation question about a derived quantity asks for its posterior distribution, equation (2.9),

$$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) \le f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

summarized by its moments (2.10) and its central credible interval (2.11).

Specialized form. The two questions are equations (2.9), (2.10) and (2.11) with the state $\mathbf{z}$ the reduced state (21.1); for hall 2 the derived quantity $f(\mathbf{z}) = T_2(\mathbf{z}) = T_s + Q_2\,\mathrm{e}^{-g_2}$, the evidence $\mathbf{y}_-$ and $p(\mathbf{z} \mid \mathbf{y}_-)$ the posterior of step 3 without the hall-2 row, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}_-, \boldsymbol{\Sigma}_-)$; for the return header $f(\mathbf{z}) = T_r(\mathbf{z})$ of equation (21.4), the evidence $\mathbf{y}$ and $p(\mathbf{z} \mid \mathbf{y})$ the five-reading posterior $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$. The reported routes, the Laplace pushforward for hall 2 and the delta method for the return header, follow from the forms below.

The gradients of T_2 and T_r, in the order of the state (21.1), are

$$\mathbf{g}_{T_2} = \big(0,\ -Q_2\,\mathrm{e}^{-g_2},\ 0,\ \mathrm{e}^{-g_2},\ 1\big)^{\mathsf{T}}, \qquad \mathbf{g}_{T_r} = \big(0,\ 0,\ 1/(\dot{m}c_p),\ 1/(\dot{m}c_p),\ 1\big)^{\mathsf{T}}, \tag{21.13}$$

the first evaluated at $\hat{\mathbf{z}}_-$, the second constant. The second entry of $\mathbf{g}_{T_2}$ is minus the temperature rise of hall 2: it is the lever through which a relative error of the conductance becomes an error of the temperature. By the delta method of Chapter 20, equation (20.15), the virtual reading of a quantity f and its standard deviation are

$$f \mid \mathbf{y} \approx \mathcal{N}(f(\hat{\mathbf{z}}),\ \mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}), \qquad \mathrm{sd}(f \mid \mathbf{y}) = \sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}, \tag{21.14}$$

with $(\hat{\mathbf{z}}_-, \boldsymbol{\Sigma}_-)$ in place of $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ for hall 2, and with the interval of equation (20.16). When f is an entry of the state, $f = z_i$, the gradient is the i-th unit vector and the variance is Σ_{ii}, the marginal of an unmeasured entry. For the return header the delta method is exact under the Laplace posterior, since T_r is linear in the state, and it is the reported route.

For hall 2 it is not. The quantity depends on g_2 through e^{-g_2}, and by the pattern (21.7) the column of g_2 holds only its prior row once the hall-2 row is removed. The posterior under $\mathbf{y}_-$ therefore factorizes as in equation (19.24): g_2 is independent of the other entries and its posterior is its prior $\mathcal{N}(\ln 5,\ 0.3^2)$, exactly and not only under the linearization, because the objective $\mathcal{L}_-$ of equation (21.6) is then a sum of the prior term $\frac{1}{2}((g_2 - \ln 5)/0.3)^2$ and terms that do not contain g_2, so that $\mathrm{e}^{-\mathcal{L}_-}$ is a product of a factor in g_2 alone and a factor in the other four entries. The first integral of the general form, under the Laplace posterior, is

$$\Pr(T_2 \le t \mid \mathbf{y}_-) = \int \mathbf{1}\big[T_2(\mathbf{z}) \le t\big]\, p(\mathbf{z} \mid \mathbf{y}_-)\, \mathrm{d}\mathbf{z}, \tag{21.15}$$

and conditioning on $g_2 = x$ reduces it to one dimension. Given x, $T_2 = T_s + Q_2\,\mathrm{e}^{-x}$ is linear in the normal pair (Q_2, T_s), so it is normal with mean $m(x) = \hat{T}_s + \hat{Q}_2\,\mathrm{e}^{-x}$ and variance $v(x) = \Sigma_{55} + \mathrm{e}^{-2x}\Sigma_{44} + 2\,\mathrm{e}^{-x}\Sigma_{45}$, the hats and the Σ_{ij} those of $\hat{\mathbf{z}}_-$ and $\boldsymbol{\Sigma}_-$, and by the law of total probability over x,

$$\Pr(T_2 \le t \mid \mathbf{y}_-) = \int_{-\infty}^{\infty} \Phi\!\left(\frac{t - m(x)}{\sqrt{v(x)}}\right) \mathcal{N}(x;\ \ln 5,\ 0.3^2)\, \mathrm{d}x, \tag{21.16}$$

the form of equation (20.8) with g_2 in place of g_1. Its derivative in t is the density of the virtual reading, its quantiles follow by solving $\Pr(T_2 \le t \mid \mathbf{y}_-) = q$, and its mean has the closed form of equation (20.9), $\hat{T}_s + \hat{Q}_2\,\mathrm{e}^{-\hat{g}_2 + \Sigma_{22}/2}$, and by the law of total variance its variance is the prior average of $v(x)$ plus the prior variance of $m(x)$, $\mathrm{Var}(T_2 \mid \mathbf{y}_-) = \mathbb{E}[v(g_2)] + \mathrm{Var}(m(g_2))$ under $g_2 \sim \mathcal{N}(\ln 5,\ 0.3^2)$. The size of the departure from the delta method is announced by the second-order term of equation (20.6), whose only nonzero second derivative is $\partial^2 T_2/\partial g_2^2 = Q_2\,\mathrm{e}^{-g_2}$:

$$\mathbb{E}[T_2 \mid \mathbf{y}_-] - T_2(\hat{\mathbf{z}}_-) \approx \tfrac{1}{2}\,\mathrm{tr}(\boldsymbol{\Sigma}_-\nabla^2 T_2) = \tfrac{1}{2}\,\Sigma_{22}\,\hat{Q}_2\,\mathrm{e}^{-\hat{g}_2}. \tag{21.17}$$

When this term is more than a few percent of the standard deviation, the delta interval is not reported without the pushforward, the rule of Chapter 20.

Building blocks. The MAP states $\hat{\mathbf{z}}$, $\hat{\mathbf{z}}_-$ and the covariances $\boldsymbol{\Sigma}$, $\boldsymbol{\Sigma}_-$ (Section 21.2.1, with the plant's values in Section 21.2.2); the gradients (21.13); the normal distribution function Φ and a one-dimensional quadrature for equation (21.16) and for the two prior averages of the variance, with a root search for its quantiles. For the values beside them, the reweighted draws of the full posterior (step 3).

The virtual reading of hall 2. At $\hat{\mathbf{z}}_-$ the conductance is $\mathrm{e}^{1.6094} = 5.000\,\mathrm{kW/K}$, the center of the delta method is

$$T_2(\hat{\mathbf{z}}_-) = 12.076 + 21.577/5.000 = 12.076 + 4.315 = 16.392^{\circ}\mathrm{C},$$

and the gradient (21.13) is $\mathbf{g}_{T_2} = (0,\ -4.315\,\mathrm{K},\ 0,\ 0.200\,\mathrm{K/kW},\ 1)^{\mathsf{T}}$. With $\Sigma_{22} = 0.09$, $\Sigma_{44} = 6.1985\,\mathrm{kW}^2$, $\Sigma_{55} = 0.08047\,\mathrm{K}^2$, $\Sigma_{45} = 0.01038\,\mathrm{kW\,K}$ and zero covariance of g_2 with the rest, the budget (20.20) is

$$\begin{aligned}\mathbf{g}_{T_2}^{\mathsf{T}}\boldsymbol{\Sigma}_{-}\mathbf{g}_{T_2} &= 4.315^2 \times 0.09 + 0.200^2 \times 6.1985 + 0.08047 + 2 \times 0.200 \times 1 \times 0.01038 \\ &= 1.6760 + 0.2479 + 0.0805 + 0.0042 = 2.0086\,\mathrm{K}^2,\end{aligned}$$

a delta standard deviation of 1.417 K. The second-order term (21.17) is $\frac{1}{2} \times 0.09 \times 4.315 = 0.194\,\mathrm{K}$, 14 % of that standard deviation, against 6 % for hall 1 at 60 kW in Chapter 20, so the distribution (21.15) is reported by the pushforward (21.16). Its quantiles at 2.5, 50 and 97.5 % are 14.26, 16.38 and 20.15°C; its mean is $12.076 + 21.577\,\mathrm{e}^{-1.6094+0.045} = 16.590$°C, the delta center plus 0.198 K, close to the second-order term; its variance is $2.30\,\mathrm{K}^2$, a standard deviation of 1.517 K. With the sensor in the fit the same quantity is 17.594°C with a standard deviation of 0.481 K, Chapter 18's fitted value of the hall-2 row. The virtual sensor puts hall 2 at a median of 16.38°C, with the 95 % band [14.26, 20.15]°C and a standard deviation of 1.52 K, three times the sensor's error. The value beside it comes from the full posterior: the draws of $\mathcal{N}(\hat{\mathbf{z}}_{-}, \boldsymbol{\Sigma}_{-})$ reweighted by equation (20.11), with an effective sample size of 191 123, give the median 16.397°C, the band [14.28, 20.17]°C and the mean 16.609°C, each within 0.02 K of the pushforward's. The gap is that small because the only entry that enters T_2 nonlinearly, g_2, has an exactly Gaussian posterior under $\mathbf{y}_{-}$.

The virtual reading of the return header. At $\hat{\mathbf{z}}$,

$$T_r(\hat{\mathbf{z}}) = 12.111 + \frac{30.688 + 22.231}{10} = 12.111 + 5.292 = 17.403°\mathrm{C},$$

with the constant gradient $\mathbf{g}_{T_r} = (0,\ 0,\ 0.1,\ 0.1,\ 1)^{\mathsf{T}}$, the entries for the loads in K/kW. With the covariance of Table 19.3, the budget (20.20) is

$$\begin{aligned}\mathbf{g}_{T_r}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}_{T_r} &= 0.01\,\Sigma_{33} + 0.01\,\Sigma_{44} + \Sigma_{55} + 2 \times 0.01\,\Sigma_{34} + 2 \times 0.1\,\Sigma_{35} + 2 \times 0.1\,\Sigma_{45} \\ &= 0.02109 + 0.05505 + 0.07843 - 0.03898 + 0.00058 - 0.00538 = 0.11079\,\mathrm{K}^2,\end{aligned}$$

a standard deviation of 0.333 K, and the central 95 % interval is $17.403 \pm 1.960 \times 0.333 = [16.75,\ 18.06]$°C. The value beside it comes from the full posterior, the reweighted draws of $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ (effective sample size 180 403): mean 17.427°C, standard deviation 0.332 K, band [16.77, 18.08]°C, the same spread and a center 0.023 K higher. In the coordinates $(T_s, Q_1 + Q_2)$, with the total load known to 1.93 kW as in Chapter 20, the same variance is 0.0784 from the supply, 0.0372 from the total load and $-0.0048\,\mathrm{K}^2$ from their covariance. Under $\mathbf{y}_{-}$ the same computation gives 17.33°C with a standard deviation of 0.35 K, the total load being then the heat meter's 52.50 kW against the five-reading fit's 52.92 kW.

21.3.2 The variance of a virtual reading by entry and by row

Engineering questions. Which entries carry each virtual reading's uncertainty? How much did the removed sensor know, and what does the virtual sensor lose? The first asks what each variance is made of, entry by entry of the state; the second what one row contributed, by comparing the posteriors with and without it.

General form. Both ask about the posterior variance of a derived quantity, equation (2.10),

$$\mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y}) = \int \big(f(\mathbf{z}) - \mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]\big)^2\, p(\mathbf{z} \mid \mathbf{y})\,\mathrm{d}\mathbf{z},$$

the second under two evidences, $\mathbf{y}$ and $\mathbf{y}_{-}$.

Specialized form. The two questions are the variance of equation (2.10) with the state $\mathbf{z}$ the reduced state, $f(\mathbf{z}) = T_2$ under $\mathbf{y}_-$ and $f(\mathbf{z}) = T_r$ under $\mathbf{y}$, and $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normals of step 3, $\mathcal{N}(\hat{\mathbf{z}}_-, \boldsymbol{\Sigma}_-)$ and $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$, under which the delta method gives the variance $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$; the second question compares the variance under $\mathbf{y}$ and under $\mathbf{y}_-$. Two choices enter here. The budget is read in the coordinates of the reduced state, $(g_1, g_2, Q_1, Q_2, T_s)$, at the linearization point of the evidence in use: its shares depend on the coordinates, as Chapter 20 warned, and Section 21.3.1 gives the return temperature's in the coordinates $(T_s, Q_1 + Q_2)$ as well. The split by row is derived in the linear model, rows linear near the MAP state: this makes it exact and gives the weights below, and it is broken to the extent that the hall-2 row's slope changes between $\hat{\mathbf{z}}$ and $\hat{\mathbf{z}}_-$. The split is reported at the refit's point, with V_- from $\boldsymbol{\Sigma}_-$ and the gradient at $\hat{\mathbf{z}}_-$; the same identities at $\hat{\mathbf{z}}$, through the leverage, are set beside it in Section 21.4.5.

By entry, the variance of the delta method is the budget (20.20), one diagonal term $[\mathbf{g}]_i^2\Sigma_{ii}$ per entry and one cross term $2[\mathbf{g}]_i[\mathbf{g}]_j\Sigma_{ij}$ per pair, or the attribution (20.24), $[\mathbf{g}]_i(\boldsymbol{\Sigma}\mathbf{g})_i$ per entry, with $\mathbf{g}$ of equation (21.13) and the covariance of the evidence in use. For the state itself the two posteriors are compared entry by entry through the ratio of prior to posterior spread, equation (19.25), read by the rule of Section 19.3.2: a ratio of one marks an entry that only its prior informs.

By row, on the scale of the quantity the reading measures, let reading k measure $f_k(\mathbf{z})$ with error σ_k, so that its normalized row is $r_k = (f_k(\mathbf{z}) - y_k)/\sigma_k$ with gradient $\mathbf{j}_k = \mathbf{g}/\sigma_k$, $\mathbf{g} = \nabla f_k$. The leverage of the row is the diagonal entry of the hat matrix $\mathbf{K} = \mathbf{J}\boldsymbol{\Sigma}\mathbf{J}^\mathsf{T}$ of Chapter 18,

$$K_{kk} = \mathbf{j}_k^\mathsf{T}\boldsymbol{\Sigma}\mathbf{j}_k = \frac{\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}}{\sigma_k^2}, \qquad 0 \le K_{kk} \le 1, \tag{21.18}$$

whose complement $1 - K_{kk}$ is the row's redundancy number, Section 18.3.2. The bounds hold because $\mathbf{K}$ is a symmetric projection: its diagonal entries are $K_{kk} = \sum_l K_{kl}^2 \ge K_{kk}^2$. The posterior with the row is the posterior without it conditioned on the row. Write $V_- = \mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}_-\mathbf{g}$ for the variance of the virtual sensor without the reading. In the linear model the rank-one update of equation (20.28), the conditioning of equation (4.63), applied to $\boldsymbol{\Sigma}_-$ with the row $\mathbf{j}_k$ and sandwiched between $\mathbf{g}^\mathsf{T}$ and $\mathbf{g}$ as in equation (20.29), gives

$$\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g} = V_- - \frac{\left(\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}_-\mathbf{j}_k\right)^2}{1 + \mathbf{j}_k^\mathsf{T}\boldsymbol{\Sigma}_-\mathbf{j}_k} = V_- - \frac{V_-^2/\sigma_k^2}{1 + V_-/\sigma_k^2} = \frac{V_-\,\sigma_k^2}{V_- + \sigma_k^2},$$

and, in reciprocals,

$$\frac{1}{\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}} = \frac{1}{V_-} + \frac{1}{\sigma_k^2}\,: \tag{21.19}$$

the reading and the virtual sensor of the same quantity add their precisions, like two independent readings of it. With equation (21.18), $K_{kk} = V_-/(V_- + \sigma_k^2)$, the reading's share of the precision, and $1 - K_{kk} = \sigma_k^2/(V_- + \sigma_k^2)$ the virtual sensor's; equivalently $V_- = \sigma_k^2 K_{kk}/(1 - K_{kk})$. The gain form of equation (4.63) moves the state by $\boldsymbol{\Sigma}_-\mathbf{g}\,(y_k - f_k(\hat{\mathbf{z}}_-))/(V_- + \sigma_k^2)$, so the quantity moves by $V_-/(V_- + \sigma_k^2) = K_{kk}$ times the innovation, and the fitted value is the weighted mean

$$f_k(\hat{\mathbf{z}}) = K_{kk}\,y_k + (1 - K_{kk})\,f_k(\hat{\mathbf{z}}_-). \tag{21.20}$$

The leverage K_{kk} is the weight the fit gives the reading, and the redundancy number $1 - K_{kk}$ the weight it gives the virtual sensor. A redundancy number near zero says that the rest of the plant knows almost nothing about the quantity the reading measures, and a number near one that the reading is nearly redundant with a good virtual sensor.

Building blocks. The gradients (21.13); the covariances $\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}_-$ (Section 21.2.1); V_- at $\hat{\mathbf{z}}_-$ and the sensor's σ_T; the prior spreads of Table 17.2.

The two budgets. Table 21.3 itemizes the variance of both virtual readings by equation (20.20), one row per term, with its value and its share of the total, and in the lower rows the per-entry attribution of equation (20.24). Figure 21.2 draws the attributions as bars, one panel per virtual sensor: the horizontal axis lists the five entries and the vertical axis is the share of the variance in percent.

Table 21.3: The variance budgets of the two virtual readings: hall 2 without its sensor ($\mathbf{g} = (0, -4.315, 0, 0.200, 1)^{\mathsf{T}}$ at $\hat{\mathbf{z}}_-$, with $\boldsymbol{\Sigma}_-$) and the return header ($\mathbf{g} = (0, 0, 0.1, 0.1, 1)^{\mathsf{T}}$, with $\boldsymbol{\Sigma}$). Upper rows: the terms of equation (20.20); lower rows: the attribution (20.24).

Term	Hall 2, virtual Value [K^2]	Share [%]	Return header Value [K^2]	Share [%]
$[\mathbf{g}]_2^2\Sigma_{22}$, conductance g_2	1.6760	83.4	0	0
$[\mathbf{g}]_3^2\Sigma_{33}$, load Q_1	0	0	0.0211	19.0
$[\mathbf{g}]_4^2\Sigma_{44}$, load Q_2	0.2479	12.3	0.0550	49.7
$[\mathbf{g}]_5^2\Sigma_{55}$, supply T_s	0.0805	4.0	0.0784	70.8
cross term Q_1, Q_2	0	0	−0.0390	−35.2
cross term Q_2, T_s	0.0042	0.2	−0.0054	−4.9
cross term Q_1, T_s	0	0	0.0006	0.5
total $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$	2.0086	sd 1.417 K	0.1108	sd 0.333 K
attribution to g_2, Q_1, Q_2, T_s	83.4, 0, 12.4, 4.1 %		0, 1.7, 29.7, 68.6 %	

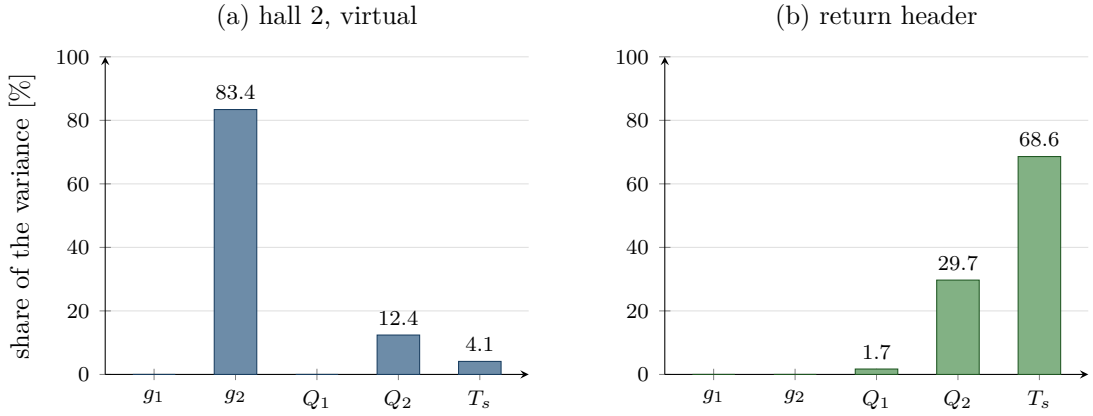


Figure 21.2: The variance budgets of the two virtual readings as the per-entry attribution $[\mathbf{g}]_i(\boldsymbol{\Sigma}\mathbf{g})_i/\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ of equation (20.24) (Table 21.3). (a) Hall 2 without its sensor, total 2.009 K^2. (b) The return header, total 0.111 K^2.

Hall 2's variance is 83 % the conductance term, the handbook's spread of 0.3 in the logarithm times the temperature rise of 4.315 K, 1.29 K of the delta standard deviation's 1.417 K; with g_2 uncorrelated with the rest, the attribution gives it the same 83 %. The return header's variance is, by attribution, 69 % the supply temperature and 30 % the load of hall 2; the loads' diagonal terms, 19 and 50 %, are offset by their cross term of −35 %, the correlation of −0.572 of Chapter 19, and Q_1 receives 2 % because $(\boldsymbol{\Sigma}\mathbf{g})_3 = 0.1\,\Sigma_{33} + 0.1\,\Sigma_{34} + \Sigma_{35} = 0.2109 - 0.1949 + 0.0029 = 0.0189$ K kW against $(\boldsymbol{\Sigma}\mathbf{g})_4 = 0.3287$ K kW. In the coordinates $(T_s, Q_1 + Q_2)$ the diagonal form gives 71 % to the supply, 34 % to the total load and −4 % to their covariance.

What the removed sensor knew. Table 21.2 answers entry by entry. With the sensor the spread of g_2 was 0.137, 2.19 times inside the prior; without it the ratio is exactly one, which by the rule of Section 19.3.2 marks an entry the readings do not inform. The spread of Q_2 grows

from 2.35 to 2.49 kW, those of Q_1 and T_s by 1 to 2 %, and g_1 barely changes.

On hall 2's temperature, at the refit's point, the virtual sensor has the variance $V_- = 2.009\,\mathrm{K}^2$ of Section 21.3.1, a precision of $0.498\,\mathrm{K}^{-2}$, against the sensor's $1/\sigma_T^2 = 4\,\mathrm{K}^{-2}$. By the precision split (21.19) their sum, $4.498\,\mathrm{K}^{-2}$, is the precision with the reading restored, a standard deviation of 0.472 K, and the sensor's share is $V_-/(V_- + \sigma_T^2) = 2.009/2.259 = 0.889$, the virtual sensor's 0.111. The weighted mean (21.20) with these weights, $0.889 \times 17.700 + 0.111 \times 16.392 = 17.555$°C, lies 0.04 K from the five-reading fit's 17.594°C, which is linearized at $\hat{\mathbf{z}}$ instead (Section 21.4.5). The hall-2 sensor carried nine tenths of the precision on hall 2's temperature and all that is known of hall 2's conductance beyond the handbook.

21.3.3 The held-out reading against its predictive

Engineering question. Does the sensor's last reading agree with the rest of the plant? The reading $y_{T_2} = 17.70$°C is held out of the fit, and the question is whether it lies where the other rows predict a reading of hall 2 should lie.

General form. A future or held-out observation is judged against its predictive distribution, equation (2.13),

$$p(y_{\text{new}} \mid \mathbf{y}) = \int p(y_{\text{new}} \mid \mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

and the question is a tail of it.

Specialized form. The question is a tail of equation (2.13) with $y_{\text{new}} = y_{T_2}$, the held-out hall-2 reading, $p(y_{\text{new}} \mid \mathbf{z}) = \mathcal{N}(T_2(\mathbf{z}),\ \sigma_T^2)$, the evidence $\mathbf{y}_-$ and $p(\mathbf{z} \mid \mathbf{y}_-)$ the posterior of step 3 without the row, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}_-, \boldsymbol{\Sigma}_-)$. One modeling choice enters here. The held-out reading is taken to be of the plant in the condition the other readings describe, with a normal error independent of the state and of the other errors: this is what makes its predictive the virtual variance plus the sensor's own, and it fails if the reading was taken before a change of the plant that the other readings have since followed (Section 21.6.6). The predictive is reported by the delta method at $\hat{\mathbf{z}}_-$; the linear identity at $\hat{\mathbf{z}}$ and the pushforward are set beside it in Section 21.4.5.

Here y_{new} is the reading $y_k = f_k(\mathbf{z}) + \varepsilon_k$, with $f_k = T_2$ and $\varepsilon_k \sim \mathcal{N}(0, \sigma_k^2)$. Its predictive is that of Chapter 4, equation (4.61), under $\mathbf{y}_-$. By the law of total variance of Chapter 3, equation (3.31), conditioning on the state,

$$\mathrm{Var}(y_k \mid \mathbf{y}_-) = \mathbb{E}\big[\mathrm{Var}(y_k \mid \mathbf{z}) \mid \mathbf{y}_-\big] + \mathrm{Var}\big(\mathbb{E}[y_k \mid \mathbf{z}] \mid \mathbf{y}_-\big) = \sigma_k^2 + \mathrm{Var}\big(f_k(\mathbf{z}) \mid \mathbf{y}_-\big),$$

and under the delta method the second term is $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}_-\mathbf{g}$, with $\mathbf{g} = \nabla f_k(\hat{\mathbf{z}}_-)$. The predictive distribution of the reading is therefore

$$y_k \mid \mathbf{y}_- \approx \mathcal{N}\big(f_k(\hat{\mathbf{z}}_-),\ \sigma_k^2 + \mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}_-\mathbf{g}\big), \qquad \mathrm{sd}(y_k \mid \mathbf{y}_-) = \sqrt{\sigma_k^2 + \mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}_-\mathbf{g}}, \tag{21.21}$$

the reading's own error variance added to the virtual sensor's. The check of the reading against the rest of the plant is its standardized departure and the tail beyond it,

$$d_k = \frac{y_k - f_k(\hat{\mathbf{z}}_-)}{\mathrm{sd}(y_k \mid \mathbf{y}_-)}, \qquad \Pr(Y_k \geq y_k \mid \mathbf{y}_-) \approx 1 - \Phi(d_k), \tag{21.22}$$

with Y_k the reading as a random variable; d_k is standard normal if the model and every row, the held-out one included, are correct. A reading with $|d_k|$ near three or beyond disagrees with what the other rows predict; a reading with $|d_k|$ below one agrees with them, and the agreement is weak evidence when $\mathrm{sd}(y_k \mid \mathbf{y}_-)$ is large. A bias b of the sensor adds b to the reading and $b/\mathrm{sd}(y_k \mid \mathbf{y}_-)$ to the mean of d_k. Flagged at $d_k \geq 3$, the bias is caught with probability $\Phi(b/\mathrm{sd}(y_k \mid \mathbf{y}_-) - 3)$: half the time at $b = 3\,\mathrm{sd}(y_k \mid \mathbf{y}_-)$ and 80 % of the time at $b = (3 + 0.84)\,\mathrm{sd}(y_k \mid \mathbf{y}_-)$, 0.84 being the 80 % quantile of the standard normal.

The check is the normalized residual. In the linear model the departure has a closed form in the objects of the full fit. With $V_- = \mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}_-\mathbf{g} = \sigma_k^2 K_{kk}/(1-K_{kk})$ from equation (21.18) and the precision split (21.19), the predictive variance in units of the reading's standard deviation is

$$\frac{\sigma_k^2 + V_-}{\sigma_k^2} = 1 + \frac{K_{kk}}{1-K_{kk}} = \frac{1}{1-K_{kk}}.$$

Subtracting y_k from both sides of the weighted mean (21.20) gives $f_k(\hat{\mathbf{z}}) - y_k = (1-K_{kk})\,(f_k(\hat{\mathbf{z}}_-) - y_k)$, and dividing by σ_k gives the residual of the reading against the fit that has not seen it,

$$r_k(\hat{\mathbf{z}}_-) = \frac{r_k(\hat{\mathbf{z}})}{1-K_{kk}}\,: \tag{21.23}$$

the leave-one-out residual (linear rows) is the residual at the full fit divided by the redundancy number. A reading whose redundancy number is a few hundredths is fitted with a residual tens of times smaller than its departure from what the rest of the plant predicts. The departure in units of σ_k is $-r_k(\hat{\mathbf{z}}_-)$, and dividing by the predictive standard deviation in the same units,

$$d_k = -\,\frac{r_k(\hat{\mathbf{z}})/(1-K_{kk})}{1/\sqrt{1-K_{kk}}} = -\,\frac{r_k(\hat{\mathbf{z}})}{\sqrt{1-K_{kk}}} = -\,r_k^{\mathrm{N}}, \tag{21.24}$$

the normalized residual of equation (18.25) with its sign reversed (the residual is predicted minus observed, the departure observed minus predicted). The two checks are one: comparing a reading with a virtual sensor built from the other rows is the bad-data statistic of state estimation, and the redundancy number that scales it, through $\mathrm{sd}(y_k \mid \mathbf{y}_-) = \sigma_k/\sqrt{1-K_{kk}}$, says before any reading is taken how large a bias the comparison can reveal.

Building blocks. $\hat{\mathbf{z}}_-$, $\boldsymbol{\Sigma}_-$ and $\mathbf{g}_{T_2}$ at $\hat{\mathbf{z}}_-$; σ_T; Φ.

The held-out reading. At the refit, by equation (21.21), the predictive standard deviation of the hall-2 reading is $\sqrt{0.5^2 + 1.417^2} = \sqrt{2.259} = 1.503\,\mathrm{K}$, and by equation (21.22)

$$d_2 = \frac{17.700 - 16.392}{1.503} = \frac{1.308}{1.503} = 0.87,$$

a one-sided tail of 0.192 and a two-sided one of 0.38: the reading and the other rows agree. Flagged at three predictive standard deviations, a bias of the hall-2 sensor is caught half the time at $3 \times 1.503 = 4.5\,\mathrm{K}$ and 80 % of the time at $3.84 \times 1.503 = 5.8\,\mathrm{K}$; a bias of 3 K, six times the sensor's error, would have put this reading at $(1.308 + 3)/1.503 = 2.87$ predictive standard deviations, inside the threshold. The cause is the virtual sensor's 11 % share of the precision: the rest of the plant checks this row only through the handbook's prior on g_2. The chapter computes no full-posterior value of this tail.

21.3.4 The return temperature at the chiller's limit

Engineering question. Is the return temperature below the chiller's limit of 18°C?

General form. A prediction under a threshold is the posterior probability that a requirement holds, compared with an assurance level, equation (2.12),

$$\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y}) = \int \mathbf{1}\,[f(\mathbf{z}) \geq f^*]\; p(\mathbf{z} \mid \mathbf{y})\,\mathrm{d}\mathbf{z}, \qquad \text{accept if } \Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y}) \geq \kappa.$$

Specialized form. The question is equation (2.12) with $f(\mathbf{z}) = -T_r(\mathbf{z})$, $f^* = -T_{\max} = -18°\mathrm{C}$, no assurance level κ, and $p(\mathbf{z} \mid \mathbf{y})$ the five-reading posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$. One choice enters here: the limit is compared with the steady return temperature of the loop balance (21.3). A load step, which warms the return water over the loop's transit and the halls' thermal time constants, would need the transient balance; Section 21.6.6 returns to it. An upper limit $T_{\max}$ on a temperature is the requirement $-T_r \geq -T_{\max}$, so $f = -T_r$, $f^* = -T_{\max}$, and the probability is

$$\Pr(T_r < T_{\max} \mid \mathbf{y}) = \int \mathbf{1}\big[T_r(\mathbf{z}) < T_{\max}\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}. \tag{21.25}$$

The chiller states no assurance level κ, so the probability itself is the answer. Since T_r is linear, its Laplace pushforward is the normal of equation (21.14), and the tail is that of equation (20.18),

$$\Pr(T_r < T_{\max} \mid \mathbf{y}) \approx \Phi\left(\frac{T_{\max} - T_r(\hat{\mathbf{z}})}{\sqrt{\mathbf{g}_{T_r}^{\mathsf{T}} \boldsymbol{\Sigma} \mathbf{g}_{T_r}}}\right),$$

exact under the Laplace posterior; the only approximation left is the Gaussian posterior itself, and the reweighted draws of equation (20.11) give the value beside it.

Building blocks. $T_r(\hat{\mathbf{z}})$ and $\mathbf{g}_{T_r}^{\mathsf{T}} \boldsymbol{\Sigma} \mathbf{g}_{T_r}$ from Section 21.3.1; Φ. For the value beside it, the reweighted draws of the full posterior.

The probability of staying below the chiller's limit. The probability (21.25), by the normal tail of Section 21.3.4, has the standardized distance to the limit $(18 - 17.4033)/0.33285 = 1.793$, and

$$\Pr(T_r < 18°\mathrm{C} \mid \mathbf{y}) \approx \Phi(1.793) = 0.963. \tag{21.26}$$

The probability that the return water exceeds the chiller's limit is 0.037. The value beside it comes from the full posterior: the reweighted draws give $\Pr(T_r < 18°\mathrm{C} \mid \mathbf{y}) = 0.958 \pm 0.0005$, an exceedance of 0.042, a gap of 0.005 in probability. The state route errs in that direction because the full posterior's center lies 0.023 K higher, the supply temperature being pulled up by the right skew of both conductances, while its spread is the same.

21.4 The checks (step 5)

Each reported value has beside it a route that does not share its approximation, and the hall-2 answer has behind it a prior that no remaining reading checks; both are tested here.

21.4.1 The downdate against the refit

The downdate (21.9) and the step (21.12), computed from the five-reading fit alone, should reproduce the refit. The hall-2 row (21.5) at $\hat{\mathbf{z}}$ has the normalized gradient

$$\mathbf{j}_2 = \frac{1}{\sigma_T}\left(0,\ -Q_2 \mathrm{e}^{-g_2},\ 0,\ \mathrm{e}^{-g_2},\ 1\right)^{\mathsf{T}} = (0,\ -10.965,\ 0,\ 0.493,\ 2)^{\mathsf{T}},$$

in the inverse units of the entries, and with the covariance of Table 19.3,

$$\boldsymbol{\Sigma}\mathbf{j}_2 = (-0.0004,\ -0.0722,\ -0.0808,\ 0.2252,\ 0.0122)^{\mathsf{T}}, \qquad K_{22} = \mathbf{j}_2^{\mathsf{T}} \boldsymbol{\Sigma} \mathbf{j}_2 = 0.9269,$$

so the redundancy number is 0.0731, the value of Table 18.5; the row is not a critical measurement, since the prior keeps g_2 proper. The downdated variance of g_2 is

$$\Sigma_{22} + \frac{(\boldsymbol{\Sigma}\mathbf{j}_2)_2^2}{1 - K_{22}} = 0.018765 + \frac{0.0722^2}{0.0731} = 0.018765 + 0.0712 = 0.0900,$$

the prior variance 0.3^2, and with the residual $r_2(\hat{\mathbf{z}}) = -0.212$ of Table 18.5 the step (21.12) moves g_2 by $-0.0722 \times (-0.212)/0.0731 = +0.209$, from 1.3999 to $1.6094 = \ln 5$, the prior mean. The other entries move by the same factor $r_2(\hat{\mathbf{z}})/(1 - K_{22}) = -2.90$ times their component of $\boldsymbol{\Sigma}\mathbf{j}_2$: Q_2 falls by 0.65 kW, Q_1 rises by 0.23 kW, T_s falls by 0.035 K and g_1 rises by 0.0012. Table 21.4 sets the result beside the refit: one row per entry, the MAP value and standard deviation by the downdate and by the refit.

Table 21.4: The fit without the hall-2 row by the downdate from the five-reading fit, $\tilde{\boldsymbol{\Sigma}}_-$ of equation (21.9) and the step (21.12), and by the refit, $\boldsymbol{\Sigma}_-$ (Table 21.2). Loads in kW, T_s in °C and K.

	Downdate		Refit	
Entry	MAP	sd	MAP	sd
g_1 [–]	1.7042	0.1067	1.7043	0.1061
g_2 [–]	1.6094	0.3000	1.6094	0.3000
Q_1 [kW]	30.92	1.483	30.92	1.483
Q_2 [kW]	21.58	2.490	21.58	2.490
T_s [°C]	12.076	0.2836	12.076	0.2837

The downdate reproduces the refit to the digits of the table for g_2, Q_1 and Q_2. It differs for g_1, by 0.0001 in the value and by 0.0006 in the spread (0.1067 against the refit's 0.1061), and by one unit in the last digit of the spread of T_s. The difference is the nonlinearity of the hall-1 row: its sensitivity to g_1 is $Q_1 e^{-g_1}/\sigma_T$, which rises from 11.18 at $\hat{\mathbf{z}}$ to 11.25 at $\hat{\mathbf{z}}_-$ because the refit raises Q_1, and the refit evaluates the information at the new state while the downdate holds it at $\hat{\mathbf{z}}$. Inequality (21.11) holds at a fixed linearization point, and the downdate column satisfies it entry by entry; a refit can narrow an entry slightly by moving the point, as it does for g_1.

21.4.2 The limit of exact balances and the dispersion of the refit

The reduced state is the limit $\sigma_p \to 0$ of the full state of Chapter 17, in which the hall temperature T_2 is the seventh entry. The fit of the full state without the hall-2 reading row, with its three physics rows at the tolerance σ_p, gives hall 2's entry the center 16.3916°C at every tolerance tried and the spread $\sqrt{\Sigma_{77}} = 1.4209$ K at $\sigma_p = 0.1$ K, 1.4173 K at 0.01 K and 1.4172 K at 0.001 K, the reduced $\sqrt{\mathbf{g}_{T_2}^{\mathsf{T}} \boldsymbol{\Sigma}_- \mathbf{g}_{T_2}}$, with g_2's spread at 0.3000 throughout. The reason is the structure of the rows: with the reading removed, the entry T_2 appears only in its own balance row, so that row is satisfied exactly at the MAP state for every tolerance, which leaves the center unchanged, and it adds its own σ_p^2 to the variance, $\sqrt{2.0086 + \sigma_p^2} = 1.4208$, 1.4173 and 1.4172 K, the hall-1 and header balances accounting for the last digit at $\sigma_p = 0.1$ K. The virtual reading of an unmeasured function of the reduced state is the marginal of an unmeasured entry of the full state, as Section 21.3.1 states. The return header is not a node of the full state, and no tolerance enters its quantity. The refit itself shows no dispersion beyond what its rows allow: $2\mathcal{L}_-(\hat{\mathbf{z}}_-) = 0.123$ against two degrees of freedom, a tail probability of 0.94.

21.4.3 The shape of the hall-2 virtual distribution

The delta interval of equation (20.16) for hall 2 is $16.392 \pm 1.960 \times 1.417 = [13.61,\ 19.17]$°C, against the pushforward's [14.26, 20.15]°C and standard deviation 1.517 K. $N = 200\,000$ draws of $\mathcal{N}(\hat{\mathbf{z}}_-, \boldsymbol{\Sigma}_-)$, seeded with 21, give mean 16.595°C, standard deviation 1.519 K, skewness 0.96 and the band [14.27, 20.16]°C, the pushforward to the precision of the draws; the same draws reweighted to the full reduced posterior by equation (20.11), with an effective sample size of 191 123, give [14.28, 20.17]°C, mean 16.609°C and median 16.397°C. The Laplace and full

posteriors agree, as the structure predicts: the only entry that enters T_2 nonlinearly, g_2, has an exactly Gaussian posterior.

Figure 21.3 draws the result. The horizontal axis is hall 2's temperature and the vertical axis its density in K^{-1}. The solid blue curve is the virtual sensor by the delta method, the dashed red curve the virtual sensor by the Laplace pushforward, and the narrow green curve the fitted value with the sensor in the fit; the vertical gray line marks the reading.

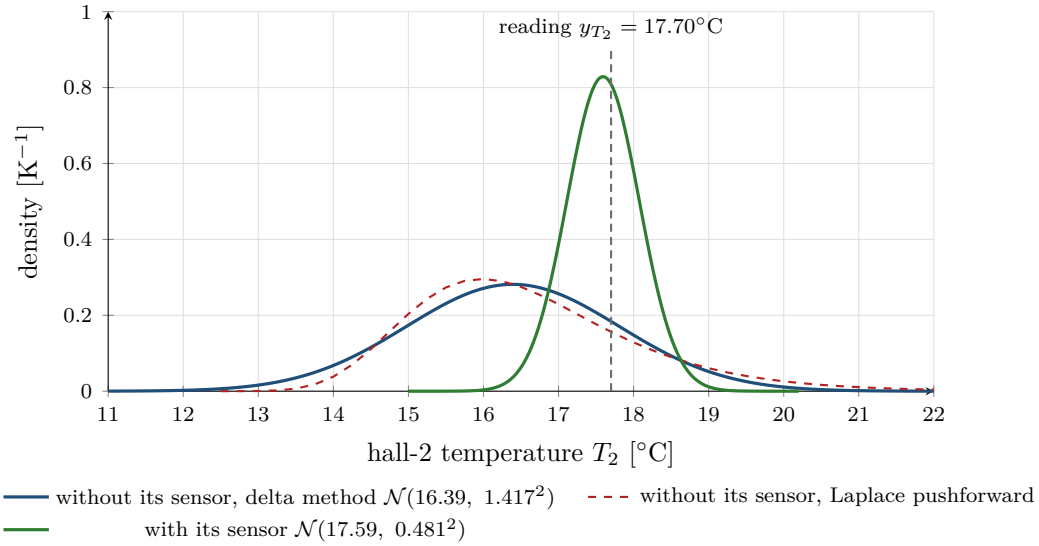


Figure 21.3: Hall 2's temperature with and without its sensor. Blue: the virtual sensor by the delta method, $\mathcal{N}(16.392, 1.417^2)$, equation (21.14) under $\mathbf{y}_{-}$. Red dashed: the virtual sensor by the Laplace pushforward, the one-dimensional integral (21.16) over $g_2 \sim \mathcal{N}(\ln 5, 0.3^2)$, evaluated on a grid of 0.25 K; the reweighted draws of the full reduced posterior coincide with it to the width of the line. Green: the fitted value with the sensor, $\mathcal{N}(17.594, 0.481^2)$. Gray: the reading.

The virtual distribution is wide and skewed to the right: its upper tail is longer, because a conductance below the handbook's raises the temperature more than a conductance above it lowers it. The delta interval is shifted 0.6 to 1.0 K to the left of the pushforward's, and it puts the probability that hall 2 is warmer than the reading at 0.178, against 0.204 by the pushforward and 0.207 under the full posterior. These are posterior probabilities of the quantity T_2, not of the reading y_{T_2}, which carries the sensor's error as well and was checked against its own predictive in Section 21.3.3. The reading and the fitted value both lie inside the virtual 95 % band, the reading at about its 80th percentile.

21.4.4 The return temperature against the full posterior

Because T_r is linear in the state, the Laplace pushforward is the delta method, and draws of the Laplace Gaussian can only reproduce it: 200 000 draws of $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$, seeded with 21, give $\Pr(T_r < 18°\mathrm{C}) = 0.9639 \pm 0.0004$. The only check with content is against the full posterior, and the same draws reweighted by equation (20.11), with log weights spanning 9.5 and an effective sample size of 180 403, give

$$\Pr(T_r < 18°\mathrm{C} \mid \mathbf{y}) = 0.958 \pm 0.0005,$$

with mean 17.427°C, standard deviation 0.332 K, skewness 0.00 and the band [16.77, 18.08]°C; draws from a multivariate Student t with five degrees of freedom and 1.5 times the Laplace spread give 0.958 again. The full posterior has the Laplace spread and no skewness, and its center lies 0.023 K higher. The shift comes from the supply temperature: its full-posterior

mean is 12.135°C against the MAP value 12.111, because the posteriors of both conductances are skewed toward large values, as Chapter 19 found for g_1, and T_s is positively correlated with both (0.43 and 0.31). A linear output inherits the posterior's shape unaltered, so its only approximation is the Gaussian posterior itself.

Table 21.5 gives the probability of staying below limits from 17.25 to 18.5°C by the three routes, each limit labeled with its standardized distance $(T_{\max} - 17.403)/0.333$; one row of the arithmetic, for 17.75°C: $0.347/0.333 = 1.041$ and $\Phi(1.041) = 0.851$. Figure 21.4 draws the densities. The horizontal axis is the return temperature and the vertical axis its density in K^{-1}; the solid blue curve is the delta-method normal, the green dotted curve the full posterior, a kernel estimate of the reweighted draws, and the shaded area to the right of the red line at 18°C is the probability of exceeding the limit under the normal.

Table 21.5: The probability that the return temperature stays below a limit, by the delta method (equal to the Laplace pushforward, since T_r is linear), by 200 000 draws of the Laplace Gaussian (seed 21) and by the same draws reweighted to the full posterior ($N_{\text{eff}} = 180\,403$; standard errors at most 0.0012). In brackets, the standardized distance of each limit from the center.

Probability [–]	Delta method	Laplace draws	Full posterior
$\Pr(T_r < 17.25°\mathrm{C})$ [−0.46]	0.323	0.323	0.298
$\Pr(T_r < 17.50°\mathrm{C})$ [0.29]	0.614	0.614	0.587
$\Pr(T_r < 17.75°\mathrm{C})$ [1.04]	0.851	0.851	0.834
$\Pr(T_r < 18.00°\mathrm{C})$ [1.79]	0.963	0.964	0.958
$\Pr(T_r < 18.25°\mathrm{C})$ [2.54]	0.9945	0.9944	0.9933
$\Pr(T_r < 18.50°\mathrm{C})$ [3.29]	0.9995	0.9996	0.9994

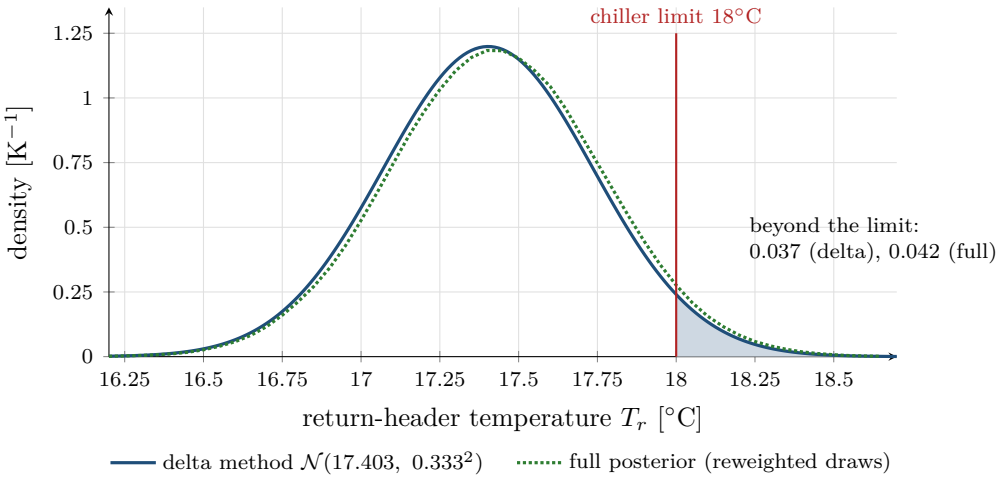


Figure 21.4: Density of the return-header temperature. Blue: the normal of the delta method, $\mathcal{N}(17.403, 0.333^2)$, exact under the Laplace posterior because T_r is linear in the state; shaded, its mass beyond the limit, 0.037. Green dotted: the full posterior, a Gaussian kernel estimate (bandwidth 0.05 K) of 200 000 draws reweighted by equation (20.11), whose mass beyond the limit is 0.042. Red: the chiller's limit on its return water.

The two curves have the same width and shape, and the green one lies a little to the right. Near the center the delta method overstates the probability of staying below a limit by up to 0.027; at the chiller's limit the error is +0.005, and beyond 2.5 standard deviations it is about 0.001. The delta answer (21.26) is adequate to two digits: the return water stays below 18°C with probability 0.96, and exceeds it with probability about 0.04, which the full posterior puts

at 0.042.

Table 21.6 collects the virtual readings of the chapter with the fitted hall-2 value for comparison, by the three routes: the value at the MAP state and its delta standard deviation and interval, the Laplace pushforward's interval and the full posterior's interval, with the medians of the last two.

Table 21.6: The virtual readings and the fitted hall-2 value: value at the MAP state, delta-method standard deviation and 95 % interval, and the 95 % intervals (median in brackets) of the Laplace pushforward and of the full posterior, from 200 000 draws (for the virtual hall-2 reading the pushforward is also computed by the one-dimensional integral (21.16), with the same interval).

Quantity [°C]	$f(\hat{\mathbf{z}})$	$\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}$ [K]	Delta interval	Laplace pushforward	Full posterior
T_2, virtual ($\mathbf{y}_-$)	16.392	1.417	[13.61, 19.17]	[14.26, 20.15] (16.38)	[14.28, 20.17] (16.40)
T_2, fitted ($\mathbf{y}$)	17.594	0.481	[16.65, 18.54]	[16.70, 18.60] (17.57)	[16.61, 18.51] (17.56)
T_r, virtual ($\mathbf{y}$)	17.403	0.333	[16.75, 18.06]	[16.75, 18.06] (17.40)	[16.77, 18.08] (17.43)
T_r, virtual ($\mathbf{y}_-$)	17.326	0.347	[16.65, 18.01]	equal to delta	–

21.4.5 The held-out check and the precision split by the other routes

Table 21.7 sets the check of Section 21.3.3 out by three routes. The columns are the linear identity at the full fit, equations (21.23) and (21.24); the refit, which is the reported value; and the Laplace pushforward of the reduced posterior through the hall balance, the integral (21.16) with the sensor's variance added to $v(x)$. The rows give the virtual reading, its variance, the predictive standard deviation, the departure, the standardized departure and the one-sided probability of a reading at least as high.

Table 21.7: The held-out check of the hall-2 sensor: the reading $y_{T_2} = 17.70$°C against its predictive distribution given the other evidence, equation (21.21), by the linear identity at the full fit, by the refit without the row and by the Laplace pushforward of the reduced posterior.

Quantity	Linear identity at $\hat{\mathbf{z}}$	Refit at $\hat{\mathbf{z}}_-$	Laplace pushforward
virtual reading [°C]	16.249	16.392	median 16.375, mean 16.590
virtual variance [K^2]	$\sigma_T^2 K_{22}/(1-K_{22}) = 3.168$	2.009	2.30
predictive sd [K]	$\sigma_T/\sqrt{1-K_{22}} = 1.849$	1.503	1.60
departure $y_{T_2} - T_2$ [K]	1.451	1.308	–
standardized departure d_2 [–]	$-r_2^{\mathrm{N}} = 0.78$	0.87	0.79 (equivalent)
$\Pr(Y_{T_2} \geq 17.70 \mid \mathbf{y}_-)$ [–]	0.216	0.192	0.216

The first column is the normalized residual of the hall-2 row in Table 18.5, −0.78, with its sign reversed, as equation (21.24) requires: the held-out check and the normalized residual are the same statistic in the linear model. The refit differs from it by 0.09 because the predictive standard deviation depends on the temperature rise at the point of evaluation, 5.48 K at $\hat{\mathbf{z}}$ and 4.32 K at $\hat{\mathbf{z}}_-$, the rise being the lever of the conductance's spread; the pushforward, which keeps

the skewness of the virtual distribution, gives a tail of 0.216, the probability of a standard normal beyond 0.79. All three say the same: the reading agrees with the rest of the plant. The linear identity's wider predictive moves the bias flagged half the time from 4.5 to $3 \times 1.849 = 5.5\,\mathrm{K}$, and the bias flagged 80 % of the time from 5.8 to $3.84 \times 1.849 = 7.1\,\mathrm{K}$.

The same point moves the precision split of Section 21.3.2. At $\hat{\mathbf{z}}$, by equation (21.18) with the leverage of Section 21.4.1, the virtual variance is $V_- = \sigma_T^2 K_{22}/(1-K_{22}) = 0.25 \times 0.9269/0.0731 = 3.168\,\mathrm{K}^2$, a precision of $0.316\,\mathrm{K}^{-2}$; with the sensor's $4\,\mathrm{K}^{-2}$ the sum, $4.316\,\mathrm{K}^{-2}$, is the reciprocal of $\sigma_T^2 K_{22} = 0.2317\,\mathrm{K}^2$, the square of the fitted value's 0.481 K, as the precision split (21.19) requires. The shares are 93 % for the sensor and 7 % for the rest of the plant, against 89 and 11 % at $\hat{\mathbf{z}}_-$, and at this point the weighted mean (21.20) holds exactly: with the linearized virtual reading $17.594 + \sigma_T\,\mathbf{j}_2^{\mathsf{T}}(\hat{\mathbf{z}}_- - \hat{\mathbf{z}}) = 16.249°\mathrm{C}$,

$$0.9269 \times 17.700 + 0.0731 \times 16.249 = 16.406 + 1.188 = 17.594°\mathrm{C}.$$

Either way the sensor carried about nine tenths of the precision.

21.4.6 Sensitivity to the handbook prior

Hall 2's virtual reading rests on the handbook's conductance, and the check that matters most is what a different handbook would have said. Table 21.8 repeats the two virtual readings under five handbook priors, the same prior on g_1 and g_2 in each row. The first column is the handbook conductance G^0 and its spread σ_g in the logarithm; the next three give the hall-2 virtual reading under $\mathbf{y}_-$, its pushforward median, 95 % band and standard deviation; the last two the return temperature under $\mathbf{y}$, its value with standard deviation and the probability of staying below 18°C. The first row is the handbook of Table 17.2.

Table 21.8: The two virtual readings under five handbook priors on the conductances, $g_i \sim \mathcal{N}(\ln G^0,\ \sigma_g^2)$ for both halls. Hall 2 under $\mathbf{y}_-$ by the pushforward (21.16), median, interval and standard deviation; the return header under $\mathbf{y}$ by the delta method.

	Hall 2, virtual			Return header	
Handbook G^0 [kW/K], σ_g [–]	Median [°C]	95 % band [°C]	sd [K]	T_r [°C]	$\Pr(T_r < 18)$ [–]
5.0, 0.30 (reported)	16.38	[14.26, 20.15]	1.517	17.403 ± 0.333	0.963
5.0, 0.15	16.38	[14.88, 18.35]	0.885	17.486 ± 0.305	0.954
5.0, 0.50	16.37	[13.57, 23.80]	2.694	17.368 ± 0.343	0.967
4.0, 0.30	17.46	[14.84, 22.19]	1.895	17.307 ± 0.333	0.981
6.0, 0.30	15.67	[13.88, 18.81]	1.268	17.483 ± 0.333	0.940

The hall-2 virtual reading follows the handbook. Its median moves with the handbook conductance, from 15.67°C at 6 kW/K to 17.46°C at 4 kW/K, and its band with the handbook's spread, from [14.88, 18.35] at 0.15 to [13.57, 23.80]°C at 0.5, whose upper end lies above the reading by six kelvin. No alternative handbook can be told from the reported one by the remaining readings, which do not see g_2. The readings of Chapter 17 were generated with $G_2 = 4.0\,\mathrm{kW/K}$ and $Q_2 = 24\,\mathrm{kW}$, a hall at 18.0°C; the reported band contains it, and the handbook at the true conductance would have put the median at 17.46°C. The return temperature depends on the handbook only through the correlation of the supply temperature with the conductances: its probability of staying below the limit moves between 0.940 and 0.981, its standard deviation between 0.305 and 0.343 K. The return answer is a property of the readings; the hall-2 answer is a property of the handbook.

21.5 The report (step 6)

The shift log's entry for hall 2 does not hold. With its sensor out, hall 2 is known not to the sensor's 0.5 K but to a 95 % band from 14.3 to 20.2°C, a standard deviation of 1.52 K, three times the sensor's error, because nothing but the handbook's conductance stands behind it; and the last reading of 17.70°C, which lies inside that band, is not thereby confirmed, since the rest of the plant would flag a bias of that sensor only half the time at 4.5 K and 80 % of the time at 5.8 K. The return header, which has never been measured, is known to 0.33 K, better than the hall sensors read their halls. The answers below are given in the units in which the questions were asked.

21.5.1 The answers and their bands

Table 21.9 gives the answers to the six questions of Table 21.1, in their order: one row per question, with the answer from the posterior of the state, the full-posterior value beside it where the chapter computes one, and its band with the check or comparison behind it. Bands are central 95 % posterior intervals unless stated.

Table 21.9: The answers to the questions of the two-hall plant with the hall-2 sensor out of service: answers from the posterior of the state, with the full-posterior value (reweighted draws) beside each where one is computed. Bands are central 95 % posterior intervals.

Question	Answer (state route)	Full posterior	Band, check and comparison
The temperature where there is no sensor (hall 2, sensor out)	16.4°C (median of the pushforward; standard deviation 1.52 K)	median 16.40°C	[14.3, 20.2]°C, skewed (full posterior the same); the delta band [13.6, 19.2] is not reported; with the sensor 17.59 ± 0.48 K
The return-header temperature	17.40°C, standard deviation 0.33 K	17.43°C, 0.33 K	[16.75, 18.06]°C; full posterior [16.77, 18.08], center 0.02 K higher
What carries each variance	hall 2: the conductance prior; return: the supply temperature and hall 2's load	–	83 % from g_2; 69 % from T_s, 30 % from Q_2, 2 % from Q_1
What the removed sensor knew	nine tenths of the precision on T_2; all of what is known of g_2 beyond the handbook	–	g_2 back to its prior (0.300); Q_2 2.35 → 2.49 kW; precision shares 89/11 % (sensor/rest of the plant)
Does the last reading agree with the rest of the plant?	yes, 0.87 predictive standard deviations above the prediction	–	one-sided tail 0.19; a bias is flagged half the time at 4.5 K, 80 % of the time at 5.8 K
Return temperature against the chiller's limit of 18°C	$\Pr(T_r < 18°\text{C} \mid \mathbf{y}) =$ 0.963, about 0.96	0.958	exceedance 0.037, about 0.04 (0.042); 0.94 to 0.98 over the handbooks of Table 21.8

21.5.2 The report to the operators

With its sensor out of service, hall 2 is estimated at 16.4°C, with a 95 % band from 14.3 to 20.2°C; the band is set by the handbook's value of the hall's conductance, which no other reading informs, and the estimate is the supply temperature plus the hall's load, known by difference from the heat meter and the hall-1 load meter, over that handbook conductance. The sensor's last reading, 17.70°C, lies inside the band, 0.87 predictive standard deviations above its center; the check would flag a bias of the hall-2 sensor only half the time at 4.5 K and 80 %

of the time at 5.8 K. The return header, never measured, is at 17.40°C, with a band from 16.8 to 18.1°C set by the supply temperature and the total load, and the probability that the chiller receives water above its limit of 18°C is about 0.04. These answers assume the steady state and a loop heat-capacity rate of 10 kW/K, with the chiller's heat meter independent of the loop's flow meter; a flow meter that reads 25 % high (true 8 kW/K) would put the return water above the limit with no change in the reported band.

Common mistake

Reporting a virtual reading with the error of the sensor it replaces, or with no spread. The standard deviation of hall 2 without its sensor is 1.52 K, three times the sensor's 0.5 K, and its distribution is skewed; that of the return header, 0.33 K, is smaller than a cheap sensor's error. Each comes from its own budget, and neither can be guessed from the other.

Common mistake

Checking a reading against the fitted value of a fit that includes it. The hall-2 reading lies 0.21 of its standard deviation from the fitted value, 0.106 K, which looks like confirmation; held out, it lies 1.31 K from the prediction of the rest of the plant. The fit follows a reading in proportion to its leverage, equation (21.20), so the residual of a reading of small redundancy number is small whatever its error; the check is the held-out departure, or equivalently the normalized residual.

Common mistake

Reading a symmetric delta band for a virtual quantity that depends nonlinearly on an entry known only through its prior. For hall 2 the delta interval is [13.6, 19.2]°C and the pushforward's [14.3, 20.2]°C; the second-order term, 14 % of the delta standard deviation, announced the difference before any draw was made.

21.6 The next measurement (step 7)

The report leaves hall 2's temperature resting on the handbook and the return answer resting on a pinned flow. The candidate measurements are the two real sensors that would stand where the virtual ones stand: the hall-2 sensor restored ($\sigma_T = 0.5$ K), and a sensor in the return header of $\sigma_{T_r} = 0.3$ K, which the plant has never had, each alone and the two together. Each is valued for the two virtual quantities and for hall 1 at 60 kW, the question of Chapter 20, $T_1^* = T_s + 60\,\mathrm{e}^{-g_1}$.

21.6.1 The preposterior of a real sensor at a virtual sensor's place

A candidate sensor a that would read the quantity f_a with error σ_a adds the normalized row $(f_a(\mathbf{z}) - y_a)/\sigma_a$, of gradient $\mathbf{j}_a = \mathbf{g}_a/\sigma_a$ with $\mathbf{g}_a = \nabla f_a$. Its preposterior covariance is that of equation (2.17), $\boldsymbol{\Sigma}_a = (\boldsymbol{\Sigma}^{-1} + \mathbf{j}_a\mathbf{j}_a^{\mathsf{T}})^{-1}$, which the rank-one form of equation (20.28) gives without an inversion. For another output with gradient $\mathbf{g}$, equation (20.29) written in the units of the reading gives

$$\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}_a\mathbf{g} = \mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g} - \frac{\left(\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}_a\right)^2}{\sigma_a^2 + V_a}, \qquad V_a = \mathbf{g}_a^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}_a, \tag{21.27}$$

with V_a the variance of the virtual sensor at the candidate's place. For the candidate's own quantity, $\mathbf{g} = \mathbf{g}_a$, it reduces to

$$\mathbf{g}_a^{\mathsf{T}}\boldsymbol{\Sigma}_a\mathbf{g}_a = V_a - \frac{V_a^2}{\sigma_a^2 + V_a} = \frac{V_a\,\sigma_a^2}{V_a + \sigma_a^2}, \tag{21.28}$$

the precision sum of equation (21.19) again: a real sensor reduces its own quantity's variance by combining with the virtual sensor, and it helps most where the virtual variance V_a is large against σ_a^2. For another output the fraction of the variance removed is the squared correlation between that output and the quantity, times $V_a/(V_a+\sigma_a^2)$, so a sensor whose quantity is nearly uncorrelated with the output under the posterior removes almost nothing from it, however precise it is. The present state is the refit under $\mathbf{y}_-$, and both candidates are evaluated from it, at $\hat{\mathbf{z}}_-$ with $\boldsymbol{\Sigma}_-$, so that they are compared on the same footing.

21.6.2 The candidate sensors

Table 21.10 lists the present state, the two candidates and both together. The columns give the standard deviation of hall 1 at 60 kW, of hall 2's temperature and of the return temperature after each, with the width each removes from the hall at 60 kW; the preposterior is linear, so the hall-2 column holds delta-method standard deviations. Figure 21.5 draws the same numbers: in panel (a) the horizontal axis lists the present state and the two candidates and the bars are the standard deviation of T_1^*; in panel (b), for the same three cases, the blue bars are the standard deviation of T_2 and the green bars that of T_r.

The return sensor, worked. The candidate's gradient is that of T_r, $\mathbf{g}_a=(0,0,0.1,0.1,1)^{\mathsf{T}}$, and its virtual variance under $\mathbf{y}_-$ is $V_a=0.1205\,\mathrm{K}^2$. For the hall at 60 kW the gradient at $\hat{\mathbf{z}}_-$ is $\mathbf{g}=(-10.914,\ 0,\ 0,\ 0,\ 1)^{\mathsf{T}}$, the variance under $\mathbf{y}_-$ is $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}_-\mathbf{g}=1.1408\,\mathrm{K}^2$, a standard deviation of 1.068 K, and the covariance of the two quantities is

$$\begin{aligned}\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}_-\mathbf{g}_a &= -10.914\,(0.1\,\Sigma_{13}+0.1\,\Sigma_{14}+\Sigma_{15})+(0.1\,\Sigma_{53}+0.1\,\Sigma_{54}+\Sigma_{55})\\ &= -0.1402+0.0805=-0.0597\,\mathrm{K}^2,\end{aligned}$$

a correlation of -0.161. By equation (21.27),

$$\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}_a\mathbf{g}=1.1408-\frac{0.0597^2}{0.09+0.1205}=1.1408-0.0169=1.1239\,\mathrm{K}^2,$$

a standard deviation of 1.060 K, a width removed of 0.74 %. For its own quantity, equation (21.28) gives $0.1205\times 0.09/(0.1205+0.09)=0.0515\,\mathrm{K}^2$, a standard deviation of 0.227 against 0.347 K, a width removed of 35 %.

Table 21.10: The preposterior of two candidate sensors from the plant with the hall-2 sensor out ($\hat{\mathbf{z}}_-$, $\boldsymbol{\Sigma}_-$), equations (21.27) and (21.28): standard deviations of hall 1 at 60 kW, of hall 2's temperature and of the return temperature, and the width removed from the first. The standard deviations of T_2 are those of the delta method (1.417 present, against the pushforward's 1.517 K).

Case	sd(T_1^*) [K]	Width removed [%]	sd(T_2) [K]	sd(T_r) [K]
present, hall-2 sensor out	1.068	–	1.417	0.347
hall-2 sensor restored, 0.5 K	1.067	0.13	0.472	0.330
return sensor, 0.3 K	1.060	0.74	1.372	0.227
both	1.057	1.1	0.470	0.222

21.6.3 Relative value of a sensor

Neither sensor helps the hall-1 question. Restoring the hall-2 sensor removes 0.13 % of the standard deviation of T_1^* and the return sensor 0.74 %, against 18.7 % for a second hall-1 sensor in Table 20.9. The reason is the correlation of equation (21.27): T_1^* depends on g_1 and T_s, and

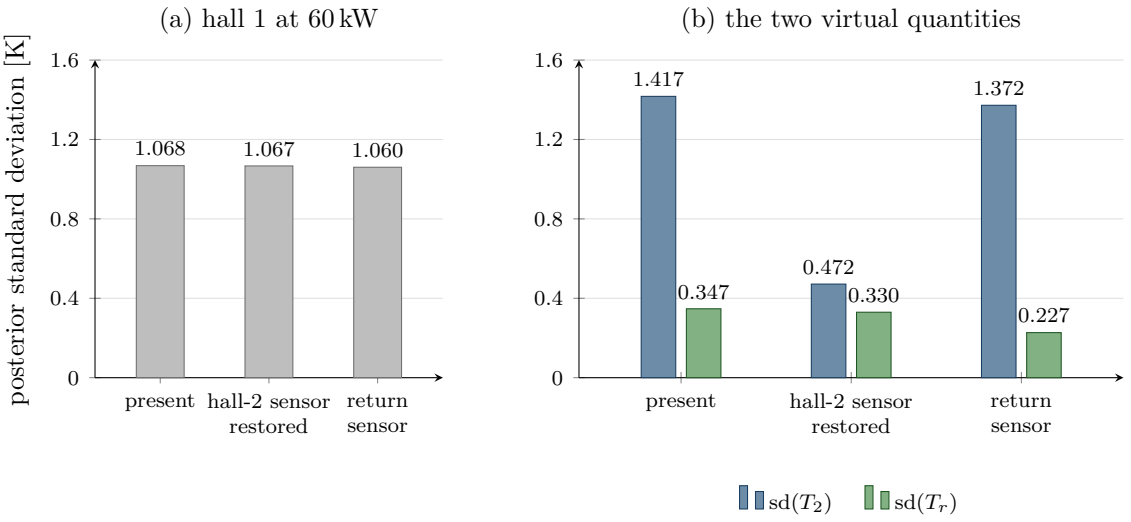


Figure 21.5: Preposterior standard deviations from the plant with the hall-2 sensor out (Table 21.10). (a) Hall 1 at 60 kW, $T_1^* = T_s + 60\,\mathrm{e}^{-g_1}$. (b) Hall 2's temperature (blue) and the return temperature (green). Candidates: the hall-2 sensor restored ($\sigma_T = 0.5\,\mathrm{K}$) and a return sensor ($\sigma_{T_r} = 0.3\,\mathrm{K}$).

the entries the two candidates tighten, g_2 and Q_2 for the hall-2 sensor, T_s and the total load for the return sensor, barely enter it. The hall-2 sensor's quantity has a correlation of 0.054 with T_1^*. Its covariance, $0.081\,\mathrm{K}^2$, comes through the g_1–Q_2 covariance, $+0.139\,\mathrm{K}^2$ (with the hall-2 reading out, Q_2 is the heat meter's reading less Q_1, and Q_1 trades with g_1 through the hall-1 reading), partly offset by the terms through T_s, $-0.058\,\mathrm{K}^2$. The return sensor's quantity has a correlation of -0.161, through T_s and, by the hall-1 reading, the trade between T_s and g_1.

Both help their own questions, and by much. The restored sensor takes hall 2's delta standard deviation from 1.42 to 0.47 K, a width removed of 67 %, because the virtual variance it combines with is large: by equation (21.28) the sensor's precision of $4\,\mathrm{K}^{-2}$ is added to the virtual sensor's 0.50. The return sensor takes the return standard deviation from 0.35 to 0.23 K, a width removed of 35 %: its precision of $11.1\,\mathrm{K}^{-2}$ adds to a virtual precision of 8.3, already comparable. Each sensor does almost nothing for the other's quantity: the return sensor takes hall 2's standard deviation only from 1.42 to 1.37 K, by sharpening Q_2, and the hall-2 sensor takes the return one from 0.35 to 0.33 K. A sensor's value is relative to a question; the ranking of sensors without a question, which Chapter 20 found undefined, is undefined here too, and a later chapter on sensor placement and the value of information formalizes the choice. On the healthy plant, with all five readings, the same return sensor would take the standard deviation of T_1^* from 1.074 to 1.064 K and that of the return temperature from 0.333 to 0.223 K; the conclusion does not depend on which plant it is added to.

The return sensor has a value the preposterior does not show: its reading does not pass through the loop balance. If the loop's flow meter read 25 % high, so that the true heat-capacity rate were 8 kW/K, the true return temperature on the healthy plant would be $12.111 + 52.92/8 = 18.73$°C, above the chiller's limit, while the virtual sensor reports 17.40 ± 0.33 K. A return sensor of 0.3 K would read 1.32 K above the virtual prediction, which against the predictive standard deviation $\sqrt{0.3^2 + 0.333^2} = 0.448$ K of equation (21.21) is 2.95 standard deviations.

The preposterior against the refit. The preposterior is computed at $\hat{\mathbf{z}}_-$ before the reading exists. The hall-2 reading does exist, and refitting with it gives back the five-reading fit, in which T_1^* has the standard deviation 1.074 K and T_2 the standard deviation 0.481 K, against the preposterior's 1.067 and 0.472. The reading of 17.70°C moves the state, g_2 from 1.609 to 1.400

and Q_1 from 30.92 to 30.69 kW, and the spreads are evaluated at the new point: the hall-1 row's sensitivity to g_1 falls with Q_1, and the standard deviation of T_1^* comes out 0.7 % larger than the preposterior's. A return reading of 17.50°C, refitted with the four other readings, gives 1.057 K against the preposterior's 1.060. Differences of a few tenths of a percent are the resolution of a linearized preposterior; the conclusion that neither sensor serves the hall-1 question is far outside them.

21.6.4 The recommendation

The next measurement for hall 2 is its own sensor, restored: it takes the standard deviation of hall 2's temperature from 1.52 K by the pushforward to 0.47 K, and no virtual sensor can stand in for it while the handbook alone holds g_2. For the return header, a sensor of 0.3 K takes the return standard deviation from 0.35 to 0.23 K and is the only instrument of the plant that would see the loop balance fail, a flow meter 25 % high showing as a departure of 2.95 predictive standard deviations; it is worth installing because the chiller's limit lies only 1.8 standard deviations above the virtual reading. A cheaper return sensor of 1 K would serve neither purpose, as Exercise 21.6 shows. Neither sensor is a measurement for hall 1's margin at 60 kW; that question's next measurement remains the second hall-1 sensor of Chapter 20.

21.6.5 What the analysis showed

The shift log carried hall 2 at its sensor's last reading with the sensor's 0.5 K; the analysis found that, with the sensor away, the plant knows hall 2 only to 1.52 K, a band from 14.3 to 20.2°C, and that the number is the handbook's conductance speaking rather than the plant: another handbook moves its median by up to a kelvin, and no remaining reading can tell the difference. The hall-2 sensor carried nine tenths of what the plant knows about that hall, and the rest of the plant checks it only through the same handbook, so its last reading's agreement, 0.87 predictive standard deviations, would have survived an error of 3 K. The return header, never measured, is the opposite case: the supply sensor, the heat meter and the load meter fix the supply temperature and the total load, and the return water is known to 0.33 K and below the chiller's 18°C with probability 0.96, on the condition that the flow meter is in calibration, which no virtual sensor can confirm. On site this means restoring the hall-2 sensor rather than trusting its virtual replacement, and installing a 0.3 K return sensor as the check on the loop balance; neither is the measurement for hall 1's margin at 60 kW.

21.6.6 Limits of the model

Every number of this chapter is conditional on the model. A virtual sensor inherits the steady state, the balances and every pinned quantity, and the checks of step 5 test how well the formulas evaluate the posterior of the model, not whether the model holds. Three cases on this plant show what that means.

The pinned heat-capacity rate. The return temperature divides the total load by $\dot{m}c_p$. The chiller's heat meter is taken to measure its duty without the loop's flow meter; a heat meter that computed the duty from the same flow reading would carry the same error into y_P, the fitted total load would scale with it, and the error would cancel in T_r. A flow meter 25 % high puts the return water above the limit while the virtual sensor reports a probability of 0.963 of staying below it (Section 21.6.3): the wrong answer with a narrow band. Carrying $\dot{m}c_p$ as an entry of the state with a prior from the flow meter would put its uncertainty in the band: the gradient of T_r with respect to it is $-(Q_1 + Q_2)/(\dot{m}c_p)^2 = -0.529$ K per kW/K, and a prior spread of 2 % (0.2 kW/K) adds $0.529^2 \times 0.2^2 = 0.0112\,\mathrm{K}^2$, raising the standard deviation from 0.333 to 0.349 K, while a spread of 5 % raises it to 0.425 K. A prior describes the instrument in

calibration; it does not describe a meter that has failed, which only a reading that does not pass through the loop balance can reveal.

The conductance of hall 2. The virtual hall-2 sensor reports the handbook's conductance. If hall 2's coil were fouled to 3 kW/K while its sensor was away, the hall would run at $12.08 + 21.58/3 = 19.27$°C, and the virtual sensor would still report 16.4°C; the true value would lie inside its band, near the upper end, so the band would not be wrong, only uninformative. A virtual sensor whose band is set by a prior cannot detect a change of the quantity the prior describes, and the held-out check of a reading taken before such a change compares it with a plant that no longer exists.

The steady state. The balances and the loop balance are steady. A hall whose load has just stepped is not at the temperature the balance gives: the virtual reading jumps to the new balance at once, while the hall warms over its thermal time constant, and the return water follows over the loop's transit time. A limit on the return temperature is then a question about a transient, which the steady rows of this plant cannot answer.

> **Principle 21.2.** *A virtual sensor is as right as the model it is computed from. When the model is wrong, through a pinned quantity, a balance or a prior that no longer describes the plant, the virtual reading is wrong with the band it would have if the model were right, and only a real sensor, whose reading does not pass through the model, can say so.*

A virtual sensor has now been seen to agree with a reading without being able to vouch for it. Chapter 22 installs the return sensor, reading 17.50°C with $\sigma_{T_r} = 0.3$ K through the loop balance with $\dot{m}c_p = 10$ kW/K, as a sixth reading, and asks which reading of the plant is lying.

21.7 What to retain

Definition 21.1 (Virtual sensor and removed row)**.** Let $f(\mathbf{z})$ be a quantity of the plant without a reading row, and let $\hat{\mathbf{z}}$, $\boldsymbol{\Sigma}$ be the MAP state and covariance of the evidence in use, on the reduced state of Chapter 17. The virtual reading is $f(\hat{\mathbf{z}})$ with variance $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$, $\mathbf{g} = \nabla f(\hat{\mathbf{z}})$, checked like any derived quantity, and reported by its pushforward when f is nonlinear in an entry known only through its prior. Removing a row k with normalized gradient $\mathbf{j}_k$ and leverage $K_{kk} = \mathbf{j}_k^\mathsf{T}\boldsymbol{\Sigma}\mathbf{j}_k < 1$ gives the downdated covariance $\tilde{\boldsymbol{\Sigma}}_- = \boldsymbol{\Sigma} + \boldsymbol{\Sigma}\mathbf{j}_k\mathbf{j}_k^\mathsf{T}\boldsymbol{\Sigma}/(1 - K_{kk})$ and $\hat{\mathbf{z}}_- \approx \hat{\mathbf{z}} + \boldsymbol{\Sigma}\mathbf{j}_k\, r_k(\hat{\mathbf{z}})/(1 - K_{kk})$, both equal to the refit's for linear rows; no spread shrinks at a fixed linearization point, and $K_{kk} = 1$ marks a critical measurement. A reading not yet seen has the predictive standard deviation $\mathrm{sd}(y_k \mid \mathbf{y}_-) = \sqrt{\sigma_k^2 + \mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}_-\mathbf{g}}$, and its standardized departure is, in the linear model, the normalized residual $r_k/\sqrt{1 - K_{kk}}$ with its sign reversed. The fitted value of a measured quantity is $K_{kk}y_k + (1 - K_{kk})f_k(\hat{\mathbf{z}}_-)$, its variance $\sigma_k^2 K_{kk}$, and the precisions of the reading and of the virtual sensor add. A sensor of error σ_a at a place with virtual variance V_a leaves its quantity the variance $V_a\sigma_a^2/(V_a + \sigma_a^2)$.

Three points carry beyond the two halls. A virtual sensor's spread is what the remaining rows say about its quantity, and its budget tells whether that is measurement or handbook: a quantity that several instruments fix in combination can be known better than a sensor would read it, while a quantity behind which only a prior remains reports the prior, and a different prior moves it with no reading able to object. The redundancy number of a reading is, on one scale, the weight of the virtual sensor of its quantity: it divides the downdate, it scales the held-out check, which is the normalized residual, and it sets the smallest error the rest of the plant can detect in that reading; a reading of small redundancy number cannot be checked by a

virtual sensor, which is the problem of bad data. And a real sensor at a virtual sensor's place is valued for a stated question: it adds its precision to the virtual one for its own quantity and removes from another output only what that output shares with it.

Table 21.11 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 21.11: The terms of virtual sensors and removed rows and where each is defined.

Term	Expression	Units	Meaning	Defined in
Virtual reading	$f(\hat{\mathbf{z}})$, variance $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$	varies	a quantity without a row	21.3.1
Loop balance	$T_r = T_s + (Q_1 + Q_2)/(\dot{m}c_p)$	°C	return-header temperature	21.1.3
Reduced objective	$\mathcal{L}_- = \mathcal{L} - \frac{1}{2}r_k^2$	–	fit without row k	21.1.6
Pushforward of T_2	$\int \Phi((t - m(x))/\sqrt{v(x)})\,\mathcal{N}(x; \ln 5, 0.3^2)\,\mathrm{d}x$	–	distribution of the hall-2 reading	21.3.1
Leverage	$K_{kk} = \mathbf{j}_k^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{j}_k$	–	complement of the redundancy number	21.3.2
Precision split	$1/(\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}) = 1/V_- + 1/\sigma_k^2$	varies	reading and virtual add	21.3.2
Fitted value	$K_{kk}y_k + (1 - K_{kk})f_k(\hat{\mathbf{z}}_-)$	reading	weights of reading and virtual	21.3.2
Predictive sd of a reading	$\mathrm{sd}(y_k \mid \mathbf{y}_-) = \sqrt{\sigma_k^2 + \mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}_-\mathbf{g}}$	reading	before it is seen	21.3.3
Held-out departure	$d_k = (y_k - f_k(\hat{\mathbf{z}}_-))/\mathrm{sd}(y_k \mid \mathbf{y}_-)$	–	check; linear: $-r_k^{\mathrm{N}}$	21.3.3
Leave-one-out residual	$r_k(\hat{\mathbf{z}}_-) = r_k(\hat{\mathbf{z}})/(1 - K_{kk})$ (linear rows)	–	residual held out	21.3.3
Rank-one downdate	$\tilde{\boldsymbol{\Sigma}}_- = \boldsymbol{\Sigma} + \boldsymbol{\Sigma}\mathbf{j}_k\mathbf{j}_k^{\mathsf{T}}\boldsymbol{\Sigma}/(1 - K_{kk})$	varies	covariance without row k	21.2.1
Critical measurement	$1 - K_{kk} = 0$	–	removal loses a direction	21.2.1
State without the row	$\hat{\mathbf{z}} + \boldsymbol{\Sigma}\mathbf{j}_k\, r_k/(1 - K_{kk})$	varies	one Gauss–Newton step	21.2.1
Sensor at a virtual place	$V_a\sigma_a^2/(V_a + \sigma_a^2)$	varies	own-quantity preposterior	21.6.1
Variance removed from f	$(\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}_a)^2/(\sigma_a^2 + V_a)$	varies	shared with the candidate	21.6.1

21.8 Exercises

Exercise 21.1. *Formulation.* A loop with $\dot{m}c_p = 10\,\mathrm{kW/K}$ has a fitted supply temperature $\hat{T}_s = 12.1°\mathrm{C}$ with standard deviation 0.3 K and a fitted total load $\hat{P} = 53\,\mathrm{kW}$ with standard deviation 2 kW, their covariance being $-0.03\,\mathrm{K\,kW}$. (a) Write the return temperature as a function of the state and the question whether it stays below 18°C as a probability. (b) Compute the virtual reading, its standard deviation and the probability. (c) Give the attribution shares of the supply temperature and the load. ((a) $T_r = T_s + P/10$; $\Pr(T_s + P/10 < 18 \mid \mathbf{y})$. (b) $12.1{+}5.3 = 17.4°\mathrm{C}$; $\mathbf{g} = (1, 0.1)$, $0.09{+}0.01\times4{+}2\times0.1\times(-0.03) = 0.124\,\mathrm{K}^2$, $\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}} = 0.352\,\mathrm{K}$; $\Phi(0.6/0.352) = \Phi(1.704) = 0.956$. (c) $(0.09 - 0.003)/0.124 = 70\,\%$ and $(0.04 - 0.003)/0.124 = 30\,\%$.)

Exercise 21.2. *Formulation.* The hall-1 sensor of the two halls is taken out instead of the hall-2 sensor. (a) State which entry of the state returns to its prior and why. (b) The refit gives $\hat{g}_1 = \ln 5$, $\hat{Q}_1 = 30.76\,\mathrm{kW}$, $\hat{T}_s = 12.127°\mathrm{C}$ with standard deviations 0.300, 1.470 kW and 0.284 K, and zero correlation of g_1 with the rest. Compute the virtual hall-1 temperature and

its delta standard deviation, and the held-out check of the reading 17.655°C. (c) What does the loss do to the owner's question of Chapter 20, hall 1 at 60 kW below 25°C? ((a) g_1: the hall-1 reading was the only reading row in which it appears, so its posterior factorizes as its prior. (b) $12.127 + 30.76/5 = 18.28$°C; $\mathbf{g} = (-6.15, 0, 0.2, 0, 1)$ and with $\mathrm{Cov}(Q_1, T_s) = 0.0135$ kW K from the refit, $6.15^2 \times 0.09 + 0.04 \times 2.159 + 0.0806 + 2 \times 0.2 \times 0.0135 = 3.407 + 0.086 + 0.081 + 0.005 = 3.58\,\mathrm{K}^2$, a standard deviation of 1.89 K; $d_1 = (17.655 - 18.280)/\sqrt{0.25 + 3.58} = -0.32$. (c) $T_1^* = 12.127 + 60/5 = 24.13$°C with standard deviation $\sqrt{12^2 \times 0.09 + 0.0806} = 3.61$ K, and $\Phi((25-24.13)/3.61) = 0.60$ (pushforward 0.59, the second-order term being 15 % of the standard deviation) against 0.966 with the sensor: a virtual sensor cannot replace a reading that is the only path to an entry.)

Exercise 21.3. *Derivation.* (a) Derive the rank-one downdate of the covariance when a row of normalized gradient $\mathbf{j}_k$ is removed, and the state $\hat{\mathbf{z}}_-$ after the removal, for linear rows. (b) Show that no output's variance decreases, and that the formula fails exactly when the row is a critical measurement. (c) Evaluate the variance of g_2 after removing the hall-2 row from $\Sigma_{22} = 0.018765$, $(\mathbf{\Sigma j}_2)_2 = -0.0722$ and $K_{22} = 0.9269$. ((a) Multiplying $\mathbf{\Sigma} + \mathbf{\Sigma j}_k \mathbf{j}_k^\mathsf{T} \mathbf{\Sigma}/(1 - K_{kk})$ by $\mathbf{\Sigma}^{-1} - \mathbf{j}_k \mathbf{j}_k^\mathsf{T}$ gives $\mathbf{I} + \mathbf{j}_k \mathbf{j}_k^\mathsf{T} \mathbf{\Sigma}\,(-1 + (1 - K_{kk})/(1 - K_{kk})) = \mathbf{I}$; the reduced gradient at $\hat{\mathbf{z}}$ is $-r_k \mathbf{j}_k$, so $\hat{\mathbf{z}}_- = \hat{\mathbf{z}} + \mathbf{\Sigma}_- \mathbf{j}_k r_k = \hat{\mathbf{z}} + \mathbf{\Sigma j}_k r_k/(1 - K_{kk})$. (b) $\mathbf{g}^\mathsf{T} \mathbf{\Sigma}_- \mathbf{g} - \mathbf{g}^\mathsf{T} \mathbf{\Sigma g} = (\mathbf{g}^\mathsf{T} \mathbf{\Sigma j}_k)^2/(1 - K_{kk}) \geq 0$; at $K_{kk} = 1$, $\mathbf{\Lambda}_- \mathbf{\Sigma j}_k = \mathbf{j}_k(1 - K_{kk}) = \mathbf{0}$ and $\mathbf{\Lambda}_-$ is singular. (c) $0.018765 + 0.0722^2/0.0731 = 0.0900$, the prior variance.)

Exercise 21.4. *Derivation.* (a) Derive the predictive standard deviation of a reading not yet seen, $\sqrt{\sigma_k^2 + \mathbf{g}^\mathsf{T} \mathbf{\Sigma}_- \mathbf{g}}$, from the law of total variance. (b) Show that in the linear model the standardized departure of the held-out reading equals minus its normalized residual at the full fit, and that the fitted value is the weighted mean $K_{kk} y_k + (1 - K_{kk}) f_k(\hat{\mathbf{z}}_-)$. (c) With the refit's $\mathrm{sd}(y_{T_2} \mid \mathbf{y}_-) = 1.503$ K and the threshold $d_2 \geq 3$, state the bias of the hall-2 sensor that the check flags half the time and the bias it flags 80 % of the time. ((a) $\mathrm{Var}(y_k \mid \mathbf{y}_-) = \mathbb{E}[\mathrm{Var}(y_k \mid \mathbf{z})] + \mathrm{Var}(\mathbb{E}[y_k \mid \mathbf{z}]) = \sigma_k^2 + \mathrm{Var}(f_k \mid \mathbf{y}_-) \approx \sigma_k^2 + \mathbf{g}^\mathsf{T} \mathbf{\Sigma}_- \mathbf{g}$. (b) The rank-one update of $\mathbf{\Sigma}_-$ by the row gives the gain $K_{kk} = V_-/(V_- + \sigma_k^2)$ on the quantity, hence the weighted mean; with $V_- = \sigma_k^2 K_{kk}/(1 - K_{kk})$ the predictive variance is $\sigma_k^2/(1 - K_{kk})$ and the departure $-\sigma_k r_k/(1 - K_{kk})$, so $d_k = -r_k/\sqrt{1 - K_{kk}}$. (c) A bias b moves the mean of d_2 by $b/1.503$, so it is flagged with probability $\Phi(b/1.503 - 3)$: half the time at $3 \times 1.503 = 4.5$ K and 80 % of the time at $(3 + 0.84) \times 1.503 = 5.8$ K; the reading's departure, 0.87 standard deviations, is well inside the threshold.)

Exercise 21.5. *Application.* A virtual flow sensor on the loop combines the chiller's heat meter ($y_P = 52.5$ kW, $\sigma = 2$ kW), a return sensor ($y_{T_r} = 17.50$°C, 0.3 K) and the supply sensor ($y_{T_s} = 12.10$°C, 0.3 K) through the loop balance, $\dot{m} c_p = P/(T_r - T_s)$, treating the three readings as independent (the heat meter does not use the loop's flow meter). (a) Compute the virtual heat-capacity rate and its delta band, and the budget. (b) Draws of the three readings give the 95 % interval [8.25, 11.66] kW/K; compare and explain. (c) Is a flow meter that reads 25 % high (true 8 kW/K) distinguishable by this virtual sensor? ((a) $52.5/5.40 = 9.72$ kW/K; gradient $(1/5.40,\ -52.5/5.40^2,\ +52.5/5.40^2) = (0.185,\ -1.800,\ 1.800)$; terms 0.137, 0.292, 0.292, total 0.720, a standard deviation of 0.85 kW/K, 81 % from the two temperatures; delta interval [8.06, 11.39]. (b) The ratio is convex in the temperature difference, so its distribution is skewed to the right and the draws' interval lies 0.2 to 0.3 kW/K above the delta interval. (c) 8 lies just below the draws' lower end, about 2 delta standard deviations from 9.72: barely; a temperature difference of 5.4 K read with two 0.3 K sensors is known to 8 %, and the flow meter, at 1 to 2 %, is the better instrument for the flow.)

Exercise 21.6. *Application.* On the healthy plant the return temperature is known to 0.333 K without any sensor. A cheap return sensor of $\sigma = 1.0$ K is offered. (a) Compute the preposterior standard deviation of the return temperature with it, and with a 0.3 K sensor. (b) Compute,

for each, the departure a flow meter that reads 25 % high (true 8 kW/K) would produce, in predictive standard deviations, the chiller's heat meter being independent of the flow meter. (c) Decide whether to install the cheap sensor. ((a) $0.1108 \times 1/(0.1108 + 1) = 0.0997\,\mathrm{K}^2$, a standard deviation of 0.316 K, a width removed of 5 %; with 0.3 K, $0.1108 \times 0.09/0.2008 = 0.0497$, a standard deviation of 0.223 K, 33 %. (b) The true return temperature would be 18.73°C, 1.32 K above the virtual reading; the predictive standard deviation $\sqrt{1 + 0.1108} = 1.054\,\mathrm{K}$ gives 1.25 standard deviations, and $\sqrt{0.09 + 0.1108} = 0.448\,\mathrm{K}$ gives 2.95. (c) No: the virtual standard deviation is already a third of the cheap sensor's error, so the sensor narrows it by only 5 %, and it cannot flag the model error that would matter for the 18°C limit; if a check of the loop balance is the purpose, the 0.3 K sensor serves it, and otherwise the virtual sensor suffices.)

Chapter 22

Detecting Sensor Bias and Bad Data

22.1 Readings that may be lying

The two halls of Chapter 17 now have a sensor on the return header as well: six readings and three handbook priors for the five free quantities of the loop. One morning the hall-2 sensor reads 19.90°C, 2.2 K above its reading of the day before, while every other reading of the loop is where it was. The operator fits the plant as Chapter 18 does and runs its check: twice the minimized objective is 3.40 on four degrees of freedom, a chi-square tail of 0.49, and the largest normalized residual is 1.77, well under the threshold of 3 the control room uses. The first reading of the fit is that the readings agree with the physics, so the sensor is right and hall 2 has warmed; at the 40 kW the hall is to carry next quarter, the same fit puts it at 25.26°C.

The engineer who maintains the sensors reads the same numbers differently. A passed check says that the record as a whole is plausible under the model; it does not say that each reading is, and the hall-2 sensor is the only reading of the hall-2 temperature, checked only through the handbook value of the hall's conductance. Whether this one reading is lying, by how much, and whether the record could have shown it at all are questions the check does not settle. The question of this chapter is

how a biased or grossly wrong reading is detected, how it is told apart from the other readings and from a change of the plant, how large the bias is and how sure that is, what the bias costs when the reading is kept, and which measurements make a sensor checkable at all.

Jargon: bad data, gross error, bias, interacting bad data, bad-data detection, identification, normalized residual, largest-normalized-residual test, redundancy number, residual sensitivity matrix, critical measurement, critical pair, leave-one-out residual, minimum detectable bias, Bonferroni bound, bias drawing, Occam factor

Bad data is a reading whose error is not described by the row's error model: a *gross error* is a single reading far off, a *bias* is an offset that persists from one reading to the next; *interacting bad data* are several bad readings whose effects on the residuals overlap. *Bad-data detection* decides whether a record contains bad data; *identification* decides which reading it is. The *normalized residual* of a row is its standardized residual at the fit divided by the root of its *redundancy number* $1 - K_{kk}$, the diagonal entry of the *residual sensitivity matrix* $\mathbf{I} - \mathbf{K}$, the complement of the hat matrix (Chapter 18); the *largest-normalized-residual test* flags the row whose normalized residual is largest in magnitude when it exceeds a threshold. A *critical measurement* is a reading whose removal leaves the state unobservable; its redundancy number is zero. A *critical pair* is two readings that check only each other: their normalized residuals are equal in magnitude for every record. A *leave-one-out residual* is the prediction of a reading from the fit without it minus the reading, with the sign of every residual of this book. The *minimum detectable bias* of a reading is the bias at which its expected normalized residual reaches the test's threshold. The *Bonferroni bound* limits the probability that any of several statistics crosses its threshold by the sum of the individual probabilities. A *bias drawing* is the plant model with one bias entry added to one reading row. The *Occam factor* is the penalty a log evidence pays for an added entry, the logarithm of the ratio of the entry's prior width to its posterior width, subtracted from the gain in fit.

Key idea

A bias on one reading is found by freeing a bias entry on that reading and asking whether the evidence prefers the drawing with it; for linear rows that evidence is a function of the reading's normalized residual and of how well the rest of the plant predicts the reading. A reading that the other rows predict poorly, a reading of small redundancy number, carries a bias that no residual and no evidence can see: the fit absorbs it into the entry only that reading informs. Whether a sensor can be checked is decided by the design before any reading is taken.

22.1.1 From detection to identification

The chi-square check of Chapter 18 is a detection test. It compares twice the minimized objective with its reference distribution and answers whether the record, taken whole, is consistent with the model. Chapter 18's own gross-error case, a hall-2 reading of 22.0°C in place of 17.70°C (Exercise 18.5), showed both what such a test misses and what a residual table cannot do: the fault of 8.6 standard deviations of the sensor raised the misfit only to a tail probability of 0.083, and the normalized residuals pointed at a chain of three rows with the same magnitude, the faulty reading, the chiller meter and the prior on the hall-2 conductance, without naming one. Identification asks which reading is at fault, and it is a different question from detection: a test can detect without identifying, when several rows share the disagreement, and a reading can be wrong without the record failing detection, when the state absorbs its error.

The question *is this reading lying* is therefore three questions: whether any reading is (detection), which one (identification), and by how much (estimation of the bias). The first is a tail of the kind Chapter 18 computed. The second and third are a diagnosis in the sense of Chapter 2: a family of candidate explanations generated from the graph, one per sensor, ranked by their evidence, with the bias of the leading candidate estimated as an entry of its state.

22.1.2 The two methods

The residual test uses only the fit without bias. At the MAP state every row has a residual, and the hat matrix says how much of each row's error the residual can show; dividing by that share puts every row on a common scale, and the row furthest out is the suspect. The test needs one fit and is the standard practice of power-system and process state estimation. It returns a suspect and a threshold crossed or not, not a magnitude with a band, and no odds of the suspect against the next row.

The evidence comparison fits one drawing per sensor, each with a bias entry carrying a prior, and compares the evidence of each with the evidence of the drawing without bias. It returns the bias with its band, a Bayes factor per sensor and posterior probabilities over the family.

22.1.3 The plant with the return sensor (step 0)

The plant is the two-hall loop of Figure 17.1: two hall nodes with their heat loads as sources, two heat-removal paths of conductance $G_i = G_{\text{ref}} e^{g_i}$ to the supply header at T_s, and the chiller, which removes the heat $P = Q_1 + Q_2$. Chapter 21 added the return header, whose temperature follows from the energy balance of the loop water: with the loop's heat capacity rate $\dot{m}c_p$, in kW/K, the water warms from T_s to T_r while it collects the two loads,

$$\dot{m}c_p\,(T_r - T_s) = Q_1 + Q_2, \qquad T_r(\mathbf{z}) = T_s + \frac{Q_1 + Q_2}{\dot{m}c_p}, \tag{22.1}$$

with $\dot{m}c_p = 10\,\text{kW/K}$ pinned from the loop's flow meter and the specific heat of water. In this chapter the return header has a sensor of its own, $\sigma_{T_r} = 0.3\,\text{K}$, the candidate of Chapter 21 installed. The loop balance is a fourth balance beside the two hall balances and the header balance, and the return temperature a quantity it determines, as the hall balances determine the hall temperatures. The free quantities are unchanged,

$$\mathbf{z} = (g_1,\ g_2,\ Q_1,\ Q_2,\ T_s)^{\mathsf{T}},$$

the five entries of the reduced state of Chapter 17.

Figure 22.1 draws the plant graph. The green boxes are the free nodes: the two halls, the supply header and the return header, whose temperatures follow from the balances; the white boxes are the sources, the two loads and the chiller; the gray box is the pinned heat capacity rate of the loop. Solid green arrows are the paths of heat, and the solid green line from the chiller to the return header is the loop balance (22.1). The six dashed boxes are the readings, joined by dashed arrows to the quantity each measures. The red dashed box is a bias entry attached to the supply sensor's row: the drawing M_{T_s} of Section 22.1.6, one member of a family that attaches one such box to each reading in turn.

22.1.4 The clean record and the two biased records

Table 22.1 gives the three records of the chapter. Each row is one reading with its standard deviation; the columns are the clean record, the five readings of Table 18.2 with the return sensor reading 17.50°C, and two records built from it by adding a bias to one sensor. Record (a) has the supply sensor reading 1.0 K high, 13.10°C in place of 12.10: an offset of 3.3 standard deviations of that sensor, the size of a drifted calibration. Record (b) has the hall-2 sensor reading 2.2 K high, 19.90°C in place of 17.70: 4.4 standard deviations of the sensor, the size of a sensor mounted in the warm-air stream of a rack, and the record of the opening of this chapter. The records are illustrative; the clean one is the record of Chapter 18 with the new reading added, and the biases are known by construction, which is what lets the chapter score the methods.

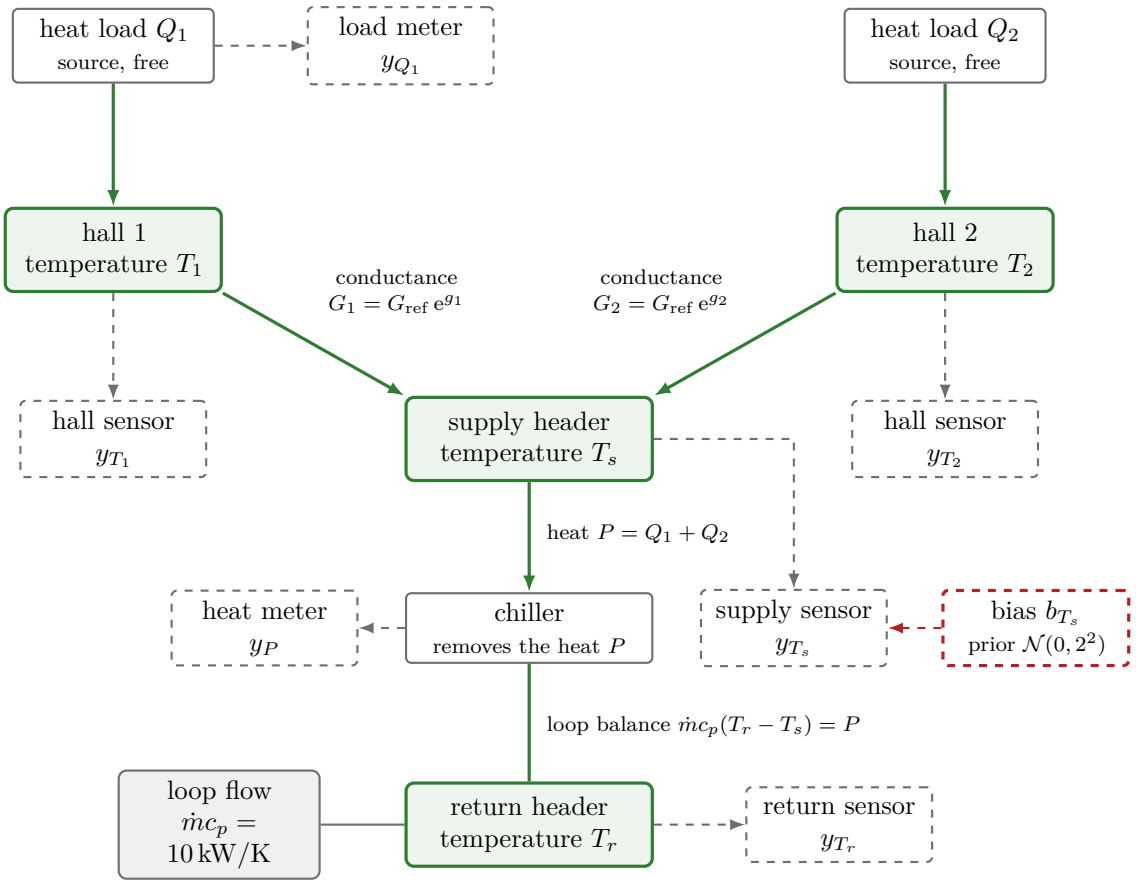


Figure 22.1: Plant graph of the two halls with the return-header sensor installed. Free: the log-conductances g_1, g_2, the loads Q_1, Q_2 and the supply temperature T_s, the five entries of the reduced state; the hall, return and chiller quantities are carried entries that the four balances determine. Pinned: $G_{\text{ref}} = 1\,\text{kW/K}$ and the loop's heat capacity rate $\dot{m}c_p = 10\,\text{kW/K}$. Observed (dashed): T_1, T_2, T_s, P, Q_1 and T_r. Red dashed: the bias entry of the drawing M_{T_s}, a sixth free quantity with its own prior row, attached to the supply sensor's row; the other drawings attach it to another reading. The three priors are rows of the evidence and are not drawn.

Table 22.1: The three records of the chapter. The clean record is that of Chapter 18 with the return sensor added; records (a) and (b) change one reading each (in bold). Illustrative values.

Reading	σ_k	Clean record	Record (a), supply +1.0 K	Record (b), hall 2 +2.2 K
hall-1 temperature y_{T_1}	0.5 K	17.655°C	17.655°C	17.655°C
hall-2 temperature y_{T_2}	0.5 K	17.70°C	17.70°C	**19.90**°C
supply temperature y_{T_s}	0.3 K	12.10°C	**13.10**°C	12.10°C
chiller heat meter y_P	2 kW	52.5 kW	52.5 kW	52.5 kW
hall-1 load meter y_{Q_1}	1.5 kW	31.0 kW	31.0 kW	31.0 kW
return temperature y_{T_r}	0.3 K	17.50°C	17.50°C	17.50°C

The priors are those of Table 17.2: $g_i \sim \mathcal{N}(\ln 5,\ 0.3^2)$ on each log-conductance, from the handbook value of 5 kW/K, and $T_s \sim \mathcal{N}(12^\circ\text{C},\ 1\,\text{K}^2)$ from the chiller's set point; the loads carry none. Each record is one steady operating condition, read once.

22.1.5 The questions (step 1)

Table 22.2 lists the six questions of the chapter. The first column states each in engineering language; the second gives the general form of its category in the notation of Chapter 2; the third its mathematical form for the loop; the fourth its category; and the last the subsection where it is answered. The first is the detection check of Chapter 18. The second is the owner's question, in the owner's words, answered by the residual test: a threshold r^* on the largest normalized residual, set so that a record without bad data crosses it with probability at most p_{fa}, a tail of the predictive under the drawing without bias. The third asks the same question of the evidence: a hypothesis comparison run over a family of drawings, one bias per reading (Section 22.1.6), which is the diagnosis category, with the bias of each drawing estimated as an entry of its state. The fourth applies the observability of Chapter 19 to a bias: the probability that a bias Δ on the hall-2 sensor pushes its normalized residual across the threshold. The fifth is an estimation question asked under the drawing that keeps the biased reading, the drawing that frees its bias and the drawing that drops the reading. The sixth is a design question, answered by quantities of a row not yet installed. Two kinds of star appear. The scenario values are starred as in Chapter 20: T_2^* is the hall-2 temperature at a load $Q_2^* = 40\,\text{kW}$ and T_1^* the hall-1 temperature at $Q_1^* = 60\,\text{kW}$. The threshold r^* and the minimum detectable bias b_k^* are starred as the levels against which a statistic and a bias are compared. The atlas of Appendix A lists all six questions.

Table 22.2: The questions asked of the loop, the general form of their category, their mathematical form, their category and where each is answered. $\mathbf{y}$ is the six readings and three priors of Section 22.1.6; $\nu = 4$; $t(y_{\text{new}})$ the statistic of a replicate record; V_{-2} the variance of the hall-2 reading's prediction from the other rows; M_0 the drawing without bias and M_k the drawing with a bias on reading k.

Engineering question	General form	Mathematical form	Category	Answered in
Do the readings disagree with the model?	$\Pr(t(y_{\text{new}}) \geq t^* \mid M_0)$, a tail of the predictive	$\Pr(\chi^2_4 > 2\mathcal{L}(\hat{\mathbf{z}}_0))$	prediction under a threshold (detection)	22.3.1
Is this reading lying? Which one?	$\Pr(t(y_{\text{new}}) \geq t^* \mid M_0)$, a tail of the predictive	$\Pr\big(\max_k \lvert r_k^{\text{N}}\rvert > r^* \mid M_0\big) \leq p_{\text{fa}}$; the row of largest $\lvert r_k^{\text{N}}\rvert$	prediction under a threshold (a test)	22.3.1
Which sensor is biased, by how much, and how sure is that?	$\Pr(M_j \mid \mathbf{y}) \propto p(\mathbf{y} \mid M_j)\ \Pr(M_j)$; $\Pr(f(\mathbf{z}) \leq f^* \mid \mathbf{y})$	$p(\mathbf{y} \mid M_k)/p(\mathbf{y} \mid M_0)$ for each k; $\hat{b}_k$ and $\text{sd}(b_k \mid \mathbf{y}, M_k)$	diagnosis by evidence; estimation	22.3.2
Can a bias on the hall-2 sensor be detected at all?	$\Pr(f \geq f^* \mid \cdot)$ on a predictive	$\Pr(\text{detect} \mid \Delta) \approx \Phi\big(\Delta/\sqrt{\sigma_T^2 + V_{-2}} - r^*\big)$; the redundancy number $1 - K_{22}$	prediction under a threshold (detectability)	22.3.3
What does the biased reading do to the answers if it is kept?	$\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]$, $\text{Var}(f(\mathbf{z}) \mid \mathbf{y})$	$p(T_2^* \mid \mathbf{y}, M_0)$ against $p(T_2^* \mid \mathbf{y}, M_{T_2})$; the same for T_1^*	estimation under three drawings	22.3.4
Which row makes the hall-2 sensor checkable?	$U(a)$, the design form	$1 - K_{22}$ and $b_{T_2}^*$ after a candidate row a, from $\mathbf{\Sigma}_a$	design	22.3.5

22.1.6 The evidence as rows and the bias drawings (step 2)

The evidence is nine rows in normalized form, the six readings and the three priors of Chapter 18, written on the reduced state:

$$\begin{aligned} &r_1 = (T_s + Q_1 \mathrm{e}^{-g_1} - y_{T_1})/\sigma_T, && r_2 = (T_s + Q_2 \mathrm{e}^{-g_2} - y_{T_2})/\sigma_T, && r_3 = (T_s - y_{T_s})/\sigma_s, \\ &r_4 = (Q_1 + Q_2 - y_P)/\sigma_P, && r_5 = (Q_1 - y_{Q_1})/\sigma_Q, && r_6 = (T_r(\mathbf{z}) - y_{T_r})/\sigma_{T_r}, \\ &r_7 = (g_1 - g^0)/\sigma_g, && r_8 = (g_2 - g^0)/\sigma_g, && r_9 = (T_s - T_s^0)/\sigma_0, \end{aligned} \tag{22.2}$$

with $G_{\mathrm{ref}} = 1\,\mathrm{kW/K}$ absorbed, $\sigma_T = 0.5\,\mathrm{K}$, $\sigma_s = 0.3\,\mathrm{K}$, $\sigma_P = 2\,\mathrm{kW}$, $\sigma_Q = 1.5\,\mathrm{kW}$, $\sigma_{T_r} = 0.3\,\mathrm{K}$, $\sigma_g = 0.3$, $\sigma_0 = 1\,\mathrm{K}$, $g^0 = \ln 5$ and $T_s^0 = 12°\mathrm{C}$, as in Table 18.2; the loads carry no prior. The redundancy is

$$\nu = n_r - n_z = 9 - 5 = 4,$$

one more than in Chapter 18. Which row acts on which entry fixes where the Jacobian of the rows can be nonzero, marked $*$:

$$\mathbf{J} = \begin{bmatrix} * & 0 & * & 0 & * \\ 0 & * & 0 & * & * \\ 0 & 0 & 0 & 0 & * \\ 0 & 0 & * & * & 0 \\ 0 & 0 & * & 0 & 0 \\ 0 & 0 & * & * & * \\ * & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * \end{bmatrix} \begin{matrix} \text{hall-1 reading} \\ \text{hall-2 reading} \\ \text{supply reading} \\ \text{heat meter} \\ \text{load meter} \\ \text{return reading} \\ \text{prior on } g_1 \\ \text{prior on } g_2 \\ \text{prior on } T_s \end{matrix}, \qquad \text{columns } g_1, g_2, Q_1, Q_2, T_s. \tag{22.3}$$

Two columns are short. The column of g_2 has two entries, the hall-2 reading and the prior on g_2, and the column of g_1 the same pair for hall 1: each conductance is seen by one reading and one handbook value, and nothing else. The return row shares the columns of Q_1, Q_2 and T_s with the heat meter and the supply sensor, which is the loop balance (22.1) seen in the pattern: the return sensor, the supply sensor and the chiller meter are tied, so each is predicted by the other two, and the supply sensor becomes checkable where it was not in Chapter 18.

The candidate explanations are generated from the graph before the records are examined: every reading may carry a bias. For reading k the drawing M_k adds one free quantity, a bias b_k in the units of the reading, to the value the model predicts for it, and one prior row on that bias:

$$r_k^{(M_k)} = \frac{h_k(\mathbf{z}) + b_k - y_k}{\sigma_k}, \qquad r_{\mathrm{b}} = \frac{b_k}{\sigma_b}, \qquad \sigma_b = 2, \tag{22.4}$$

where $h_k(\mathbf{z})$ is the predicted reading of row k in equation (22.2) and $\sigma_b = 2$ is in the reading's units, 2 K for a temperature sensor and 2 kW for a meter. The prior says that a bias, if present, is of the order of a few standard deviations of the sensor and is equally likely to be positive or negative. The state of M_k is $(\mathbf{z}, b_k)$, six entries, and its rows are the ten of equations (22.2) and (22.4); the redundancy is still four. In the pattern (22.3) the bias adds a sixth column with two entries, one in row k and one in the bias prior row. The drawing M_0 has no bias entry and is the plant of equation (22.2). The family $\{M_0, M_{T_1}, M_{T_2}, M_{T_s}, M_P, M_{Q_1}, M_{T_r}\}$ has seven members, one per reading plus the drawing without bias, and it is fixed here, before any record is looked at: a family chosen after the records would be chosen to fit them. The priors on the conductances and the supply temperature are the same in every drawing, so the drawings differ only in the bias entry and its row. Chapter 23 reuses the sensor-bias entries: they are among the candidate faults it ranks against changes of the plant itself.

The modeling choices. Four modeling choices make the rows and the drawings, and each buys something and can fail. The chapter works on the reduced state of Chapter 17: the full state $(g_1, g_2, Q_1, Q_2, T_s, T_1, T_2, P, T_r)$ carries the two hall balances, the header balance and the loop balance (22.1) as physics rows at the tolerance σ_p, $\mathbf{N}$ is the null-space basis of their Jacobian, the Jacobian of the nine rows (22.2) is $\mathbf{J} = \mathbf{J}_\mathrm{e}\mathbf{N}$ (equation (17.18)), and $\mathbf{\Sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}$ is the free-entry marginal of the full covariance in the limit $\sigma_p \to 0$ (equation (17.19)); this keeps the hat matrix 9×9 over five entries, and the limit is checked for this chapter's quantities in Section 22.4.5. The rows are taken as linear over the spread of the fit: this is what makes the residuals at the fit Gaussian with covariance $\mathbf{I} - \mathbf{K}$ and twice the misfit chi-square, and it holds only approximately for the two hall reading rows, which on the reduced state carry the products $Q_i \mathrm{e}^{-g_i}$, to the extent that the fit moves when the errors are redrawn; Table 18.8 checked it by simulation for this plant. The family holds one constant bias on one reading at a time, fixed before the records are seen: this keeps it to seven drawings of one added entry each, and it cannot explain two biased readings at once, a bias that drifts over the record, or a change of the plant itself, which Section 22.6.5 discusses. The bias carries the prior $\mathcal{N}(0, \sigma_b^2)$ with $\sigma_b = 2$ in the reading's units, the calibration tolerance of a field sensor: this makes the evidence of a drawing finite and lets it pay for its added entry, and the Bayes factor depends on it, which Section 22.4.4 measures.

22.2 The posterior of the state under each drawing (step 3)

In the book's notation $p(\mathbf{z} \mid \mathbf{y})$ is the posterior density of the state, $f(\mathbf{z})$ a quantity derived from it, y_new a future or replicate record and M_j a drawing.

All six questions are read from the posteriors of the drawings: the posterior of the state that the nine rows of step 2 give under the drawing without bias, and the posterior of the state and the bias that each bias drawing gives, each with its evidence. This section derives them and computes them for the three records; Section 22.3 reads every answer from them.

Four objects of the fit recur in the specialized forms. The *MAP state* $\hat{\mathbf{z}}_0$ of the drawing without bias is the minimum of the objective that the nine rows put on the state; the *information matrix* $\mathbf{\Lambda} = \mathbf{J}^\mathsf{T}\mathbf{J}$ is the Gauss–Newton curvature of that objective at the minimum, with $\mathbf{J}$ the 9×5 Jacobian of the normalized rows; the *Laplace covariance* $\mathbf{\Sigma} = \mathbf{\Lambda}^{-1}$ makes the posterior approximately normal, $\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}_0, \mathbf{\Sigma})$; and the *hat matrix* $\mathbf{K} = \mathbf{J}\mathbf{\Sigma}\mathbf{J}^\mathsf{T}$ carries the residual tests. Each bias drawing M_k has its own MAP state $\hat{\mathbf{z}}_k$, information $\mathbf{\Lambda}_k$ and covariance $\mathbf{\Sigma}_k$, six entries each, and its evidence.

22.2.1 The rows, the fit and the residuals at the fit

Every question of step 4 uses the fit of the nine rows and the distribution of their residuals at it. The rows of equation (22.2) are Gaussian and normalized, so each puts the term $\mathcal{L}_k = \frac{1}{2} r_k(\mathbf{z})^2$ in the objective and the objective is $\mathcal{L}_0 = \frac{1}{2}\mathbf{r}^\mathsf{T}\mathbf{r}$. Chapter 18 derived its minimization by Gauss–Newton: at each iterate the normal equations

$$(\mathbf{J}^\mathsf{T}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^\mathsf{T}\mathbf{r}, \qquad \mathbf{z}_{j+1} = \mathbf{z}_j + \Delta\mathbf{z},$$

with $\mathbf{J}$ the 9×5 Jacobian of the pattern (22.3) evaluated at $\mathbf{z}_j$; at the minimum the information matrix is $\mathbf{\Lambda} = \mathbf{J}^\mathsf{T}\mathbf{J}$ and the Laplace covariance $\mathbf{\Sigma} = \mathbf{\Lambda}^{-1}$. A bias drawing M_k adds the column $\mathbf{e}_k/\sigma_k$ for b_k and the row b_k/σ_b, and the same iteration on its six-entry state $(\mathbf{z}, b_k)$ gives $\hat{\mathbf{z}}_k$, $\mathbf{\Lambda}_k$ and $\mathbf{\Sigma}_k$: the posterior of a bias drawing is the same construction as that of M_0, on the extended state.

For linear rows Chapter 18 derived the residual vector at the MAP state, equation (18.23),

$$\mathbf{r}(\hat{\mathbf{z}}_0) = -(\mathbf{I} - \mathbf{K})\,\boldsymbol{\zeta}, \qquad \mathbf{K} = \mathbf{J}(\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}\mathbf{J}^\mathsf{T} = \mathbf{J}\mathbf{\Sigma}\mathbf{J}^\mathsf{T},$$

with $\boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ the normalized errors of the rows and $\mathbf{K}$ the hat matrix, symmetric and idempotent; equally, $\mathbf{r}(\hat{\mathbf{z}}_0) = -(\mathbf{I} - \mathbf{K})\tilde{\mathbf{y}}$, with $\tilde{\mathbf{y}}$ the vector of readings and prior means divided by their standard deviations (written $\mathbf{b}$ in Chapter 18; renamed here to keep b for a bias), whose entries are $\tilde{y}_k$. State-estimation practice calls $\mathbf{I} - \mathbf{K}$ the *residual sensitivity matrix*: its column k is the response of every residual to a unit error in row k. A linear map of a Gaussian vector is Gaussian, with mean the map of the mean and covariance the map applied on both sides:

$$\mathbb{E}[\mathbf{r}(\hat{\mathbf{z}}_0)] = \mathbf{0}, \qquad \operatorname{Cov}[\mathbf{r}(\hat{\mathbf{z}}_0)] = (\mathbf{I} - \mathbf{K})\,\mathbf{I}\,(\mathbf{I} - \mathbf{K})^{\mathsf{T}} = (\mathbf{I} - \mathbf{K})^2 = \mathbf{I} - \mathbf{K}.$$

Row k's residual is therefore normal with variance $1 - K_{kk}$, its redundancy number, and the normalized residual of equation (18.25) is a standard normal variable under the model,

$$r_k^{\mathrm{N}} = \frac{r_k(\hat{\mathbf{z}}_0)}{\sqrt{1 - K_{kk}}} \sim \mathcal{N}(0, 1), \qquad \operatorname{Cov}(r_k^{\mathrm{N}}, r_l^{\mathrm{N}}) = \frac{-K_{kl}}{\sqrt{(1 - K_{kk})(1 - K_{ll})}}. \tag{22.5}$$

The normalized residuals of different rows are correlated, through the off-diagonal entries of the hat matrix, and their number is $n_r = 9$ while only $\nu = 4$ of them are free: the residual vector lies in the four-dimensional space onto which $\mathbf{I} - \mathbf{K}$ projects, and twice the misfit, $\boldsymbol{\zeta}^{\mathsf{T}}(\mathbf{I} - \mathbf{K})\boldsymbol{\zeta}$, is χ^2_4 (equation (18.24)). The two hall reading rows are the rows of the reduced state that are nonlinear in it, through the products $Q_i \mathrm{e}^{-g_i}$; for them the result holds to the extent that the rows are linear over the spread of the fit, the caveat that Table 18.8 checked by simulation for this plant.

The evidence of a drawing. The evidence of a drawing is the density of the record under it, averaged over the drawing's prior (equation (4.18)). For M_0 with its nine Gaussian rows and the flat prior of constant density c_Q on the two loads,

$$p(\mathbf{y} \mid M_0) = c_Q \int \prod_{k=1}^{9} \frac{1}{\sqrt{2\pi}\,\sigma_k} \mathrm{e}^{-r_k(\mathbf{z})^2/2}\,\mathrm{d}\mathbf{z} = c_Q \prod_{k=1}^{9} \frac{1}{\sqrt{2\pi}\,\sigma_k} \int \mathrm{e}^{-\mathcal{L}_0(\mathbf{z})}\,\mathrm{d}\mathbf{z},$$

with $\mathcal{L}_0 = \frac{1}{2}\sum_k r_k^2$ the objective of the normalized rows. Expanding $\mathcal{L}_0$ to second order about its minimum and integrating, equation (19.9) with the Gauss–Newton curvature $\boldsymbol{\Lambda}_0 = \mathbf{J}^{\mathsf{T}}\mathbf{J}$ in place of the Hessian, gives the Laplace evidence

$$\ln p(\mathbf{y} \mid M_0) \approx \ln c_Q - \sum_{k=1}^{9} \ln(\sqrt{2\pi}\,\sigma_k) - \mathcal{L}_0(\hat{\mathbf{z}}_0) + \frac{5}{2}\ln 2\pi - \frac{1}{2}\ln\det \boldsymbol{\Lambda}_0. \tag{22.6}$$

22.2.2 The redundancy number as a prediction from the other rows

The redundancy number has a reading in physical units that makes it a design quantity. Remove row k and fit the other rows: call $\hat{\mathbf{z}}_{-k}$ the MAP state without it, $\boldsymbol{\Sigma}_{-k}$ its covariance and $\hat{h}_{-k} = h_k(\hat{\mathbf{z}}_{-k})$ the value the rest of the plant predicts for reading k, a virtual sensor in the sense of Chapter 21. With $\mathbf{j}_k^{\mathsf{T}}$ the row of $\mathbf{J}$ belonging to row k, normalized by σ_k, the variance of that prediction in the units of the reading is $V_{-k} = \sigma_k^2\,\mathbf{j}_k^{\mathsf{T}}\boldsymbol{\Sigma}_{-k}\mathbf{j}_k$, the band of the virtual sensor. Chapter 21 showed, equation (21.19), that for linear rows the reading and the virtual sensor of the same quantity add their precisions, $1/(\sigma_k^2 K_{kk}) = 1/V_{-k} + 1/\sigma_k^2$, and that $V_{-k} = \sigma_k^2 K_{kk}/(1 - K_{kk})$. Solving the latter for the redundancy number,

$$1 - K_{kk} = \frac{\sigma_k^2}{\sigma_k^2 + V_{-k}}. \tag{22.7}$$

The redundancy number of a reading is the share of the sensor's own variance in the variance of the reading as the rest of the plant predicts it. It is near one when the other rows predict

the reading much better than the sensor measures it, and near zero when they predict it much worse.

The residual has the same reading. Chapter 21 showed, equation (21.24), that the standardized departure of a reading from its virtual sensor is minus the normalized residual. Written with the sign of every residual of this book, predicted minus observed, and with equation (22.7) for the variance of the difference,

$$r_k^{\mathrm{N}} = \frac{\hat{h}_{-k} - y_k}{\sqrt{\sigma_k^2 + V_{-k}}} : \tag{22.8}$$

the normalized residual is the *leave-one-out residual*, the prediction of the reading from the other rows minus the reading, divided by the standard deviation of that difference, the sensor's variance plus the prediction's. It is the predictive check of a reading not yet seen, applied to every reading after the fact, and it explains the second form of equation (22.18): the other rows do not see a bias on reading k, so the prediction $\hat{h}_{-k}$ is unchanged on average while the reading moves by Δ. Two consequences follow. Deleting row k lowers the minimized misfit by

$$2\mathcal{L}(\hat{\mathbf{z}}_0) - 2\mathcal{L}_{-k}(\hat{\mathbf{z}}_{-k}) = \left(r_k^{\mathrm{N}}\right)^2, \tag{22.9}$$

the standard deletion identity of least squares (Exercise 22.3), so that the row of largest normalized residual is the row whose removal explains the most misfit. And a reading that the other rows cannot predict at all, $V_{-k} = \infty$, has $1 - K_{kk} = 0$: it is a *critical measurement*, its removal leaves an entry unobserved, and by $\mathrm{Var}[r_k(\hat{\mathbf{z}}_0)] = 1 - K_{kk} = 0$ with mean zero its residual is zero for every record, whatever error the reading carries.

Critical pairs. If two rows k and l have proportional columns in the residual sensitivity matrix, $(\mathbf{I} - \mathbf{K})\mathbf{e}_k = \varphi\,(\mathbf{I} - \mathbf{K})\mathbf{e}_l$ for a scalar φ, then a bias on either produces the same pattern of residuals up to scale, and for every record

$$r_k(\hat{\mathbf{z}}_0) = \varphi\, r_l(\hat{\mathbf{z}}_0), \qquad 1 - K_{kk} = \varphi^2(1 - K_{ll}), \qquad r_k^{\mathrm{N}} = \mathrm{sign}(\varphi)\, r_l^{\mathrm{N}}, \tag{22.10}$$

the first from $r_k = -\mathbf{e}_k^{\mathsf{T}}(\mathbf{I}-\mathbf{K})\tilde{\mathbf{y}}$, the second from $\mathbf{e}_k^{\mathsf{T}}(\mathbf{I}-\mathbf{K})\mathbf{e}_k = \varphi\,\mathbf{e}_l^{\mathsf{T}}(\mathbf{I}-\mathbf{K})\mathbf{e}_k = \varphi^2\,\mathbf{e}_l^{\mathsf{T}}(\mathbf{I}-\mathbf{K})\mathbf{e}_l$ by symmetry. The two rows always carry normalized residuals of the same magnitude, and by equation (22.9) deleting either removes the same misfit. State-estimation practice calls such rows a *critical pair*: removing both leaves an entry unobserved, and each is checked only by the other. This is the tie of the chain of three rows in Chapter 18's gross-error exercise, and by the pattern (22.3) it occurs here between a hall reading and the prior on its conductance, the only two rows in that conductance's column.

> **Principle 22.1.** *For linear rows the normalized residual of a reading is its leave-one-out residual: the other rows' prediction of the reading minus the reading, over the spread $\sqrt{\sigma_k^2 + V_{-k}}$ of that difference. A bias Δ shifts it by $\Delta/\sqrt{\sigma_k^2 + V_{-k}}$, and a critical measurement, which the other rows cannot predict, has a residual of zero whatever its error.*

22.2.3 The posterior of the loop for the three records

The loop's inputs are the three records of Table 22.1, the priors of Table 17.2 and the bias prior $\sigma_b = 2$ in each reading's units. Every fit starts from the prior point $\mathbf{z}^{(0)} = (\ln 5,\ \ln 5,\ 28\,\mathrm{kW},\ 28\,\mathrm{kW},\ 12°\mathrm{C})$ of Chapter 17, and every bias drawing from the fit of M_0 with its bias at zero.

Table 22.3 gives the Gauss–Newton iterations of the drawing without bias for record (b). Each row is an iterate; the columns give its five entries, twice the objective, the norm of the

Table 22.3: Gauss–Newton iterations of the fit without bias to record (b), from the prior point. Loads in kW, supply temperature in °C; $\|\mathbf{J}^{\mathsf{T}}\mathbf{r}\|$ and $\|\Delta\mathbf{z}\|$ in the mixed units of the entries.

Iteration j	g_1	g_2	Q_1	Q_2	T_s	$2\mathcal{L}$	$\|\mathbf{J}^{\mathsf{T}}\mathbf{r}\|$	$\|\Delta\mathbf{z}\|$
0	1.6094	1.6094	28.000	28.000	12.0000	28.4566	52.4	5.7
1	1.7077	1.0912	30.450	22.887	12.1633	3.4861	4.29	0.233
2	1.6990	1.1148	30.386	23.108	12.1429	3.3987	0.0892	0.0145
3	1.6993	1.1148	30.391	23.095	12.1438	3.3986	0.00278	6.7×10^{-4}
4	1.6993	1.1148	30.391	23.095	12.1438	3.3986	5.5×10^{-5}	3.4×10^{-5}

gradient $\mathbf{J}^{\mathsf{T}}\mathbf{r}$ and the length of the step taken from it, the last two in the mixed units of the entries.

The first step takes twice the objective from 28.46 to 3.49 and the second to within 10^{-4} of its minimum, 3.3986; the step length then falls by a factor of about twenty per iteration. At the minimum the Jacobian of the nine rows, in the columns $(g_1, g_2, Q_1, Q_2, T_s)$, is

$$\mathbf{J}(\hat{\mathbf{z}}_0) = \begin{bmatrix} -11.111 & 0 & 0.366 & 0 & 2 \\ 0 & -15.150 & 0 & 0.656 & 2 \\ 0 & 0 & 0 & 0 & 3.333 \\ 0 & 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0.667 & 0 & 0 \\ 0 & 0 & 0.333 & 0.333 & 3.333 \\ 3.333 & 0 & 0 & 0 & 0 \\ 0 & 3.333 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

the pattern (22.3) with its values: the hall-2 entry -15.150 is the fitted rise of 7.575 K over σ_T, and the column of g_2 holds it and the prior's 3.333 alone.

Table 22.4 gives the fits of the three records. Each row is an entry of the state, with its MAP value and the posterior standard deviation $\sqrt{\Sigma_{ii}}$ for each record, and the conductances $\mathrm{e}^{\hat{g}_i}$ beside the log-conductances; the last row is twice the minimized objective.

Table 22.4: The fits without bias of the three records: MAP entries with $\sqrt{\Sigma_{ii}}$, and the conductances $\hat{G}_i = \mathrm{e}^{\hat{g}_i}$ in kW/K.

Entry	Clean record	Record (a)	Record (b)
g_1 [–] ($\hat{G}_1$)	1.709 ± 0.103 (5.525)	1.801 ± 0.110 (6.054)	1.699 ± 0.103 (5.470)
g_2 [–] ($\hat{G}_2$)	1.413 ± 0.124 (4.106)	1.441 ± 0.133 (4.225)	1.115 ± 0.109 (3.049)
Q_1 [kW]	30.70 ± 1.45	30.64 ± 1.45	30.39 ± 1.45
Q_2 [kW]	22.39 ± 2.23	20.73 ± 2.22	23.10 ± 2.23
T_s [°C]	12.148 ± 0.223	12.698 ± 0.222	12.144 ± 0.223
$2\mathcal{L}(\hat{\mathbf{z}}_0)$ [–]	0.786	4.709	3.399

Fitted to the clean record, the plant is close to the eight-row fit of Table 18.4; the return sensor raises the hall-2 conductance from 4.055 by 1.3 %, because the loop balance now pins the total load and the supply temperature together. Fitted to record (a), the plant moves its supply temperature to 12.698°C, 0.55 K above the clean fit: the supply sensor pulls it up, and the return sensor and the hall readings pull it down. To keep both halls at their readings with a warmer supply, the fit raises the conductances to 6.054 and 4.225 kW/K, and to keep the return temperature near its reading it lowers the hall-2 load to 20.73 kW. Fitted to record (b), the plant lowers the hall-2 conductance from 4.106 to 3.049 kW/K and raises the hall-2 load

from 22.39 to 23.10 kW, and its hall-2 temperature is 19.72°C, against the reading of 19.90 and the clean fit's 17.60: the fit follows the faulty sensor almost all the way. The spreads barely differ between the records; the readings move the state, not its uncertainty.

22.3 The answers (step 4)

The six questions of Table 22.2 fall into four categories of Chapter 2: prediction under a threshold, as a tail of the predictive under the drawing without bias (the first, the second and the fourth); diagnosis, with the bias of each drawing estimated as an entry of its state (the third); estimation under several drawings (the fifth); and design (the sixth). Each is answered from the posteriors of Section 22.2: its engineering question, the general form of its category, the form it takes for the loop, the quantities of the posterior it uses, and the loop's numbers. The questions name the threshold of the residual test and, for the design question, the candidate rows. Every answer is computed from the posteriors of the state of step 3: the residual tests at the linearization of the fit $(\hat{\mathbf{z}}_0, \boldsymbol{\Sigma})$, and each Bayes factor as the ratio of the two drawings' Laplace evidences, each with the curvature at its own MAP state. Where an exact value exists, the evidence by quadrature over the log-conductances or the rate on replicate records, it is set beside the reported one in the same answer, with the gap.

22.3.1 The misfit and the largest normalized residual

Engineering questions. Do the readings disagree with the model? Is this reading lying? Which one? The first is asked of the record as a whole; the second of each reading, with the leading suspect named when one stands out.

General form. Both are tails of a predictive, equations (2.13) and (2.12): the probability, under the drawing without bias, that a replicate record y_{new} of the same plant would show a statistic t at least as large as a level t^*,

$$\Pr(t(y_{\text{new}}) \geq t^* \mid M_0) = \int \mathbf{1}\big[t(y_{\text{new}}) \geq t^*\big]\, p(y_{\text{new}} \mid M_0)\, \mathrm{d}y_{\text{new}},$$

with t^* the observed value for a detection check and a fixed threshold for a test.

Specialized form. Both questions are the tail of equations (2.13) and (2.12) written above, with y_{new} a replicate of the nine rows, $p(y_{\text{new}} \mid M_0)$ the predictive of the drawing without bias under its posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}_0, \boldsymbol{\Sigma})$ with the rows linearized at $\hat{\mathbf{z}}_0$, the statistic t the misfit $2\mathcal{L}(\hat{\mathbf{z}}_0)$ (with t^* its observed value) or the largest normalized residual (with $t^* = r^*$ a threshold). One modeling choice enters here. The family false-alarm probability of the nine statistics is controlled by the Bonferroni bound: this sets the threshold without their joint distribution, and it errs toward fewer alarms because the statistics are correlated; the clean record of Section 22.4.1 and the draws of Section 22.4.2 are the false-alarm check.

For the detection check $t = 2\mathcal{L}(\hat{\mathbf{z}}_0)$, and under the model it is χ^2_ν with $\nu = 4$ (Section 22.2.1). For $\nu = 4$ the tail has a closed form: the density of χ^2_4 is $x\,\mathrm{e}^{-x/2}/4$, and integrating by parts,

$$\Pr(\chi^2_4 > x) = \int_x^\infty \frac{t}{4}\,\mathrm{e}^{-t/2}\,\mathrm{d}t = \frac{x}{2}\,\mathrm{e}^{-x/2} + \int_x^\infty \frac{1}{2}\,\mathrm{e}^{-t/2}\,\mathrm{d}t = \Big(1 + \frac{x}{2}\Big)\mathrm{e}^{-x/2}. \tag{22.11}$$

For the test $t = \max_k |r^{\mathrm{N}}_k|$, the largest normalized residual, with $r^{\mathrm{N}}_k = r_k(\hat{\mathbf{z}}_0)/\sqrt{1 - K_{kk}}$ the standardized residual of row k at the fit divided by the root of its redundancy number. Under the model each r^{N}_k is standard normal (Section 22.2.1), so one row crosses a threshold r^* with

probability $2[1-\Phi(r^*)]$; the probability that any of n_r rows crosses is at most the sum of the individual probabilities, the Bonferroni bound, whatever their correlation:

$$\Pr\Big(\max_k |r_k^{\mathrm{N}}| > r^*\Big) \le \sum_{k=1}^{n_r} \Pr\big(|r_k^{\mathrm{N}}| > r^*\big) = 2n_r\big[1-\Phi(r^*)\big].$$

Setting the bound equal to a chosen family false-alarm probability p_{fa} and solving,

$$r^* = \Phi^{-1}\!\left(1 - \frac{p_{\mathrm{fa}}}{2n_r}\right). \tag{22.12}$$

The bound is conservative, because the normalized residuals are correlated and only ν of them are free: the two rows of a critical pair (Section 22.2.2) cross together and count once in the true false-alarm probability but twice in the bound. A row with redundancy number near zero contributes its full share to the bound while it can never cross for a bias of practical size.

Once a row is flagged the identification rule removes it, refits, and repeats the test on the remaining rows, until no normalized residual exceeds the threshold or the redundancy is exhausted. Each removal lowers $2\mathcal{L}$ by the square of the removed row's normalized residual, by the deletion identity of Section 22.2.2, and the refit's misfit is checked against $\chi^2_{\nu-1}$. The rule identifies correctly when the bad row's normalized residual exceeds every other row's, and it fails in two ways: when the bias spreads so strongly into a neighbor that the neighbor's normalized residual is larger, and when the bad row belongs to a critical pair, in which case the rule has no basis for choosing.

Building blocks. The MAP state $\hat{\mathbf{z}}_0$, the Jacobian $\mathbf{J}$ at it, the residuals $r_k(\hat{\mathbf{z}}_0)$, the hat matrix and the redundancy numbers $1-K_{kk}$ (Section 22.2.1); the inverse normal Φ^{-1}. The fits are computed in Section 22.2.3. For the rate beside the bound, replicate records drawn at the clean fit.

The misfits and the normalized residuals. By equation (22.11) the detection tails on four degrees of freedom are, for the clean record, $(1+0.393)\mathrm{e}^{-0.393} = 0.940$; for record (a), $\Pr(\chi^2_4 > 4.709) = 3.354 \times 0.09497 = 0.3185$; and for record (b), $(1+1.699)\,\mathrm{e}^{-1.699} = 0.493$. The 95 % point of χ^2_4 is 9.49. The check rejects none of the three, and it passes record (b) more easily than record (a), although the bias of record (b) is 2.2 times larger in kelvin and 4.4 standard deviations of its sensor.

By equation (22.12) for the nine rows, $p_{\mathrm{fa}} = 0.05$ gives $r^* = 2.77$, 0.025 gives 2.99 and 0.01 gives 3.26; the threshold of 3 used in practice corresponds to $9 \times 2[1-\Phi(3)] = 9 \times 0.0027 = 0.024$. The value beside the bound comes from 4000 replicate records drawn from the model at the clean fit and each refitted (NumPy's `default_rng(22)`): the family false-alarm rate is 0.031 at $r^* = 2.77$ and 0.014 at 3, below the bounds of 0.050 and 0.024 by 0.019 and 0.010. The bound errs high, as its derivation said, because the nine statistics are correlated and the two rows of a critical pair cross together.

Table 22.5 gives the residual tests for the three records. Each line is one of the nine rows; for each record the columns are the standardized residual $r_k(\hat{\mathbf{z}}_0)$, the redundancy number $1-K_{kk}$ evaluated with the Jacobian at that record's fit, and the normalized residual r_k^{N}. The last line gives $2\mathcal{L}(\hat{\mathbf{z}}_0)$, the sum of the squared standardized residuals, and the sum of the redundancy numbers, which is $\nu = 4$ in every record.

Figure 22.2 draws the normalized residuals as bars, one panel per record. The horizontal axis lists the nine rows, the six readings in blue and the three priors in gold, with the biased reading of the record in red; the vertical axis is the normalized residual, and the dashed lines mark the threshold ± 3. In panel (ii), record (a), the supply sensor leads at -1.99, followed by the return sensor at $+1.66$ and the chiller meter at -1.19. All three lie on the loop balance (22.1): the

Table 22.5: Residual tests of the fit without bias for the three records: standardized residual r_k, redundancy number $1-K_{kk}$ at the record's own fit, and normalized residual $r_k^{\mathrm{N}} = r_k/\sqrt{1-K_{kk}}$. The biased reading of each record is in bold.

	Clean record			Record (a)			Record (b)		
Row	r_k	$1-K_{kk}$	r_k^{N}	r_k	$1-K_{kk}$	r_k^{N}	r_k	$1-K_{kk}$	r_k^{N}
y_{T_1}	0.100	0.079	0.35	0.210	0.094	0.69	0.090	0.079	0.32
y_{T_2}	−0.201	0.078	−0.72	−0.191	0.093	−0.63	**−0.363**	**0.042**	**−1.77**
y_{T_s}	0.161	0.449	0.24	**−1.339**	**0.450**	**−1.99**	0.146	0.448	0.22
y_P	0.292	0.221	0.62	−0.565	0.225	−1.19	0.493	0.224	1.04
y_{Q_1}	−0.201	0.063	−0.80	−0.239	0.068	−0.92	−0.406	0.063	−1.62
y_{T_r}	−0.145	0.448	−0.22	1.118	0.452	1.66	−0.025	0.444	−0.04
prior g_1	0.333	0.882	0.35	0.638	0.865	0.69	0.300	0.882	0.32
prior g_2	−0.656	0.830	−0.72	−0.561	0.803	−0.63	−1.649	0.868	−1.77
prior T_s	0.148	0.950	0.15	0.698	0.951	0.72	0.144	0.950	0.15
$2\mathcal{L}$; sum	0.786	4.000		4.709	4.000		3.399	4.000	

readings say that the water warms by $17.50 - 13.10 = 4.40\,\mathrm{K}$ through the halls, while the chiller meter says it must warm by $52.5/10 = 5.25\,\mathrm{K}$, and the three rows share the missing $0.85\,\mathrm{K}$. No normalized residual crosses 3, nor the Bonferroni threshold 2.77 of a 5 % family false-alarm probability. The test names the supply sensor as the leading suspect and does not flag it. The correlation of the supply and return normalized residuals, by equation (22.5), is −0.83: the two rows nearly share one pattern, and a bias on either looks much like a bias on the other with the sign reversed.

In panel (iii), record (b), the hall-2 reading and the prior on g_2 stand at −1.77 each, the hall-1 meter at −1.62 and the chiller meter at +1.04; nothing crosses the threshold. The equality is exact: the hall-2 reading and the prior on g_2 are the only rows in the column of g_2, a critical pair whose columns of the residual sensitivity matrix are proportional, and the correlation of their normalized residuals is one, equation (22.10). No residual test can say whether the sensor reads high or the path conducts less than the handbook says. Of the misfit of 3.399, 2.72 sits on the prior of the hall-2 conductance, which the fitted $3.05\,\mathrm{kW/K}$ undershoots by 1.65 of its standard deviations; the faulty reading itself shows −0.36. The two meters carry the bias that leaked into them: so that the conductance need not fall as far to produce the temperature rise Q_2/G_2 the reading demands, the fit raised Q_2 by $0.7\,\mathrm{kW}$, which the chiller meter resists (+0.49), and it lowered Q_1 to share the cost with the hall-1 meter (−0.41).

22.3.2 The bias drawings and their evidence

Engineering question. Which sensor is biased, by how much, and how sure is that? The answer is a ranking of the sensors with a probability each, and for the leading one a bias with its band.

General form. Diagnosis, equation (2.15), the ranking of the family by the posterior probability of each member, built from the Bayes factors of equation (2.14),

$$\Pr(M_j \mid \mathbf{y}) \propto p(\mathbf{y} \mid M_j)\,\Pr(M_j), \qquad B_{j0} = \frac{p(\mathbf{y} \mid M_j)}{p(\mathbf{y} \mid M_0)};$$

and for the bias of a drawing, estimation, equation (2.9), with $f(\mathbf{z}) = b_k$ under M_k.

Specialized form. The ranking is equations (2.15) and (2.14) with the family the seven drawings M_j, $j \in \{0, T_1, T_2, T_s, P, Q_1, T_r\}$, each evidence $p(\mathbf{y} \mid M_j)$ the integral of the drawing's

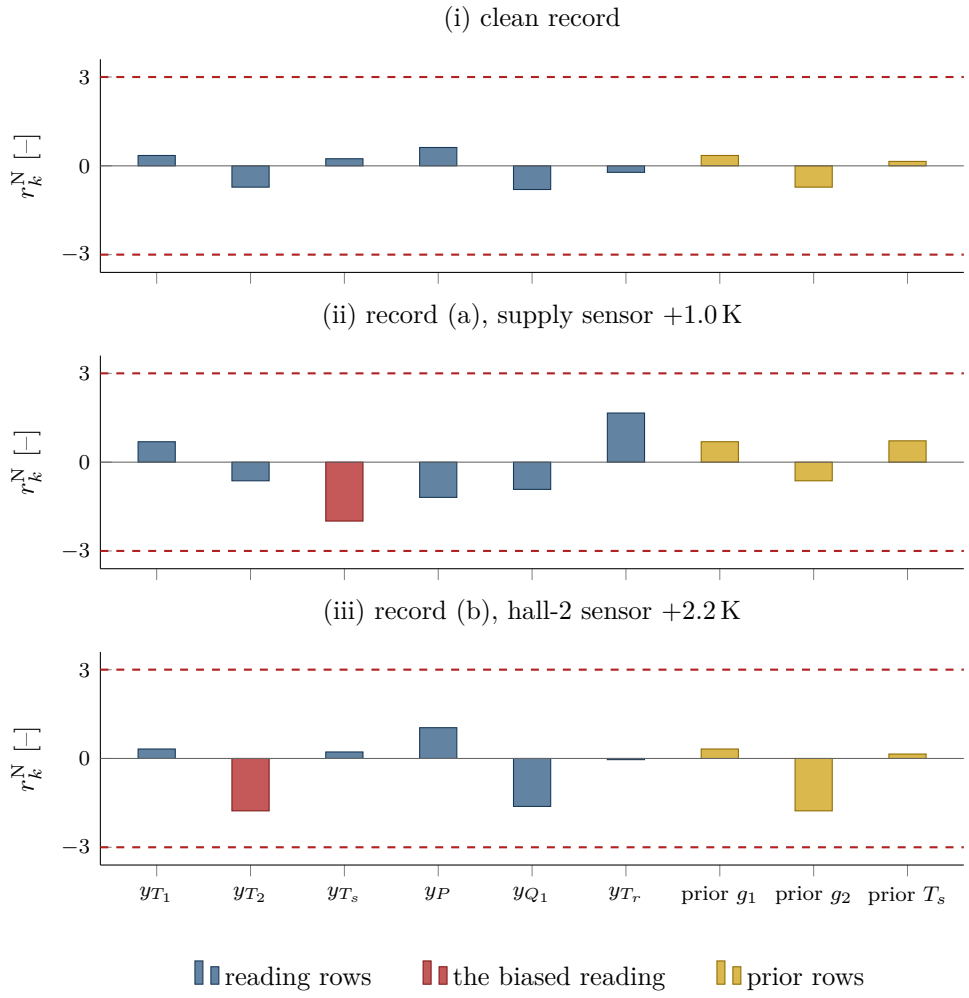


Figure 22.2: Normalized residuals $r_k^{\mathrm{N}} = r_k(\hat{\mathbf{z}}_0)/\sqrt{1 - K_{kk}}$ of the fit without bias, Table 22.5: (i) the clean record, (ii) record (a) with the supply sensor 1.0 K high, (iii) record (b) with the hall-2 sensor 2.2 K high. Blue: readings; gold: priors; red: the biased reading. Dashed: the threshold ±3 of the largest-normalized-residual test.

rows over its state, evaluated by the Laplace form (22.6) at the drawing's posterior of step 3, $(\hat{\mathbf{z}}_j, \mathbf{\Lambda}_j)$; the bias is equation (2.9) with $f(\mathbf{z}) = b_k$, the sixth entry of the state of M_k, and $p(\mathbf{z} \mid \mathbf{y}, M_k)$ the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}_k, \mathbf{\Sigma}_k)$. Two modeling choices enter here. The seven drawings are given equal prior odds: this lets the probabilities follow from the Bayes factors alone, and a maintenance history that makes one sensor more suspect would multiply its odds. The Bayes factor is reported as the ratio of the two drawings' Laplace evidences, each with its curvature at its own MAP state: this follows the curvature of the hall rows as the fitted state moves from one drawing to another, and the exact integral is set beside it.

Against the Laplace evidence (22.6) of M_0, the drawing M_k has the same nine rows with the bias added to row k, one more row, the prior on b_k with constant $-\ln(\sqrt{2\pi}\,\sigma_b)$, and one more entry, which adds $\frac{1}{2}\ln 2\pi$ to the Gaussian integral; its curvature $\mathbf{\Lambda}_k$ is 6×6, with b_k in the units of the reading. The flat-prior constant c_Q and the nine row constants are common to both drawings and cancel in the ratio, and so do the normalizers of the four physics rows of the full state, which both drawings carry:

$$\begin{aligned} \ln B_{k0} = \ln \frac{p(\mathbf{y} \mid M_k)}{p(\mathbf{y} \mid M_0)} &= -\big[\mathcal{L}_k(\hat{\mathbf{z}}_k) - \mathcal{L}_0(\hat{\mathbf{z}}_0)\big] - \frac{1}{2}\ln\frac{\det \mathbf{\Lambda}_k}{\det \mathbf{\Lambda}_0} - \ln(\sqrt{2\pi}\,\sigma_b) + \frac{1}{2}\ln 2\pi \\ &= \big[\mathcal{L}_0(\hat{\mathbf{z}}_0) - \mathcal{L}_k(\hat{\mathbf{z}}_k)\big] - \frac{1}{2}\ln\frac{\det \mathbf{\Lambda}_k}{\det \mathbf{\Lambda}_0} - \ln \sigma_b. \end{aligned} \tag{22.13}$$

The $\sqrt{2\pi}$ of the bias row's normalizer cancels against the $\frac{1}{2}\ln 2\pi$ of the added dimension, and what is left of the added prior is $-\ln \sigma_b$. The logarithm of σ_b carries the unit of the reading and so does the extra factor of $\det \mathbf{\Lambda}_k$, whose b_k column is in inverse units of the reading; their combination is dimensionless.

The same result follows from the other bookkeeping, in which the bias is carried as the normalized entry $b'_k = b_k/\sigma_b$ with a prior row of unit spread. Its constant is then $-\ln\sqrt{2\pi}$, which cancels the $\frac{1}{2}\ln 2\pi$ of the added dimension exactly; but the column of b'_k in the Jacobian is σ_b times the column of b_k, so that $\det \mathbf{\Lambda}'_k = \sigma_b^2 \det \mathbf{\Lambda}_k$ and $-\frac{1}{2}\ln\det \mathbf{\Lambda}'_k = -\frac{1}{2}\ln\det \mathbf{\Lambda}_k - \ln\sigma_b$. The $-\ln\sigma_b$ reappears from the determinant, and the two bookkeepings agree. Read as a whole, equation (22.13) is a gain in fit, the drop in the objective that the bias buys, minus an Occam factor, $\frac{1}{2}\ln(\sigma_b^2 \det \mathbf{\Lambda}_k / \det \mathbf{\Lambda}_0)$, the logarithm of the ratio of the bias's prior width to its posterior width.

Because M_0 is M_k with $b_k = 0$ and the same priors on the other entries, the Bayes factor is also the Savage–Dickey ratio of equation (7.68), $B_{0k} = p(b_k = 0 \mid \mathbf{y}, M_k)/p(b_k = 0 \mid M_k)$: the posterior density of the bias at zero over its prior density there. Chapter 7 warned that the Laplace normal read at the nested value can be poor when the nested value lies in its tail; Section 22.4 measures that error here against the exact integral.

At equal prior odds the probability of a member of the family is its Bayes factor over the sum of all seven, with $B_{00} = 1$,

$$\Pr(M_k \mid \mathbf{y}) = \frac{B_{k0}}{\sum_j B_{j0}}, \qquad j \in \{0, T_1, T_2, T_s, P, Q_1, T_r\}, \tag{22.14}$$

and against M_0 alone, at even odds, it is $B_{k0}/(1 + B_{k0})$. The bias itself is the sixth entry of the state of M_k: its estimate is the entry $\hat{b}_k$ of $\hat{\mathbf{z}}_k$ and its posterior standard deviation $\mathrm{sd}(b_k \mid \mathbf{y}, M_k) = \sqrt{(\mathbf{\Sigma}_k)_{66}}$, read off the fit of the drawing.

The evidence for linear rows. For linear rows the evidence of a bias drawing has a closed form in the quantities of the residual test. Carry the bias as $b'_k = b_k/\sigma_b$, so that its column in the normalized rows is $\mathbf{a} = (\sigma_b/\sigma_k)\mathbf{e}_k$ and its prior row is b'_k itself. Write the rows of M_k as $\mathbf{J}\mathbf{z} + \mathbf{a}b'_k - \tilde{\mathbf{y}}$, with $\tilde{\mathbf{y}}$ the normalized data, and the residual of M_0 at its fit as $\mathbf{r}_0 = -(\mathbf{I} - \mathbf{K})\tilde{\mathbf{y}}$. Define

$$\psi_k = \frac{\sigma_b^2\,(1 - K_{kk})}{\sigma_k^2} = \frac{\sigma_b^2}{\sigma_k^2 + V_{-k}}, \tag{22.15}$$

the prior variance of the bias over the variance of the leave-one-out residual, by equation (22.7). The derivation runs in five steps.

1. For fixed b'_k the minimum over $\mathbf{z}$ of $\|\mathbf{J}\mathbf{z} + \mathbf{a}b'_k - \tilde{\mathbf{y}}\|^2$ is $\|(\mathbf{I}-\mathbf{K})(\mathbf{a}b'_k - \tilde{\mathbf{y}})\|^2$, the part of the vector the columns of $\mathbf{J}$ cannot absorb.
2. Expanding with $(\mathbf{I}-\mathbf{K})^2 = \mathbf{I}-\mathbf{K}$, $\mathbf{a}^\mathsf{T}(\mathbf{I}-\mathbf{K})\mathbf{a} = \psi_k$, $\mathbf{a}^\mathsf{T}(\mathbf{I}-\mathbf{K})\tilde{\mathbf{y}} = -(\sigma_b/\sigma_k)r_{0,k}$ and $\tilde{\mathbf{y}}^\mathsf{T}(\mathbf{I}-\mathbf{K})\tilde{\mathbf{y}} = 2\mathcal{L}_0(\hat{\mathbf{z}}_0)$, and adding the prior row, $2\mathcal{L}_k(b'_k) = 2\mathcal{L}_0(\hat{\mathbf{z}}_0) + 2(\sigma_b/\sigma_k)r_{0,k}\,b'_k + (1+\psi_k)\,b'^2_k$.
3. The minimum is at $\hat{b}'_k = -(\sigma_b/\sigma_k)r_{0,k}/(1+\psi_k)$, where $2\mathcal{L}_0 - 2\mathcal{L}_k = (\sigma_b/\sigma_k)^2 r^2_{0,k}/(1+\psi_k) = \psi_k\big(r^{\mathrm{N}}_k\big)^2/(1+\psi_k)$, since $(\sigma_b/\sigma_k)^2 r^2_{0,k} = \psi_k\, r^2_{0,k}/(1-K_{kk})$.
4. The curvature of M_k in $(\mathbf{z}, b'_k)$ has the block $\mathbf{\Lambda}_0$ for the state, the column $\mathbf{J}^\mathsf{T}\mathbf{a}$ coupling the state to the bias and the entry $1+\sigma_b^2/\sigma_k^2$ for the bias; its determinant is $\det\mathbf{\Lambda}_0$ times the Schur complement $1+\sigma_b^2/\sigma_k^2 - \mathbf{a}^\mathsf{T}\mathbf{K}\mathbf{a} = 1+\psi_k$.
5. In the normalized bookkeeping no constant remains, and equation (22.13) becomes
$$\ln B_{k0} = \frac{1}{2}\left[\frac{\psi_k}{1+\psi_k}\big(r^{\mathrm{N}}_k\big)^2 - \ln(1+\psi_k)\right]. \tag{22.16}$$

Returning to the units of the reading, the bias and its posterior standard deviation are

$$\hat{b}_k = \frac{\sigma_b^2\,(y_k - \hat{h}_k)}{\sigma_k^2 + \sigma_b^2(1-K_{kk})}, \qquad \mathrm{sd}(b_k \mid \mathbf{y}, M_k) = \frac{\sigma_b}{\sqrt{1+\psi_k}}, \tag{22.17}$$

with $\hat{h}_k$ the fitted value of reading k under M_0; the Savage–Dickey ratio with this Gaussian posterior gives the same equation (22.16), as it must for linear rows.

Equation (22.16) ties the two methods together. The evidence depends on the record only through r^{N}_k, the statistic of the residual test, and on the design and the bias prior only through ψ_k. It is positive when

$$\big(r^{\mathrm{N}}_k\big)^2 > \frac{1+\psi_k}{\psi_k}\,\ln(1+\psi_k),$$

a threshold that grows slowly with ψ_k. As $\sigma_b \to \infty$ the drawing becomes the deletion of row k: the fit gain tends to $(r^{\mathrm{N}}_k)^2/2$, the deletion identity (22.9), while the Occam factor $\frac{1}{2}\ln(1+\psi_k)$ grows without limit, so that a flat prior on the bias never wins. Equally, $\frac{1}{2}[\psi_k/(1+\psi_k)](r^{\mathrm{N}}_k)^2$ bounds the evidence from above, so a small ψ_k caps it at a small share of the squared normalized residual. And as $1-K_{kk} \to 0$, $\psi_k \to 0$ and $\ln B_{k0} \approx \frac{1}{2}\psi_k[(r^{\mathrm{N}}_k)^2 - 1] \to 0$: for a critical measurement the evidence is exactly indifferent, neither for nor against a bias, the evidence's form of the zero residual of Section 22.2.2.

Building blocks. The fits of M_0 and of each M_k, the determinants of their information matrices and the minimized objectives (Section 22.2.1); the closed form for linear rows (Section 22.3.2), which ties the Bayes factor to the normalized residual and sets the threshold below. For the values beside them, the exact evidence: the entries that enter the rows linearly, Q_1, Q_2, T_s and the bias, integrated in closed form for each pair (g_1, g_2), a Gaussian integral of the kind of equation (19.9), and the sum over a grid of 121×121 points spanning seven posterior standard deviations in each log-conductance.

The evidence ranking of the bias drawings. Table 22.6 gives the evidence ranking. Each line is one drawing of the family; for each biased record the columns are the fitted bias with its posterior standard deviation, twice the minimized objective of the drawing, the log Bayes factor against M_0 by equation (22.13), and the posterior probability of the drawing over the seven members of the family by equation (22.14); the last column gives the log Bayes factor for the clean record. The first line is M_0 itself.

In record (a) the supply drawing leads, with a bias of 0.85 ± 0.44 K against the 1.0 built into the record and $\ln B_{k0} = +0.23$. The arithmetic of equation (22.13) is $\mathcal{L}_0 - \mathcal{L}_{T_s} = 2.3543 - 0.4592 =$

Table 22.6: The bias drawings ranked by evidence. $\hat{b}_k \pm \mathrm{sd}(b_k \mid \mathbf{y}, M_k)$ in K for the temperature sensors and in kW for the meters; $2\mathcal{L}_k$ the minimized misfit of the drawing; $\ln B_{k0}$ by equation (22.13) with $\sigma_b = 2$; $\Pr(M_k \mid \mathbf{y})$ over the seven drawings at equal prior odds.

	Record (a), supply +1.0 K				Record (b), hall 2 +2.2 K				Clean
Drawing	$\hat{b}_k \pm \mathrm{sd}$	$2\mathcal{L}_k$	$\ln B_{k0}$	Pr	$\hat{b}_k \pm \mathrm{sd}$	$2\mathcal{L}_k$	$\ln B_{k0}$	Pr	$\ln B_{k0}$
M_0	–	4.71	0	0.15	–	3.40	0	0.17	0
M_{T_1}	-0.66 ± 1.34	4.44	−0.36	0.10	-0.31 ± 1.36	3.34	−0.40	0.11	−0.40
M_{T_2}	0.62 ± 1.19	4.46	−0.29	0.11	$\mathbf{1.89 \pm 1.37}$	1.95	**+0.61**	**0.30**	−0.20
M_{T_s}	$\mathbf{0.85 \pm 0.44}$	0.92	**+0.23**	**0.19**	-0.09 ± 0.44	3.35	−1.48	0.04	−1.48
M_P	0.91 ± 1.81	4.45	+0.04	0.16	-0.79 ± 1.81	3.20	−0.01	0.16	−0.07
M_{Q_1}	0.57 ± 1.89	4.62	−0.01	0.15	0.95 ± 1.90	3.14	+0.08	0.18	−0.02
M_{T_r}	-0.70 ± 0.44	2.08	−0.13	0.13	0.02 ± 0.44	3.40	−1.52	0.04	−1.51

1.8951 for the gain in fit, $\frac{1}{2}\ln(\det \mathbf{\Lambda}_{T_s} / \det \mathbf{\Lambda}_0) = 0.9749$ and $\ln 2 = 0.6931$ for the Occam factor, and $1.8951 - 0.9749 - 0.6931 = 0.2270$. The return drawing comes second among the biases with -0.70 ± 0.44 K, a return sensor reading low by about the same amount, and $\ln B_{k0} = -0.13$; the chiller-meter drawing third, +0.04. Both the supply and the return drawing repair the loop balance, and the difference between them lies elsewhere. Under M_{T_s} the fitted supply temperature is 12.23°C and the conductances 5.60 and 4.12 kW/K, the plant of the clean record. Under M_{T_r} the supply temperature is 12.96°C, close to the reading, and the conductances 6.36 and 4.66 kW/K; its prior rows pay for it, the set point at 0.96 and the hall-1 conductance at 0.80 of their standard deviations, and its misfit is 2.08 against 0.92. The supply sensor is preferred because a biased supply sensor leaves the rest of the plant where the handbook and the set point expect it; the same asymmetry is what makes the supply sensor's normalized residual the larger of the two, 1.99 against 1.66, so the residual test already ranks them in this order. A log Bayes factor of 0.23 is a ratio of 1.26; at even prior odds between M_0 and M_{T_s} it gives $\Pr(M_{T_s} \mid \mathbf{y}) = 1.26/2.26 = 0.56$, and over the family of seven it gives 0.19, against 0.15 for no bias and 0.13 for the return sensor: the supply sensor and its bias are found, but not firmly. The values beside them come from the exact evidence: $\ln B_{k0} = 0.233$ for the supply drawing, 0.006 nats above the reported 0.227; −0.129 for the return drawing (−0.126 reported) and 0.033 for the chiller meter (0.038). The order is the same.

In record (b) the hall-2 drawing comes first, $\ln B_{k0} = +0.61$, with a bias of 1.89 ± 1.37 K; the hall-1 meter follows at +0.08 and every other drawing is at or below zero. At even odds against no bias the hall-2 drawing has probability 0.65; over the family, 0.30 against 0.17 for no bias. The bias is estimated with the right sign and nearly the right size, but its band, about ±2.7 K at 95 %, reaches below zero: barely distinguishable from no bias. The value beside it comes from the exact evidence: $\ln B_{k0} = 0.681$, 0.08 nats above the reported 0.605, a Bayes factor of 1.98 against 1.83 and, at even odds against no bias, a probability of 0.66 against 0.65; for the hall-1 meter, 0.072 against 0.081. The gap is the largest of the chapter because the hall-2 drawing's posterior along the unobservable direction, which the priors alone hold, is the least Gaussian; it changes no ranking. On the clean record the hall-2 drawing has −0.190 exactly against −0.204 reported.

Figure 22.3 draws the ranking. The horizontal axis lists the six bias drawings, one group of bars per drawing; the vertical axis is $\ln B_{k0}$ in nats; the gray bars are the clean record, the blue bars record (a) and the gold bars record (b). In each biased record the drawing on the biased sensor is the tallest bar, and only it stands clearly above zero; in the clean record every bar lies below zero.

The closed form (22.16) sets the size of normalized residual at which a bias drawing can win at all: $|r_k^{\mathrm{N}}| > 1.79$ for the supply sensor, whose $\psi_{T_s} = (2/0.3)^2 \times 0.4505 = 20.0$, and $|r_k^{\mathrm{N}}| > 1.13$

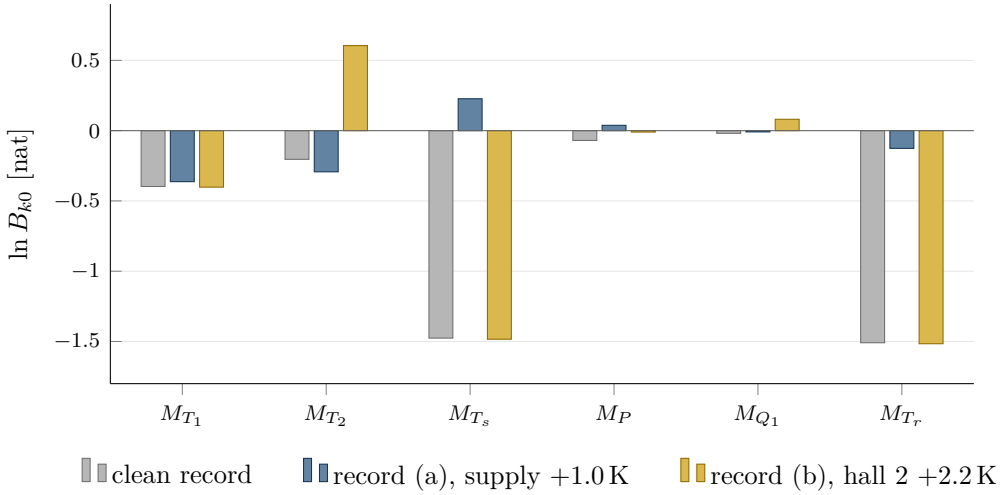


Figure 22.3: Log Bayes factors $\ln B_{k0}$ of the six bias drawings against the drawing without bias, equation (22.13) with $\sigma_b = 2$, Table 22.6. Gray: the clean record; blue: record (a); gold: record (b).

for the hall-2 sensor of record (b), whose $\psi_{T_2} = (2/0.5)^2 \times 0.042 = 0.67$. The supply sensor's 1.99 barely clears its threshold, and the hall-2 sensor's 1.77 clears its own easily but gains little from it, for the reason of Section 22.3.3.

22.3.3 The detectability of a bias

Engineering question. Can a bias on the hall-2 sensor be detected at all? The question asks, before any record, how large a bias must be for the residual test to see it, and why a bias below that size is invisible.

General form. A prediction under a threshold, equation (2.12), taken of the predictive of the test statistic when the reading carries a bias Δ:

$$\Pr(|r_k^{\mathrm{N}}| > r^* \mid \Delta) = \int \mathbf{1}\left[|r_k^{\mathrm{N}}(y_{\mathrm{new}})| > r^*\right] p(y_{\mathrm{new}} \mid \Delta)\, \mathrm{d}y_{\mathrm{new}}.$$

Specialized form. The question is equation (2.12) in the form written above, with y_{new} a replicate record whose reading k carries the bias Δ, the statistic $|r_k^{\mathrm{N}}|$, the threshold r^*, and $p(y_{\mathrm{new}} \mid \Delta)$ the predictive of the rows under the posterior of step 3 linearized at its MAP state, through which the residuals are the linear map $\mathbf{I} - \mathbf{K}$ of the errors. Two modeling choices enter here. The bias is a fixed offset added to the reading, $y_k = h_k(\mathbf{z}) + \Delta + \varepsilon_k$, the drawing M_k with $b_k = \Delta$ known: this makes the statistic a shifted standard normal, and a bias that grows with the reading or drifts during the record would need a row of its own. The variance V_{-k} with which the other rows predict the reading is taken at the refit without row k, the route Chapter 21 settled for the held-out reading: this makes the minimum detectable bias independent of the reading it is to judge, since the fit that predicts the reading does not contain it, and it fails as far as the hall rows, whose derivatives carry the temperature rise, are curved between the refit and the plant. The hat matrix at the fit with the row gives the same value for linear rows; for the hall rows it is reported beside the refit as the fit linearization (Section 22.3.3).

In the normalized data vector of Chapter 18 the bias adds $(\Delta/\sigma_k)\mathbf{e}_k$, with $\mathbf{e}_k$ the k-th unit vector, so that the residuals at the fit become

$$\mathbf{r}(\hat{\mathbf{z}}_0) = -(\mathbf{I} - \mathbf{K})\,\boldsymbol{\zeta} - \frac{\Delta}{\sigma_k}\,(\mathbf{I} - \mathbf{K})\,\mathbf{e}_k,$$

with $\boldsymbol{\zeta}$ the normalized errors. The bias shifts the mean of every residual through column k of the residual sensitivity matrix. Row k's mean is $-(\Delta/\sigma_k)(1-K_{kk})$, and row l's is $(\Delta/\sigma_k)K_{lk}$: the bias spreads to the other rows' residuals through the off-diagonal entries of the hat matrix, which is how one bad reading makes good ones look suspect, and when two or more readings are bad their effects add in the same way, the case state-estimation practice calls *interacting bad data*. Dividing row k's shift by $\sqrt{1-K_{kk}}$, and writing the redundancy number through the variance V_{-k} with which the other rows predict the reading (Section 22.2.2),

$$\mathbb{E}\big[r_k^{\mathrm{N}}\big] = -\frac{\Delta}{\sigma_k}\sqrt{1-K_{kk}} = -\frac{\Delta}{\sqrt{\sigma_k^2+V_{-k}}}, \qquad \Pr(\text{detect} \mid \Delta) \approx \Phi\left(\frac{|\Delta|}{\sqrt{\sigma_k^2+V_{-k}}} - r^*\right), \tag{22.18}$$

the second neglecting the far tail of the opposite sign. A test that flags $|r_k^{\mathrm{N}}| > r^*$ sees the bias on average when the mean reaches the threshold. The *minimum detectable bias* at that threshold is

$$b_k^* = \frac{r^*\sigma_k}{\sqrt{1-K_{kk}}} = r^*\sqrt{\sigma_k^2+V_{-k}}, \tag{22.19}$$

detected about half the time; a bias of $(r^*+0.84)\sqrt{\sigma_k^2+V_{-k}}$ is detected with probability 0.8, since the normalized residual then exceeds the threshold unless its standard normal error falls below -0.84. Equation (22.19) states the limit in the units of the reading: a bias is visible when it exceeds a few times the spread with which the sensor's own noise and the rest of the plant together predict the reading.

The bias absorbed by an entry. Under M_{T_2} the state is $(g_1, g_2, Q_1, Q_2, T_s, b_{T_2})$, and by the pattern (22.3) only two rows touch g_2, the hall-2 reading and the prior on g_2, and only two touch b_{T_2}, the hall-2 reading and the bias prior. The hall-2 reading row compares $T_s + Q_2\mathrm{e}^{-g_2} + b_{T_2}$ with the reading; its Jacobian entries in the two columns are $-(T_2-T_s)/\sigma_T$ and $1/\sigma_T$, with $T_2 - T_s = Q_2\mathrm{e}^{-g_2}$ the rise. Move the state along

$$\mathbf{d} = (\Delta g_2,\ \Delta b_{T_2}) = (1,\ T_2 - T_s) \tag{22.20}$$

with the other entries fixed: the row changes by $-(T_2-T_s)+(T_2-T_s)=0$ to first order, and no other reading involves g_2 or b_{T_2}. The direction lies in the null space of the design columns, equation (19.23): it is unobservable, and by the rule of Chapter 19, equation (19.24), only the two priors hold it, the conductance prior of spread σ_g, worth $\sigma_g(T_2-T_s)$ kelvin along the direction, and the bias prior of spread σ_b. A reading of small redundancy number and its own bias are a single unidentified direction: any bias can be traded for a change of conductance that leaves every reading as it was. The posterior ellipse of (g_2, b_{T_2}) is long along $\mathbf{d}$, and its short axis is the combination the reading does fix, the hall-2 temperature $T_s + Q_2\mathrm{e}^{-g_2}$ seen through the sensor.

The same direction sets the redundancy number of the hall-2 reading. If the reading were checked only by the conductance prior, its prediction from the other rows would carry the variance $(T_2-T_s)^2\sigma_g^2$, the prior mapped through the rise, and the redundancy number of Section 22.2.2 would be

$$1 - K_{22} \approx \frac{\sigma_T^2}{\sigma_T^2 + (T_2-T_s)^2\sigma_g^2}, \tag{22.21}$$

which falls as the rise grows. A bias that raises the reading raises the fitted rise, which makes the reading look more informative and less checked, so that a larger bias lowers the redundancy number by which it would be detected. The spreads of Q_2 and T_s add to the prediction's variance, so the exact number lies a little below the approximation.

Building blocks. The hat matrix at the fit and V_{-k} (Sections 22.2.1 and 22.2.2); the threshold r^*; the fitted rise $T_2 - T_s$; the covariance $\boldsymbol{\Sigma}_{T_2}$ of M_{T_2} and its eigenvectors. For the rate beside the reported one, replicate records drawn at the clean fit.

The detectability of the readings. Table 22.7 lists, for the six readings of the clean record, the redundancy number the readings alone would give without the three priors, then the standard deviation $\sqrt{V_{-k}}$ with which the other eight rows, refitted without the reading, predict it and the minimum detectable bias (22.19) on that route at $r^* = 3$, for detection half the time and with probability 0.8; the last two columns give the same two biases at the fit linearization, from the redundancy numbers of Table 22.5.

Table 22.7: Minimum detectable biases of the six readings, clean record, $r^* = 3$. "Readings only": the redundancy number of the design without the three priors, $\nu = 1$. Refit: V_{-k} from the fit without reading k, the route reported. Fit linearization: $V_{-k} = \sigma_k^2 K_{kk}/(1 - K_{kk})$ from the hat matrix at the clean fit. 50 % and 80 %: the biases detected with those probabilities.

		Refit			Fit linearization	
Reading	readings only [–]	$\sqrt{V_{-k}}$	50 %	80 %	50 %	80 %
hall-1 temperature y_{T_1}	0	1.88 K	5.8 K	7.5 K	5.3 K	6.8 K
hall-2 temperature y_{T_2}	0	1.39 K	4.4 K	5.7 K	5.4 K	6.9 K
supply temperature y_{T_s}	0.409	0.33 K	1.34 K	1.72 K	1.34 K	1.72 K
chiller heat meter y_P	0.182	3.81 kW	12.9 kW	16.5 kW	12.8 kW	16.3 kW
hall-1 load meter y_{Q_1}	0	5.92 kW	18.3 kW	23.5 kW	17.9 kW	22.9 kW
return temperature y_{T_r}	0.409	0.33 K	1.34 K	1.72 K	1.34 K	1.72 K

Without the priors the readings split into two kinds. The two hall sensors and the hall-1 meter are critical measurements: each is the only reading that informs one entry, g_1, g_2 and the split of the load, and its redundancy number is zero. The supply sensor, the return sensor and the chiller meter form a redundant set tied by the loop balance, with redundancy numbers 0.41, 0.41 and 0.18 summing to the single degree of freedom. The priors lift the critical readings off zero, to between 0.06 and 0.08, which is enough to give them a residual and not enough to check them: a hall-2 bias must be about 4.4 K, and a hall-1 bias 5.8 K, before the sensor's residual reaches 3 half the time, and a hall-1 meter error 18 kW. The loop's three sensors are well checked by comparison, and the supply and return sensors would show a bias of 1.34 K half the time. The 1.0 K bias of record (a) lies below that, and the 2.2 K of record (b) far below the hall-2 sensor's 4.4: the two outcomes of the residual tests were predictable from the table before either record was drawn.

For the hall-2 reading the refit without it is the row-removed fit of Section 22.3.4: the other eight rows predict the hall-2 temperature at 16.53°C with a standard deviation of 1.391 K, and the difference between the reading and that prediction has the spread $\sqrt{0.5^2 + 1.391^2} = 1.479$ K. The prediction is the same for the clean record and for record (b), since the only reading that differs is the one left out. By equation (22.19) the minimum detectable bias is $3 \times 1.479 = 4.44$ K half the time and $3.84 \times 1.479 = 5.68$ K with probability 0.8; by equation (22.18) the 2.2 K bias of record (b) shifts the expected statistic by $2.2/1.479 = 1.49$, half the threshold, and record (b)'s own leave-one-out residual is $(16.53 - 19.90)/1.479 = -2.28$. Chapter 21, without the return sensor, found that its held-out check flags a hall-2 bias half the time at 4.5 K, three predictive standard deviations of 1.503 K, and 80 % of the time at 5.8 K, and by the linear identity at the full fit 5.5 and 7.1 K; the return sensor narrows the prediction only to 1.479 K, because it informs the loads and the supply temperature and not the conductance that carries the prediction's variance.

The value beside the reported one comes from replicate records: of 4000 records drawn from the model at the clean fit, each refitted (NumPy's `default_rng(22)`), with 4.5 K added to the hall-2 reading, the leave-one-out residual crosses 3 in 0.39 of them, against $\Phi(4.5/1.479 - 3) = 0.52$ on the reported route, a gap of 0.13 in probability. The route errs through the point at which the hall-2 row is linearized: the draws are made at the clean fit, rise 5.45 K, while the refit sits

at the handbook conductance, rise 4.38 K, and at the draws' own predictive spread of 1.73 K the same formula gives $\Phi(4.5/1.73 - 3) = 0.35$. The normalized residual at the fit crosses in only 0.14 of them.

The fit linearization gives larger values for the hall-2 reading, 5.4 K at the clean fit and, at the fit of record (b), where the redundancy number is 0.042, a prediction spread of $0.5\sqrt{0.958/0.042} = 2.39$ K and $3\times0.5/\sqrt{0.042} = 7.3$ K. The difference is the point at which the hall-2 row's derivative, the rise $T_2 - T_s$, is taken: 4.38 K at the refit, where the conductance sits at its handbook 5.00 kW/K, against 5.45 K at the clean fit and 7.58 K at the biased fit, where the reading itself sets it. The two-row approximation (22.21) follows the rise: $0.25/(0.25 + 7.58^2 \times 0.09) = 0.046$ at the biased fit and $0.25/(0.25 + 5.45^2 \times 0.09) = 0.086$ at the clean one, against the exact 0.042 and 0.078. A bias that raises the reading lowers the redundancy number at the fit by which the normalized residual would detect it, which is why record (b)'s normalized residual, -1.77, is smaller than its leave-one-out residual, -2.28. The value $\psi_{T_2} = 0.67$ at the biased fit says the same to the evidence: equation (22.16) can never exceed $\frac{1}{2}[\psi/(1+\psi)](r^{\mathrm{N}})^2$, a fifth of the squared normalized residual.

The bias absorbed by the conductance. Under M_{T_2} for record (b) the fitted rise is 5.74 K, so the unobservable direction (22.20) is $\mathbf{d} = (1,\ 5.74\,\mathrm{K})$, and the conductance prior is worth $0.3 \times 5.74 = 1.72$ K along it against the bias prior's 2 K. Figure 22.4 draws the posterior of the two entries. The horizontal axis is the hall-2 log-conductance g_2 and the vertical axis the bias b_{T_2} in K; the two closed curves are the contours of the Laplace posterior at one and two standard deviations, $\Delta(2\mathcal{L}) = 1$ and 4, about the MAP point, the red star at $g_2 = 1.365$ ($G_2 = 3.92$ kW/K), $b_{T_2} = 1.89$ K. The dashed line is the unobservable direction through the MAP point, $b_{T_2} = 1.89 + 5.74\,(g_2 - 1.365)$. The gold square is the fit without a bias entry, which must sit on $b_{T_2} = 0$; the blue circle marks the clean fit's log-conductance with the 2.2 K the record was built with; the dotted vertical line is the handbook value $g^0 = \ln 5$.

The ellipse is long and tilted: $\mathrm{sd}(b_{T_2} \mid \mathbf{y}) = 1.37$ K, $\mathrm{sd}(g_2 \mid \mathbf{y}) = 0.234$, and $\mathrm{corr}(g_2, b_{T_2}) = 0.86$. Its long axis, of standard deviation 1.39 K, lies 8.4° from the bias axis in these units, and the unobservable direction 9.9° from it, within 1.5°: the long axis is the direction the readings cannot see, shortened only by the two priors. The short axis is the combination the reading fixes, the hall-2 temperature plus the bias, the value the sensor should read: 19.78°C with a standard deviation of 0.49 K, against the hall temperature's own 1.36 K. The fit without bias lies near the long axis where it meets $b_{T_2} = 0$, between the one- and two-standard-deviation contours: it slid along the unidentified direction and put the whole bias into the conductance. The construction point lies well inside the inner contour, so the drawing is right about the truth; it cannot be sure of it, because the conductance's posterior spread, 0.234, is still 78 % of its prior spread 0.3, a ratio (19.25) of 1.28. Were the conductance known, the bias would be determined to 0.71 K, its spread conditional on g_2, and 1.89 K would stand 2.7 standard deviations from zero.

Figure 22.5 draws the redundancy numbers as bars. In both panels the horizontal axis lists the rows and the vertical axis is the redundancy number; readings are blue, priors gold and the hall-2 reading red. Panel (a) is the present design, nine rows summing to four. Panel (b) adds a second hall-2 temperature sensor of the same standard deviation, ten rows summing to five, one of the candidates of Section 22.3.5. In panel (a) the readings fall into the two groups of Table 22.7, the hall sensors and the hall-1 meter near zero and the loop's three sensors well above them, and the priors carry most of the redundancy.

22.3.4 The answers with the reading kept, freed or removed

Engineering question. What does the biased reading do to the answers if it is kept? The answers of Chapter 20, the hall temperatures at a larger load, are asked again under each way

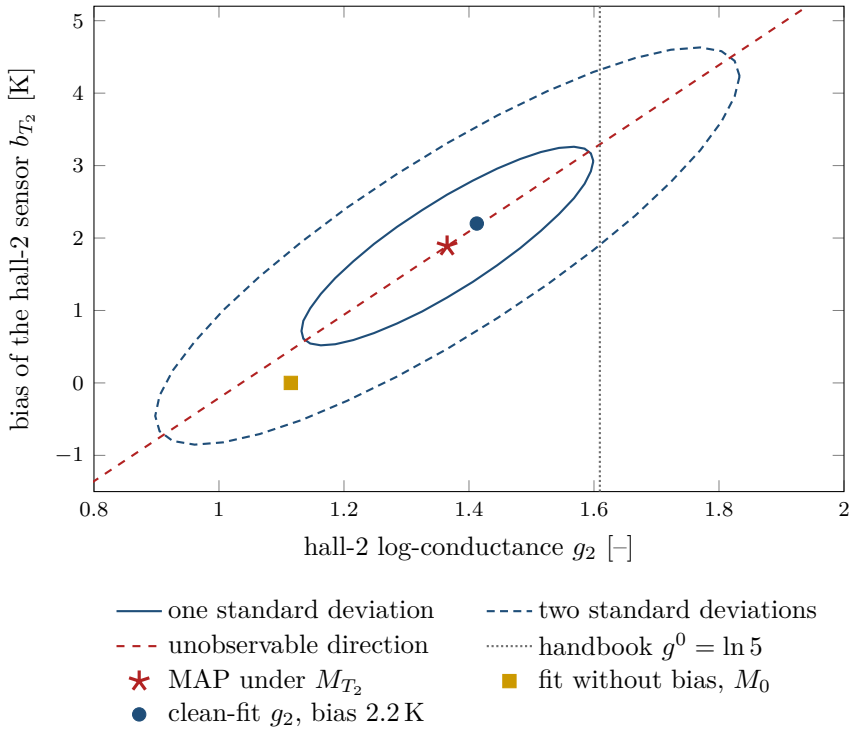


Figure 22.4: Laplace posterior of the hall-2 log-conductance g_2 and the hall-2 sensor's bias b_{T_2} under the drawing M_{T_2} for record (b): mean $(1.365, 1.89\,\mathrm{K})$, posterior standard deviations 0.234 and 1.37 K, correlation 0.86. Solid and dashed blue: contours at one and two standard deviations. Red dashed: the unobservable direction $b_{T_2} = 1.889 + 5.744\,(g_2 - 1.365)$ of equation (22.20). Gold square: the fit without bias, $g_2 = 1.115$ ($G_2 = 3.05\,\mathrm{kW/K}$). Blue circle: $g_2 = 1.4125$ of the clean fit with the 2.2 K built into the record.

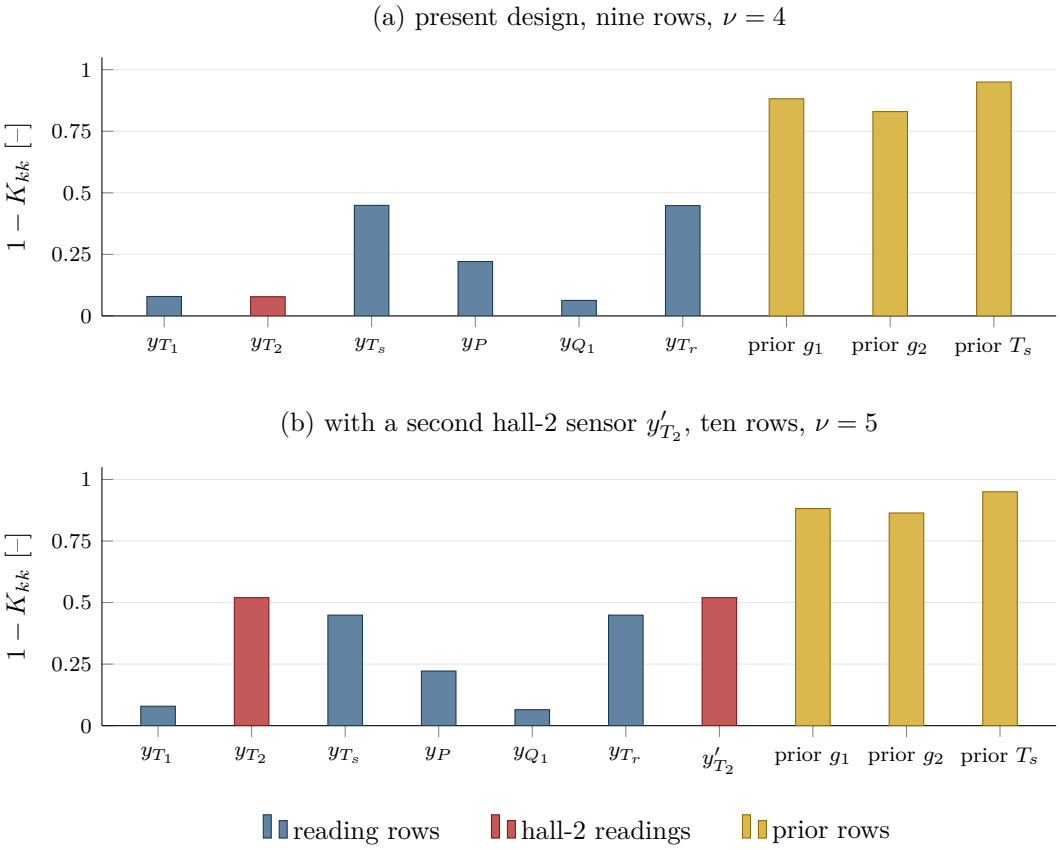


Figure 22.5: Redundancy numbers $1 - K_{kk}$ of the rows at the clean fit. (a) The present design: the hall sensors and the hall-1 meter are nearly critical, the loop's three sensors check each other, the priors carry most of the redundancy; sum 4. (b) With a second hall-2 temperature sensor, $\sigma_T = 0.5\,\mathrm{K}$, on the same balance: both hall-2 readings rise to 0.52; sum 5.

of treating the suspect reading.

General form. Estimation of a derived quantity, equations (2.9) and (2.10),

$$\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}] = \int f(\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

here under each of three drawings.

Specialized form. The question is equations (2.9) and (2.10) with $f(\mathbf{z})$ the scenario temperatures T_2^* and T_1^* below and $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal of each treatment's posterior of step 3: $\mathcal{N}(\hat{\mathbf{z}}_0, \boldsymbol{\Sigma})$ under M_0, $\mathcal{N}(\hat{\mathbf{z}}_{T_2}, \boldsymbol{\Sigma}_{T_2})$ under M_{T_2} and $\mathcal{N}(\hat{\mathbf{z}}_{-2}, \boldsymbol{\Sigma}_{-2})$ without row 2, the mean read as $f(\hat{\mathbf{z}})$ and the spread as $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$. Two modeling choices enter here. The scenario keeps the present conductance and supply temperature at the larger load, as in Chapter 20: this makes each scenario temperature a function of the present state, and it fails if the conductance itself changes with the load. Each answer is carried by the delta method on the reduced state, which Chapter 20 checked against draws for this plant: this makes every band one product $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$.

The three treatments of the hall-2 reading are three drawings fitted to record (b): M_0, which keeps the reading as it is; M_{T_2}, which frees its bias; and the drawing without row 2, which removes it, the identification rule's remedy. The clean record is the reference. The scenario temperatures are

$$T_2^*(\mathbf{z}) = T_s + Q_2^* \mathrm{e}^{-g_2}, \qquad \mathbf{g}_{T_2^*} = \big(0,\ -Q_2^*\mathrm{e}^{-g_2},\ 0,\ 0,\ 1\big)^{\mathsf{T}}, \qquad \mathrm{sd}(T_2^* \mid \mathbf{y}) \approx \sqrt{\mathbf{g}_{T_2^*}^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}_{T_2^*}}, \tag{22.22}$$

with $Q_2^* = 40\,\mathrm{kW}$, and the same with $Q_1^* = 60\,\mathrm{kW}$, g_1 and the first entry for T_1^*; under M_{T_2} the gradient has a sixth entry, zero, and $\boldsymbol{\Sigma}$ is the 5×5 block of $\boldsymbol{\Sigma}_{T_2}$, the marginal of the five plant entries. The present hall-2 temperature is $T_s + Q_2\mathrm{e}^{-g_2}$ in every drawing: under M_{T_2} it is the hall's temperature, not the reading's, and the reading less the fitted bias is a separate number. A kept bias moves the answers through the entry it is absorbed into: if the conductance moves from G_2 to $\hat{G}_2$, the scenario temperature moves by $Q_2^*(1/\hat{G}_2 - 1/G_2)$, an error that grows with the load. Whether the reference value lies inside the band of the kept reading is the normal tail $\Pr(T_2^* \le t_\mathrm{c} \mid \mathbf{y}, M_0) \approx \Phi\big((t_\mathrm{c} - \hat{T}_2^*)/\,\mathrm{sd}(T_2^* \mid \mathbf{y})\big)$, with t_c the clean answer.

Building blocks. The fits of the clean record and of the three drawings of record (b), their covariances and the gradients of equation (22.22).

The cost of keeping the biased reading. Table 22.8 gives the answers of record (b) under the three treatments of the hall-2 reading of Section 22.3.4, with the clean record for reference. Each row is one treatment; the columns give the fitted hall-2 conductance with the posterior standard deviation of its logarithm, the hall-2 temperature at the present load, the scenario temperature T_2^* at $Q_2^* = 40\,\mathrm{kW}$ and the scenario temperature T_1^* at $Q_1^* = 60\,\mathrm{kW}$, each with its delta-method band $\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}$ of equation (22.22).

Kept, the biased reading moves the hall-2 answers by about the bias at the present load and by more at a larger one. The present temperature is reported as $19.72 \pm 0.49°\mathrm{C}$, 2.1 K above the clean value with a band as narrow as the clean one. At 40 kW the fit reports $T_2^* = 25.26 \pm 1.44°\mathrm{C}$ against the clean 21.89 ± 1.21: 3.37 K too high, 2.3 of its own standard deviations, and the clean answer lies outside its 95 % band of 22.43 to 28.09°C; the biased fit gives the clean value or less a probability of $\Phi(-3.37/1.44) = 0.010$. The error grows with the load because it sits in the conductance, which alone accounts for $40\,(1/3.049 - 1/4.106) = 3.38\,\mathrm{K}$, while at the present load of 23 kW it is about 2.1 K. Freeing the bias gives $22.36 \pm 2.39°\mathrm{C}$, within 0.5 K of the clean answer and with a band twice as wide: the width is the unobservable direction of Figure 22.4, and it is the honest statement of what record (b) knows about hall 2. Removing the row returns

Table 22.8: The cost of the hall-2 bias, record (b). Conductance in kW/K; temperatures in °C with their posterior standard deviations; T_2^* at $Q_2^* = 40\,\text{kW}$ and T_1^* at $Q_1^* = 60\,\text{kW}$. Under M_{T_2} the present temperature is the hall's, the reading less the fitted bias being 18.01°C.

Treatment	$\hat{G}_2$	$\text{sd}(g_2 \mid \mathbf{y})$	T_2 now	T_2^* at 40 kW	T_1^* at 60 kW
clean record (reference)	4.106	0.124	17.60 ± 0.48	21.89 ± 1.21	23.01 ± 1.06
record (b), reading kept (M_0)	3.049	0.109	19.72 ± 0.49	25.26 ± 1.44	23.11 ± 1.07
record (b), bias freed (M_{T_2})	3.917	0.234	17.89 ± 1.36	22.36 ± 2.39	23.02 ± 1.07
record (b), row removed	5.000	0.300	16.53 ± 1.39	20.14 ± 2.41	22.94 ± 1.06

the conductance to its prior, 5.00 kW/K, and gives 20.14 ± 2.41°C, as wide and on the other side. The hall-1 answer barely moves, 23.11 against 23.01°C: a bias damages the answers that depend on the entry it is absorbed into, and leaves the rest.

22.3.5 The rows that make a reading checkable

Engineering question. Which row makes the hall-2 sensor checkable? The candidates are a hall-2 load meter, a heat meter on hall 2's coil, a second hall-2 temperature sensor and a commissioning test of hall 2's path.

General form. The design form, equation (2.16), with the preposterior covariance of equation (2.17),

$$U(a) = \mathbb{E}_{\tilde{y}_a \mid \mathbf{y}}\Big[\, u\big(p(\mathbf{z} \mid \mathbf{y}, \tilde{y}_a)\big)\Big], \qquad \boldsymbol{\Sigma}_a = \big(\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Lambda}_a\big)^{-1}.$$

Specialized form. The question is equation (2.16) with the actions a the candidate rows, the value u the redundancy number of the hall-2 reading after the row, and the preposterior covariance $\boldsymbol{\Sigma}_a$ formed from the posterior of step 3 without the hall-2 row, $\boldsymbol{\Sigma}_{-2}$, and the candidate's information $\boldsymbol{\Lambda}_a$. Two modeling choices enter here. Each candidate is a linear Gaussian row with a stated standard deviation, or, for the path test, a replacement of the handbook prior on g_2 by a measured value with a tighter spread: this makes the redundancy number after the row independent of the reading it will give, so it ranks the candidates before any is bought, and a candidate whose error is larger than stated would rank too high. The readings the candidates would give under record (b) are illustrative; they serve only to show what the record would then look like.

The quantity u valued is the redundancy number of the hall-2 reading, and with it the minimum detectable bias (22.19); a larger redundancy number is better. With $\mathbf{j}_k^{\mathsf{T}}$ the row of $\mathbf{J}$ belonging to reading k and $\boldsymbol{\Sigma}_{-k}$ the covariance of the fit without that row, a candidate of Jacobian row $\mathbf{J}_a$ adds its information $\boldsymbol{\Lambda}_a = \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a$, and the variance with which the other rows then predict reading k is

$$V_{-k,a} = \sigma_k^2\, \mathbf{j}_k^{\mathsf{T}} \big(\boldsymbol{\Sigma}_{-k}^{-1} + \boldsymbol{\Lambda}_a\big)^{-1} \mathbf{j}_k, \qquad 1 - K_{kk}(a) = \frac{\sigma_k^2}{\sigma_k^2 + V_{-k,a}}, \tag{22.23}$$

by the relation of Section 22.2.2; nothing in it depends on the outcome. The candidates' Jacobian rows on the reduced state read off the physics. A load meter of spread σ_a on Q_2 has the row $(0,\ 0,\ 0,\ 1,\ 0)/\sigma_a$; a heat meter on hall 2's coil, a water flow meter with the water's temperature rise across the coil, measures the same quantity, $Q_2 = \dot{m}_2 c_p\, \Delta T_2$, and has the same row with its own spread. A second hall-2 temperature sensor has the row of the first, $(0,\ -(T_2 - T_s),\ 0,\ \mathrm{e}^{-g_2},\ 1)/\sigma_T$. A test of the path replaces the information $1/\sigma_g^2$ of the handbook prior in the g_2 entry by $1/\sigma_{g,\text{test}}^2$. By the unobservable direction (22.20), a row that

has no entry in the column of g_2 cannot close it: the load meter and the coil meter inform Q_2, which the plant already knows, while the second sensor and the path test touch the direction itself.

Building blocks. The fit without the hall-2 row and its covariance $\mathbf{\Sigma}_{-2}$; the candidates' Jacobian rows; the fits of record (b) under each design with the illustrative readings.

Rows that make the hall-2 reading checkable. At the clean fit the hall-2 reading's prediction from the other rows has the variance $V_{-k} = 2.97\,\mathrm{K}^2$, of which the conductance prior mapped through the rise contributes $(T_2 - T_s)^2\sigma_g^2 = 5.45^2 \times 0.09 = 2.67$; the spread of the load, already held to 2.2 kW by the chiller meter and the return sensor, contributes most of the rest. Table 22.9 ranks five candidate designs by equation (22.23). Each row is one design; the columns give the redundancy ν, the redundancy number of the hall-2 reading at the clean fit and at the biased fit of record (b), its minimum detectable bias on the refit route of Table 22.7, and what record (b) then shows: the misfit with its tail, the normalized residual of the hall-2 reading and the log Bayes factor of the hall-2 drawing. The candidates are a hall-2 load meter of $\sigma = 1.5$ kW (reading 23.0 kW, illustrative); a heat meter on hall 2's coil of 1.0 kW; a second hall-2 temperature sensor of 0.5 K reading 17.70°C; and a commissioning test of hall 2's path that replaces the handbook prior on g_2 by a measured $G_2 = 4.1$ kW/K with a spread of 0.10 or 0.05 in the logarithm (illustrative).

Table 22.9: Candidate rows and the checkability of the hall-2 reading. $1 - K_{22}$ at the clean fit and at the fit of record (b); $b^*_{T_2}$ at $r^* = 3$ on the refit route, clean record; the last three columns are record (b) under the design. Illustrative readings for the candidates.

		$1 - K_{22}$			Record (b)		
Design	ν	clean	(b)	$b^*_{T_2}$	$2\mathcal{L}$ (tail)	$r^{\mathrm{N}}_{T_2}$	$\ln B_{T_2}$
present	4	0.078	0.042	4.4 K	3.40 (0.49)	−1.77	0.61
hall-2 load meter, 1.5 kW	5	0.082	0.045	4.4 K	3.40 (0.64)	−1.73	0.57
hall-2 coil heat meter, 1.0 kW	5	0.082	0.045	4.4 K	3.40 (0.64)	−1.72	0.56
second hall-2 sensor, 0.5 K	5	0.520	0.514	2.1 K	11.65 (0.040)	−3.28	3.88
path test, $\sigma_g = 0.10$	4	0.291	0.213	2.7 K	6.17 (0.19)	−2.43	1.64
path test, $\sigma_g = 0.05$	4	0.399	0.356	2.4 K	9.10 (0.059)	−2.97	2.84

A meter on the hall-2 load does not help. It raises the redundancy number from 0.078 to 0.082, and at the biased fit from 0.042 to 0.045, leaves the minimum detectable bias at 4.4 K, and the coil heat meter, a row on the same entry, does no better. A second temperature sensor on the same balance raises it to 0.52, close to the redundancy number of 0.51 in Chapter 18's faulty-sensor section, for its nine-row plant with the second sensor at the faulty fit (Section 18.3.2, Exercise 18.6); each hall-2 sensor is then predicted by the other rows to within $0.5\sqrt{0.48/0.52} = 0.48$ K. In panel (b) of Figure 22.5 the two hall-2 sensors rise together to 0.52, the prior on g_2 from 0.830 to 0.864, and the other rows move by less than 0.003. Under record (b) the misfit becomes 11.65 on five degrees of freedom, tail 0.040; the faulty sensor's normalized residual is −3.28, across the threshold, and the good one's +2.85, below it: the largest-normalized-residual test flags the right sensor. The evidence agrees, $\ln B_{k0} = 3.88$ for a bias of 2.05 ± 0.66 K on the faulty sensor against 2.44 for a bias of -1.80 ± 0.66 K on the good one, odds of $\mathrm{e}^{1.44} = 4.2$ to one for the right sensor. For the larger fault of 22.0°C Chapter 18's faulty-sensor case (Section 18.3.2) found a misfit of 40.4 on four degrees of freedom with the same pair. A test of the path is the other way to know the conductance. With a spread of 0.05 it raises the redundancy number to 0.40 and lets the evidence find the bias, $\ln B_{k0} = 2.84$; its

normalized residual, −2.97, equals that of the test row in magnitude, since the test and the reading form a critical pair, and the residual test cannot choose between them; the evidence weighs the sensor bias against the test's stated spread.

Principle 22.2. *A bias on a reading is detectable only when the quantity the reading measures is also known another way. How well it is known is computed from the design before any reading is taken, and it ranks the candidate rows: a row that informs a different entry, however good, does not make the reading checkable; a second path to the same quantity does.*

22.3.6 The computation as an algorithm

The computations of this chapter use the fit of Chapter 18, the curvature of Chapter 19 and one fit per bias drawing.

1. *Inputs.* The rows in normalized form with their kinds (reading or prior); the family of bias drawings, one per reading, with the bias prior σ_b of each in the reading's units, fixed before the record is examined; the threshold r^* from a family false-alarm probability by equation (22.12).
2. *Fit without bias.* The MAP state $\hat{\mathbf{z}}_0$, the Jacobian $\mathbf{J}$ at it, $2\mathcal{L}_0(\hat{\mathbf{z}}_0)$ against χ^2_ν.
3. *Residual tests.* The hat matrix $\mathbf{K} = \mathbf{J}\boldsymbol{\Sigma}\mathbf{J}^\mathsf{T}$ with $\boldsymbol{\Sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}$ by triangular solves; the redundancy numbers $1 - K_{kk}$; the normalized residuals; the largest against r^*; if one crosses, remove the row, refit, and repeat; report critical pairs, rows whose columns of the residual sensitivity matrix are proportional.
4. *Design quantities.* For each reading, V_{-k} from the refit without it (equal to $\sigma_k^2 K_{kk}/(1-K_{kk})$ for linear rows) and the minimum detectable bias $b_k^* = r^*\sqrt{\sigma_k^2 + V_{-k}}$; flag every reading whose b_k^* exceeds the bias that matters to the decisions the reading feeds.
5. *Bias drawings.* For each reading k, fit M_k from $\hat{\mathbf{z}}_0$ with $b_k = 0$; record $\hat{b}_k$, its posterior standard deviation and $\mathcal{L}_k$; form $\ln B_{k0}$ by equation (22.13), with the $-\ln\sigma_b$ of the bias prior; the closed form (22.16) serves as a cross-check.
6. *Ranking.* Posterior probabilities over the family at the stated prior odds, equation (22.14); the leading drawing with its bias and band.
7. *Checks.* The clean or a simulated clean record for false alarms; the exact evidence by quadrature where the bias's posterior is held by priors; the ranking under a halved and a doubled σ_b.

For the plant the algorithm returns Tables 22.5 and 22.6 with seven fits per record, and the design quantities of Table 22.7 from six refits of the clean record.

22.4 The checks (step 5)

The answers rest on a threshold and a detectability set for replicate records, on a route for the evidence, on the bias prior, on the four balances taken as a limit and on two methods that are meant to see the same thing; each can be set against an alternative on the same records.

22.4.1 The clean record

The clean record tests both methods for false alarms. Its misfit is 0.786 on four degrees of freedom, tail 0.940; its largest normalized residual is the hall-1 meter's, −0.80; and every bias drawing has a negative log Bayes factor, from −0.02 for the hall-1 meter to −1.51 for the return sensor (Table 22.6, last column, and the gray bars of Figure 22.3). Over the family, the drawing without bias has the highest posterior probability, 0.21. A bias on the hall-2 sensor, fitted to the

clean record, comes out at 0.73 ± 1.25 K with $\ln B_{k0} = -0.20$: the evidence declines it, weakly, for the same reason it accepted the real one weakly.

22.4.2 The residual test against draws

The threshold of equation (22.12) and the minimum detectable bias of Table 22.7 are statements about replicate records, and they are tested on replicates. In the manner of Table 18.8, 4000 records are drawn from the model at the clean fit, the six readings from their error distributions and the three prior means from their spreads (NumPy's `default_rng(22)`); each is refitted by Gauss–Newton, its nine normalized residuals are formed with the hat matrix at its own fit, and the hall-2 reading's leave-one-out residual is formed from a second fit without that reading. The same draws are repeated with 2.2 and 4.5 K added to the hall-2 reading. Table 22.10 gives the fractions of records crossing each threshold, with the mean of the hall-2 statistic; the sampling error of a fraction near 0.03 is about 0.003, and of one near 0.4 about 0.008.

Table 22.10: The residual test on 4000 records drawn at the clean fit. Family: the largest of the nine $|r_k^{\mathrm{N}}|$. Hall 2: the hall-2 reading's normalized residual at the fit, and its leave-one-out residual from the fit without it. Fractions of records above $r^* = 2.77$ and 3; the mean of the hall-2 statistic.

Statistic and bias on the hall-2 reading	> 2.77	> 3	mean	Bonferroni bound
family, no bias	0.031	0.014	–	0.050, 0.024
hall-2 at the fit, no bias	0.006	0.003	−0.01	–
hall-2 at the fit, 2.2 K	0.047	0.028	−1.07	–
hall-2 at the fit, 4.5 K	0.19	0.14	−1.90	–
hall-2 leave-one-out, no bias	0.018	0.013	−0.15	–
hall-2 leave-one-out, 2.2 K	0.16	0.13	−1.42	–
hall-2 leave-one-out, 4.5 K	0.45	0.39	−2.75	–

On clean records the family false-alarm rate is 0.031 at $r^* = 2.77$ and 0.014 at 3, below the bounds of 0.050 and 0.024, as the correlation of the nine statistics and the critical pairs that cross together require. The bias is another matter. On the refit route of Section 22.3.3 a 4.5 K bias would be detected at $r^* = 3$ about half the time, $\Phi(4.5/1.479 - 3) = 0.52$; the leave-one-out residual of the draws detects it 0.39 of the time, and the normalized residual at the fit only 0.14. The two departures have two causes. The draws are made at the clean fit, whose hall-2 conductance is 4.11 kW/K and rise 5.45 K, while the record's refit sits at the handbook 5.00 kW/K and a rise of 4.38 K: the drawn statistic moves by $(2.75 - 0.15)/4.5 = 0.58$ per kelvin of bias, a predictive spread of 1.73 K, near the fit linearization's 1.79, and at that spread $\Phi(4.5/1.73 - 3) = 0.35$. The normalized residual at the fit loses more, because its fit follows the biased reading and its redundancy number falls as the bias grows (equation (22.21)): its mean moves by only 0.42 per kelvin. The minimum detectable biases of Table 22.7 are therefore optimistic for the hall sensors of this plant: they are reached only by the leave-one-out residual, and only if the path conducts as the handbook says; a hall-2 bias of 2.2 K is crossed by 3 % of records at the fit and 13 % by the leave-one-out residual.

22.4.3 The Laplace evidence against the exact integral

The evidence of equation (22.13) is a Laplace approximation, and it is checked against the exact integral. The exact evidence of each drawing is computed by integrating the entries that enter the rows linearly, Q_1, Q_2, T_s and the bias, in closed form for each pair (g_1, g_2), a Gaussian integral of the kind of equation (19.9), and then summing over a grid of 121×121 points spanning seven posterior standard deviations in each log-conductance; repeating with 201×201

points over nine standard deviations changes no value in the fourth decimal. Table 22.11 compares four routes for six drawings: the closed form (22.16) with $1 - K_{kk}$ and r_k^{N} from the fit of M_0; the Savage–Dickey ratio with the Laplace normal of b_k read at zero; the ratio of the two Laplace evidences (22.13), the route reported; and the exact integral.

Table 22.11: Checks of the log Bayes factor $\ln B_{k0}$, $\sigma_b = 2$: the linear closed form at the fit of M_0, the Savage–Dickey ratio with the Laplace normal of the bias, the ratio of Laplace evidences of equation (22.13), and the exact integral by quadrature over the log-conductances.

Drawing and record	Closed form	Savage–Dickey, Laplace	Laplace ratio	Exact
M_{T_s}, record (a)	0.372	0.371	0.227	0.233
M_{T_r}, record (a)	−0.209	−0.215	−0.126	−0.129
M_P, record (a)	0.029	0.026	0.038	0.033
M_{T_2}, record (b)	0.372	0.572	0.605	0.681
M_{Q_1}, record (b)	0.079	0.074	0.081	0.072
M_{T_2}, clean record	−0.260	−0.302	−0.204	−0.190

The ratio of Laplace evidences agrees with the exact value to within 0.015 nats in five of the six cases, and is 0.08 nats low for the hall-2 drawing of record (b), whose posterior along the unobservable direction is held by priors and is the least Gaussian. The two routes that use one linearization are worse, by up to 0.14 nats in record (a) and 0.31 in record (b). The reason is physical. A bias drawing moves the fitted state, the supply temperature from 12.70 to 12.23°C under M_{T_s}, the hall-2 conductance from 3.05 to 3.92 kW/K under M_{T_2}, and the curvature of the two hall reading rows, which scales with the square of the temperature rise, moves with it: under M_{T_s} the determinant ratio times σ_b^2 is 28.1 where the linear form, at the fit of M_0, has $1 + \psi = 21.0$. Each Laplace evidence evaluates its curvature at its own MAP state and follows the change; a single linearization cannot. The ranking is the same by every route: the closed form is the right tool for the argument of Section 22.3.2, and the ratio of the two fits' evidences for the numbers.

22.4.4 Sensitivity to the bias prior

The evidence depends on the prior of the bias, as every Bayes factor depends on the prior of the entry it frees. Repeating the ranking with $\sigma_b = 0.5$, 1, 2, 4 and 8 in the reading's units, the supply drawing of record (a) has $\ln B_{k0} = 0.61$, 0.63, 0.23, −0.38 and −1.05 and the return drawing 0.41, 0.33, −0.13, −0.75 and −1.42; the hall-2 drawing of record (b) has 0.07, 0.24, 0.61, 0.69 and 0.27, with the hall-1 meter's drawing next at 0.01, 0.02, 0.08, 0.21 and 0.27; and the largest log Bayes factor of the clean record is −0.001, −0.005, −0.019, −0.080 and −0.293. Each drawing's evidence is largest when the prior's spread matches the bias the record supports, about 1 K for the supply sensor and about 4 K for the hall-2 sensor, and falls on either side: a narrow prior cannot reach the bias, a wide one spreads its mass over biases the record excludes and pays the Occam factor $\frac{1}{2}\ln(1 + \psi)$ for it. The supply sensor leads record (a) for every prior up to 2 K; at 4 and above its evidence falls below zero, and the chiller-meter drawing, whose prior in kilowatts is then wide enough to repair the loop balance by itself, leads with 0.03 and −0.21, by a margin of a few hundredths. The hall-2 sensor leads record (b) at every prior, tied with the hall-1 meter at 8. No prior produces a false alarm on the clean record. The bias prior is part of the drawing and is fixed with the family, before the record is seen, from what is known of the sensors' calibration; the 2 of this chapter says that a bias, if present, is of the order of the calibration tolerance of a field sensor.

22.4.5 The exact balances as a limit

The answers of Section 22.3 were computed on the reduced state. Table 22.12 repeats the quantities of the two leading drawings on the full state, nine entries $(g_1, g_2, Q_1, Q_2, T_s, T_1, T_2, P, T_r)$ with the two hall balances, the header balance and the loop balance as physics rows at the tolerance σ_p, thirteen rows in all and still four degrees of freedom. Each column is one tolerance, in kelvin for the three temperature balances and in kilowatts for the header, with the reduced state in the last column; each row is a quantity of record (b) or of record (a).

Table 22.12: The fits of the two leading drawings on the full state with the four balances at the tolerance σ_p, against the reduced state. Record (b): the hall-2 reading and M_{T_2}; record (a): the supply reading and M_{T_s}; $\hat{b} \pm$ sd in K.

Quantity	$\sigma_p = 0.1$	0.01	0.001	reduced
(b) $2\mathcal{L}(\hat{\mathbf{z}}_0)$ [–]	3.3924	3.3986	3.3986	3.3986
(b) $1 - K_{22}$ [–]	0.0420	0.0420	0.0420	0.0420
(b) $r^{\mathrm{N}}_{T_2}$ [–]	−1.770	−1.770	−1.770	−1.770
(b) $\ln B_{T_2}$ [nat]	0.6041	0.6053	0.6053	0.6053
(b) $\hat{b}_{T_2} \pm$ sd	1.889 ± 1.373	1.889 ± 1.371	1.889 ± 1.371	1.889 ± 1.371
(a) $2\mathcal{L}(\hat{\mathbf{z}}_0)$ [–]	4.5724	4.7071	4.7085	4.7085
(a) $1 - K_{T_sT_s}$ [–]	0.4354	0.4503	0.4505	0.4505
(a) $r^{\mathrm{N}}_{T_s}$ [–]	−1.961	−1.995	−1.995	−1.995
(a) $\ln B_{T_s}$ [nat]	0.1738	0.2265	0.2270	0.2270
(a) $\hat{b}_{T_s} \pm$ sd	0.849 ± 0.444	0.850 ± 0.437	0.850 ± 0.437	0.850 ± 0.437

At $\sigma_p = 0.01$ every quantity agrees with the reduced state to within 0.0015, and at 0.001 to the last printed digit. At 0.1 the record (a) quantities still move, the log Bayes factor of the supply drawing by 0.05 nats: a loop balance allowed to fail by 0.1 K is not small against the 0.3 K of the supply and return sensors whose disagreement it carries, and the redundancy number of the loop balance itself is then 0.048, against zero in the limit. The hall-2 quantities of record (b) hardly move at any tolerance, because a hall balance allowed to fail by 0.1 K adds only $0.01\,\mathrm{K}^2$ to the $0.25\,\mathrm{K}^2$ of the hall-2 sensor it relays. The normalizers of the physics rows, which grow as σ_p shrinks, are the same in M_0 and M_k and cancel in every Bayes factor.

22.4.6 The residual test against the evidence ranking

The residual test and the evidence ranking agree on the leading candidate in the three records of the chapter: the supply sensor in record (a), the hall-2 sensor in record (b), and nothing in the clean record; with the second hall-2 sensor installed, the faulty one. For linear rows equation (22.16) makes the log Bayes factor an increasing function of the squared normalized residual at fixed ψ_k, so that the two rank alike among rows of similar ψ_k and can differ only where ψ_k differs, through the redundancy number or the ratio σ_b/σ_k. The 16 kW meter error of Exercise 22.5 is such a case: the residual test leads with the meter, the evidence with the return sensor, whose ψ is a hundred times the meter's.

They differ in confidence and in what they return. The residual test returns a suspect and a threshold crossed or not; in both records it crossed nothing, and its answer was that no reading can be named. The evidence returns, for every sensor, a bias with its band and a Bayes factor, and over the family a posterior probability: 0.85 ± 0.44 K on the supply sensor at 0.19, 1.89 ± 1.37 K on the hall-2 sensor at 0.30. What it adds to the residual is the odds: through ψ_k it weighs the normalized residual against how well the other rows predict the reading and against the prior size of a bias, the Occam factor; and for the nonlinear hall rows each drawing's evidence carries the curvature at its own fit, which Section 22.4.3 found worth up to 0.3 nats

here. It does not see a different disagreement from the residual test: the order of the supply and return sensors in record (a), 1.99 against 1.66, was already the residual test's. Its probabilities are also conditional on the family, which in record (b) holds no drawing of a changed hall-2 path (Section 22.6.5). The residual test is the right first screen, cheap and local, run after every fit; the evidence ranking is the diagnosis, run when the screen or the operator raises a doubt.

Both fail in the same way, on a reading of small redundancy number, whose normalized residual and ψ_k are both small: neither method can be repaired by a better statistic, only by a row that predicts the reading. The standard treatment of these tests, in the vocabulary used here, bad-data detection, identification, the largest normalized residual, the redundancy number and the critical measurement, is the state-estimation literature of power systems, which the companion volume *Reliability Engineering* treats in its chapter on state estimation.

22.5 The report (step 6)

The larger fault is the better hidden one. The hall-2 sensor of record (b) reads 2.2 K high, 4.4 standard deviations of the sensor, and the record passes the chi-square check with a tail of 0.49, more easily than the 1.0 K supply bias of record (a) at 0.32; no normalized residual reaches the threshold, and kept in the fit the reading reports hall 2 at 40 kW at 25.26 ± 1.44°C, 3.4 K above the 21.89°C of the clean record, with a band as narrow as an honest one. With the present sensors no test can see a hall-2 bias smaller than about 4.4 K half the time, or 5.7 K with probability 0.8.

22.5.1 The answers and their bands

Table 22.13 is the report for the six questions of Table 22.2: one row per question, in the table's order, with the answer from the posteriors of the state, the exact value beside it where one exists (the evidence by quadrature, the rates on replicate records), and the evidence behind it.

22.5.2 The report line

Fitted to the six readings of the two-hall loop with the return sensor, neither biased record fails the chi-square check and no reading crosses the residual threshold. Freeing a bias on each sensor in turn, the evidence points at the supply sensor in the first record, reading 0.85 ± 0.44 K high, and at the hall-2 sensor in the second, reading 1.89 ± 1.37 K high, in both cases weakly: Bayes factors of 1.3 and 1.8 (1.3 and 2.0 by the exact evidence). The hall-2 sensor cannot be checked by the present design at all: the rest of the plant predicts its reading only to ± 1.4 K, a bias under about 4.4 K is absorbed into the hall-2 conductance, and kept in the fit it would report hall 2 at 40 kW 3.4 K too warm with an ordinary band. Until the hall-2 sensor is checked, the hall-2 answers are reported with the bias freed, 22.36 ± 2.39°C at 40 kW.

22.6 The next measurement (step 7)

The report leaves two questions open: which of the loop's sensors carries the bias of record (a), and whether the hall-2 sensor of record (b) is biased at all. The candidate measurements are more days of the present readings, a second reading of the supply temperature, and the candidate rows of Section 22.3.5; each is valued by the spread it leaves on the bias in question and by what it does to the ranking.

22.6.1 The record length for the supply bias

A bias persists from one reading to the next while the sensor's noise does not. Averaging n daily readings of a sensor whose errors are independent from day to day keeps the bias and

Table 22.13: The answers to the questions of this chapter: answers from the posteriors of the state, with the exact value beside each where one exists.

Question	Answer (state route)	Exact	Evidence
Do the readings disagree with the model?	no, in all three records	–	$2\mathcal{L} = 0.79$, 4.71, 3.40 on 4 degrees of freedom, tails 0.94, 0.32, 0.49
Is this reading lying? Which one?	no reading can be named by the residual test	family false-alarm rate 0.031 on draws (bound 0.050)	largest $\lvert r_k^{\mathrm{N}} \rvert$ 1.99 (supply, record (a)) and 1.77 (hall 2, a critical pair with its conductance prior, record (b)), below $r^* = 2.77$
Which sensor is biased, by how much, how sure?	record (a): the supply sensor, 0.85 ± 0.44 K; record (b): the hall-2 sensor, 1.89 ± 1.37 K; both weakly	$\ln B_{k0} = 0.233$ and 0.681	$\ln B_{k0} = +0.23$ and $+0.61$; family probabilities 0.19 and 0.30; the return sensor second in (a) at -0.13; in (b) the advantage rests on a family without a conductance-change drawing and on $\sigma_b = 2$ K against $\sigma_g = 0.3$
Can a hall-2 bias be detected at all?	not below about 4.4 K (50 %) or 5.7 K (80 %) with the present design	a 4.5 K bias crossed by 39 % of drawn records (52 % on the reported route)	the other rows predict the reading to ± 1.39 K (refit); $1 - K_{22} = 0.078$ at the clean fit and 0.042 at the biased fit (fit linearization 5.4 and 7.3 K); the bias and g_2 share an unobservable direction
What does the kept reading cost?	T_2^* at 40 kW reported 3.4 K high with a normal-width band	–	25.26 ± 1.44 against 21.89 ± 1.21°C; with the bias freed 22.36 ± 2.39; T_1^* unaffected
Which row makes the hall-2 sensor checkable?	a second hall-2 temperature sensor; a path test second; a load or coil heat meter does not	–	$1 - K_{22} = 0.52$, 0.40 and 0.08; under record (b) the second sensor gives $r^{\mathrm{N}} = -3.28$ and $\ln B_{k0} = 3.88$

divides the variance of the noise by n, $\mathrm{Var}(\bar{y}_k) = \sigma_k^2/n$, so that the reading rows of the averaged record have the standard deviations $\sigma_k/\sqrt{n}$ while the prior rows keep theirs. Split the Jacobian of a drawing into its reading rows $\mathbf{J}_\mathrm{y}$ and its prior rows $\mathbf{J}_\pi$; the averaged record multiplies the reading rows' information by n, and the preposterior covariance of the drawing is

$$\boldsymbol{\Sigma}_a(n) = \left(n\,\mathbf{J}_\mathrm{y}^\mathsf{T}\mathbf{J}_\mathrm{y} + \mathbf{J}_\pi^\mathsf{T}\mathbf{J}_\pi\right)^{-1}, \tag{22.24}$$

the case of equation (2.17) in which the candidate is $n-1$ more copies of the reading rows, outcome-free to the order of the linearization. With the averages taken equal to the readings of record (a), as Chapter 23 does for its ten-day record, every reading row's contribution to the objective is multiplied by n. The convention has a cost that Chapter 23 also states: holding the averages at the one-day values scales record (a)'s realized noise up by $\sqrt{n}$ along with the bias, where a real average would shrink that noise. The expected record, readings without noise generated from the clean fit with the supply sensor offset by 1.0 K, is the check: at three days it gives a misfit tail of 0.005 and normalized residuals of -3.79 on the supply sensor and $+3.57$ on the return sensor, a little stronger than the table's. Table 22.14 gives the result for $n = 1$ to 10 days. Each row is one record length; the columns give the misfit and its tail on four degrees of freedom, the magnitudes of the supply and return normalized residuals, the log Bayes factors of the supply, return and chiller-meter drawings, the posterior probabilities of the supply and return drawings over the family, and the fitted supply bias.

Table 22.14: Record (a) as an average of n daily readings, reading standard deviations $\sigma_k/\sqrt{n}$, priors unchanged, $\sigma_b = 2$. Pr over the seven drawings at equal prior odds.

n	$2\mathcal{L}$	tail	$\lvert r^\mathrm{N}_{T_s}\rvert$	$\lvert r^\mathrm{N}_{T_r}\rvert$	$\ln B_{T_s}$	$\ln B_{T_r}$	$\ln B_P$	$\Pr(M_{T_s})$	$\Pr(M_{T_r})$	$\hat{b}_{T_s}$ [K]
1	4.71	0.32	1.99	1.66	0.23	−0.13	0.04	0.19	0.13	0.85 ± 0.44
2	8.11	0.088	2.70	2.45	1.56	1.19	0.46	0.36	0.25	0.85 ± 0.32
3	11.44	0.022	3.26	3.05	3.01	2.64	1.14	0.50	0.34	0.85 ± 0.26
5	18.05	0.0012	4.15	3.98	6.05	5.68	3.00	0.57	0.40	0.85 ± 0.21
10	34.50	6×10^{-7}	5.80	5.68	13.92	13.55	9.16	0.59	0.41	0.85 ± 0.15

Three days of the same averages are enough for detection: the misfit's tail falls to 0.022, the supply sensor's normalized residual crosses 3 and the return sensor's does too, at 3.05, both above the Bonferroni threshold 2.77, and the supply drawing reaches $\ln B_{k0} = 3.0$. Ten days make the detection certain, a tail of 6×10^{-7} and a log Bayes factor of 13.9, and determine the bias to 0.85 ± 0.15 K, the spread (22.24) gives. They do not settle which sensor. The return drawing rises in step, to 13.55, and the two drawings share the family's probability in the ratio 0.59 to 0.41 at ten days, 0.57 to 0.40 at five.

The reason is the limit the averaging approaches. As n grows the reading rows outweigh the priors, and the design tends to the readings-only design of Table 22.7, in which the supply sensor, the return sensor and the chiller meter share a single degree of freedom, $0.41 + 0.41 + 0.18 = 1$. One degree of freedom among three rows makes their columns of the residual sensitivity matrix proportional, so every two of them form a critical pair by equation (22.10), and their normalized residuals converge: at $n = 100$ they are -18.14, $+18.11$ and -18.03, at $n = 10^4$ -181.22, $+181.22$ and -181.21. Equation (22.16) then gives the supply and return drawings the same gain in fit, and the gap between their log evidences tends to what each pays in the set-point and conductance priors, which an average of readings does not tighten: 0.367 nats at three days, 0.370 from a hundred days on. For the expected record the gap settles at 0.47 nats. Ten days settle that one of the loop's sensors is biased by about 0.85 K, not which.

The row that separates them is a second, independent reading of the supply temperature, such as the chiller's own leaving-water sensor. With one at 12.10°C and $\sigma = 0.3$ K, a single day of record (a) gives $\ln B_{k0} = +1.34$ for the supply drawing and -1.23 for the return drawing,

a gap of 2.6 nats where ten days of the present sensors left 0.4; its preposterior spread of the supply bias is 0.37 K against the present 0.44.

22.6.2 The candidates for the hall-2 bias

For record (b) the candidates are the rows of Table 22.9, valued now by what they leave of the band on the hall-2 bias. Under M_{T_2} the preposterior covariance of equation (2.17) at the present MAP state, with the candidate's Jacobian row of Section 22.3.5, gives the spread of the bias and of the conductance after the row is installed, before it gives a reading. Table 22.15 lists them with the record (a) candidates of the previous subsection. Each row is a candidate; the columns give the posterior standard deviation of the bias in question before and after it, the width removed, the posterior standard deviation of g_2 after it for the hall-2 candidates, and the log Bayes factor of the leading drawing once the candidate's illustrative reading is in the record.

Table 22.15: The next measurement. Preposterior standard deviations of the bias, in K, under M_{T_s} for record (a) and M_{T_2} for record (b), from equation (2.17) (and (22.24) for the averaged days); $\ln B$ of the leading drawing with the illustrative readings of Tables 22.14 and 22.9.

Candidate	sd now	sd after	removed	sd(g_2) after	$\ln B$
Record (a): the supply bias					
three days of the present readings	0.437	0.264	0.173	–	3.01
ten days of the present readings	0.437	0.147	0.290	–	13.92
second supply sensor, 0.3 K, one day	0.437	0.367	0.070	–	1.34
Record (b): the hall-2 bias					
hall-2 load meter, 1.5 kW	1.371	1.350	0.021	0.232	0.57
hall-2 coil heat meter, 1.0 kW	1.371	1.346	0.025	0.231	0.56
second hall-2 sensor, 0.5 K	1.371	0.656	0.715	0.120	3.88
path test, $\sigma_g = 0.10$	1.371	0.860	0.511	0.097	1.64
path test, $\sigma_g = 0.05$	1.371	0.752	0.619	0.050	2.84

For the supply bias, more days narrow the band most and settle nothing about which sensor; the second supply sensor narrows it least and settles the question, because its row has its entry in the column of T_s alone and so breaks the critical pair of the loop's sensors. For the hall-2 bias, the meters on Q_2 remove less than a fiftieth of the band, because they add nothing along the unobservable direction (22.20). The second sensor removes half of it, more than either path test, although the path test of 0.05 leaves g_2 better known: the second sensor reads the hall-2 temperature, the short axis of Figure 22.4, directly, and the bias is the difference between two readings of the same quantity, while the path test fixes the long axis only through the prior it replaces.

22.6.3 The recommendation

The next measurement is a second temperature sensor in hall 2, on the same balance as the first. It raises the hall-2 reading's redundancy number from 0.078 to 0.52 and its minimum detectable bias from 4.4 to 2.1 K, narrows the band on a bias from 1.37 to 0.66 K, and on record (b) it would have flagged the faulty sensor by its normalized residual and named it by the evidence at odds of 4.2 to one. Where a second sensor cannot be mounted, a commissioning test of hall 2's path to 0.05 in the logarithm is the second choice. A meter on the hall-2 load or on its coil is not a remedy for this sensor. For the loop, the chiller's leaving-water sensor brought into the record as a second supply reading separates a supply bias from a return bias in a single day, which no length of the present record does.

22.6.4 What the analysis showed

The first reading of the hall-2 record, a passed chi-square check and no normalized residual near the threshold, said the sensor was right and the hall had warmed; the analysis shows that this record could not have said anything else. The hall-2 sensor is the only reading of its hall and is checked only by a handbook conductance, so a bias of a few kelvin on it slides into that conductance, leaves the misfit unremarkable and moves every hall-2 answer at higher load by more than its band. The supply sensor of record (a) is the opposite case: the return sensor and the chiller meter check it through the loop balance, its bias is found and sized, if weakly, from one day, and the question left open there is which of two tied sensors carries it, a question more days cannot answer and a second supply reading can. What an engineer takes back to the site is that whether a sensor can be checked is read from the drawing before any reading is taken, by its redundancy number; that the hall sensors of this loop cannot be, as installed; and that the cure is a second path to the same temperature, not a better test or a meter on another quantity. Until that sensor is in place, a hall-2 answer is reported with the bias freed, wide but honest.

22.6.5 Limits of the model

The method finds what its family contains. The family of this chapter holds one constant bias on one reading at a time. Two biased readings at once, such as a shared calibration error of the two hall sensors, are a drawing it does not contain; the largest single drawing then absorbs part of the pair and the rest stays in the residuals, and the family must be enlarged by pairs, at the cost of a larger Occam factor for each. A bias that drifts over the record is not a constant, and an average of readings mixes its values. A gross error, a single reading far off, is a bias on one record of one sensor, and the same drawing serves it; but its size may lie far outside a prior set for calibration drift, and a prior of 2 in the reading's units cannot reach an error of 8 (Exercise 22.5). The residual tests are exact for linear rows; the two hall reading rows are nonlinear in the reduced state, so for them the tests are linearized at the fit, and the redundancy number moves with the fit, as the hall-2 reading's moved from 0.078 to 0.042.

Most important, a bias drawing explains disagreement as a fault of a sensor, and a disagreement can equally be a change of the plant. In record (b) the drawing with a biased hall-2 sensor and the drawing with a hall-2 path that conducts less than the handbook says are the same direction of Figure 22.4: the reading and the conductance prior are a critical pair, and the evidence prefers the sensor only because the bias prior is wider than what the conductance prior allows. Telling a sensor fault from a fouled path needs the drawings of the plant's own faults in the same family.

Common mistake

Treating a passed chi-square check as a clean bill for every reading. Record (b) passes detection with a tail of 0.49, better than record (a), while its kept reading moves the hall-2 conductance by 26 % and the hall-2 temperature at 40 kW by 3.4 K, outside its own band. The check sees a bias only through the rows that predict the biased reading; a reading the rest of the plant predicts poorly can be wrong by several kelvin and leave the misfit unremarkable.

Common mistake

Removing the row with the largest normalized residual and calling the record clean. In record (b) the hall-2 reading and the conductance prior are a critical pair; removing the reading returns the conductance to the handbook value and the hall-2 temperature at 40 kW to 20.14 ± 2.41°C, on the other side of the clean answer from the biased fit, and removing the prior instead would have kept the bias. Within a critical pair the residual test has no basis to choose; the identification rule is applied only to a row that stands alone above the threshold, and the choice within a pair is left to evidence or to a new measurement.

Common mistake

Ranking sensors by their normalized residuals without their redundancy numbers. A reading with a normalized residual of 1.8 and a redundancy number of 0.04 carries an expected shift of only 0.9 for a bias of 2.2 K; a small normalized residual on such a reading is not evidence that it is right. The redundancy number and the minimum detectable bias go into the report with every residual.

The next chapter widens the family from sensor biases to the faults of the plant itself, a fouled path, a derated chiller, and ranks them together by the same evidence.

22.7 What to retain

Definition 22.1 (Residual tests and bias drawings)**.** On the reduced state of a plant with Gaussian rows, at the MAP state of linear rows the residual vector is $-(\mathbf{I}-\mathbf{K})\boldsymbol{\zeta}$, with covariance $\mathbf{I}-\mathbf{K}$ and $\mathbf{K}=\mathbf{J}\boldsymbol{\Sigma}\mathbf{J}^{\mathsf{T}}$; the normalized residual $r_k^{\mathrm{N}} = r_k/\sqrt{1-K_{kk}}$ is standard normal and equals the standardized leave-one-out residual $(\hat{h}_{-k}-y_k)/\sqrt{\sigma_k^2+V_{-k}}$, with V_{-k} the variance of the reading's prediction from the other rows and $1-K_{kk}=\sigma_k^2/(\sigma_k^2+V_{-k})$. A bias Δ on reading k shifts r_k^{N} by $\Delta/\sqrt{\sigma_k^2+V_{-k}}$, and the minimum detectable bias at threshold r^* is $r^*\sqrt{\sigma_k^2+V_{-k}}$, with V_{-k} from the refit without row k; the Bonferroni threshold over n_r rows is $\Phi^{-1}(1-p_{\mathrm{fa}}/2n_r)$. The bias drawing M_k adds an entry b_k with prior $\mathcal{N}(0,\sigma_b^2)$ to row k; its Laplace log Bayes factor against M_0 is $\mathcal{L}_0-\mathcal{L}_k-\frac{1}{2}\ln(\det\boldsymbol{\Lambda}_k/\det\boldsymbol{\Lambda}_0)-\ln\sigma_b$, and for linear rows $\frac{1}{2}[\psi_k(r_k^{\mathrm{N}})^2/(1+\psi_k)-\ln(1+\psi_k)]$ with $\psi_k=\sigma_b^2(1-K_{kk})/\sigma_k^2$. A candidate row's effect on a reading's redundancy number follows from the preposterior covariance before the row gives a reading.

Three conclusions carry beyond the two halls. The residual test returns a suspect; the evidence returns the bias with its band and a probability over the family it was given. A reading is checkable only to the extent that the rest of the plant predicts it: a bias on a reading of small redundancy number is absorbed by the entry only that reading informs, and the answers that depend on that entry are reported wrong with an ordinary band. And checkability is a design property, computed from the Jacobian before any reading is taken: the redundancy number and the minimum detectable bias of each sensor rank the rows that would make it checkable, and only a second path to the quantity the sensor measures raises them.

Table 22.16 lists the terms of the chapter with their expressions, units and the subsection that defines each.

Table 22.16: The terms of bad-data detection and sensor-bias diagnosis and where each is defined.

Term	Expression	Units	Meaning	Defined in
Chi-square tail ($\nu = 4$)	$(1 + x/2)\,\mathrm{e}^{-x/2}$	–	detection check	22.3.1
Bonferroni threshold	$r^* = \Phi^{-1}(1 - p_{\mathrm{fa}}/2n_r)$	–	family false-alarm control	22.3.1
Bias drawing	$(h_k + b_k - y_k)/\sigma_k$, $b_k \sim \mathcal{N}(0, \sigma_b^2)$, eq. (22.4)	reading	one sensor biased	22.1.6
Laplace log Bayes factor	$\mathcal{L}_0 - \mathcal{L}_k - \frac{1}{2}\ln\frac{\det \mathbf{\Lambda}_k}{\det \mathbf{\Lambda}_0} - \ln\sigma_b$	nat	evidence for the bias	22.3.2
Family probability	$B_{k0}/\sum_j B_{j0}$, eq. (22.14)	–	ranking at equal odds	22.3.2
Response to a bias	$\mathbb{E}[r_k^{\mathrm{N}}] = -\Delta/\sqrt{\sigma_k^2 + V_{-k}}$	–	shift of the statistic	22.3.3
Minimum detectable bias	$b_k^* = r^*\sqrt{\sigma_k^2 + V_{-k}}$	reading	smallest visible bias	22.3.3
Unobservable bias direction	$(\Delta g_2, \Delta b_{T_2}) = (1, T_2 - T_s)$	–	bias traded for conductance	22.3.3
Redundancy after a row	$\sigma_k^2/(\sigma_k^2 + V_{-k,a})$, eq. (22.23)	–	checkability before installation	22.3.5
Residual sensitivity matrix	$\mathbf{I} - \mathbf{K} = \mathrm{Cov}[\mathbf{r}(\hat{\mathbf{z}})]$	–	response of the residuals to row errors	22.2.1
Normalized residual	$r_k/\sqrt{1 - K_{kk}} \sim \mathcal{N}(0,1)$, eq. (22.5)	–	bad-data statistic	22.2.1
Redundancy number	$1 - K_{kk} = \sigma_k^2/(\sigma_k^2 + V_{-k})$	–	how well the plant predicts a reading	22.2.2
Standardized leave-one-out residual	$(\hat{h}_{-k} - y_k)/\sqrt{\sigma_k^2 + V_{-k}} = r_k^{\mathrm{N}}$, eq. (22.8)	–	prediction against the reading	22.2.2
Deletion identity	$2\mathcal{L} - 2\mathcal{L}_{-k} = (r_k^{\mathrm{N}})^2$	–	misfit a row explains	22.2.2
Critical measurement	$1 - K_{kk} = 0$, $r_k \equiv 0$	–	reading nothing checks	22.2.2
Critical pair	$(\mathbf{I} - \mathbf{K})\mathbf{e}_k \propto (\mathbf{I} - \mathbf{K})\mathbf{e}_l$	–	rows no test separates	22.2.2
Linear closed form	$\frac{1}{2}[\psi_k (r_k^{\mathrm{N}})^2/(1 + \psi_k) - \ln(1 + \psi_k)]$	nat	evidence from the residual	22.3.2
Ratio ψ_k	$\sigma_b^2(1 - K_{kk})/\sigma_k^2 = \sigma_b^2/(\sigma_k^2 + V_{-k})$	–	bias prior against the check	22.3.2
Bias estimate	$\hat{b}_k = \sigma_b^2(y_k - \hat{h}_k)/(\sigma_k^2 + \sigma_b^2(1 - K_{kk}))$, eq. (22.17)	reading	size of the bias	22.3.2
Averaged record	$\mathbf{\Sigma}_a(n) = (n\mathbf{J}_{\mathrm{y}}^{\mathsf{T}}\mathbf{J}_{\mathrm{y}} + \mathbf{J}_{\pi}^{\mathsf{T}}\mathbf{J}_{\pi})^{-1}$, bias kept	reading	record length	22.6.1

22.8 Exercises

Exercise 22.1. *Formulation.* In record (a) the fit without bias returns the standardized residuals $r_3 = -1.339$ (supply sensor), $r_6 = +1.118$ (return sensor) and $r_4 = -0.565$ (chiller meter), with redundancy numbers 0.4505, 0.4523 and 0.2248, and $2\mathcal{L}(\hat{\mathbf{z}}) = 4.709$. Compute the three normalized residuals; the threshold of the largest-normalized-residual test for nine rows at a family false-alarm probability of 5 % by the Bonferroni bound; the verdict; and the misfit left after deleting each of the three rows. ($r^{\mathrm{N}} = -1.339/\sqrt{0.4505} = -1.995$, $1.118/\sqrt{0.4523} = 1.662$, $-0.565/\sqrt{0.2248} = -1.192$. By equation (22.12), $r^* = \Phi^{-1}(1 - 0.05/18) = \Phi^{-1}(0.99722) = 2.77$;

no row crosses it, and no reading is identified. By equation (22.9) the misfits after deletion are $4.709 - 3.980 = 0.729$, $4.709 - 2.763 = 1.946$ and $4.709 - 1.421 = 3.288$: deleting the supply sensor explains the most, the return sensor nearly as much.)

Exercise 22.2. *Formulation.* Under record (b) the hat matrix has $K_{22} = 0.958$ for the hall-2 reading, $\sigma_T = 0.5\,\mathrm{K}$. State the redundancy number; the standard deviation with which the other eight rows predict the hall-2 temperature; the minimum detectable bias at $r^* = 3$ at this fit linearization; and the mean normalized residual that the record's bias of 2.2 K produces. ($1 - K_{22} = 0.042$. By equation (22.7), $V_{-k} = \sigma_T^2 K_{22}/(1 - K_{22}) = 0.25 \times 0.958/0.042 = 5.70\,\mathrm{K}^2$, a standard deviation of 2.39 K. By equation (22.19), $b^* = 3\sqrt{0.25 + 5.70} = 7.32\,\mathrm{K}$. By equation (22.18), $2.2\sqrt{0.042}/0.5 = 0.90$ in magnitude: the expected statistic is a third of the threshold.)

Exercise 22.3. *Derivation.* Using $\mathbf{r}(\hat{\mathbf{z}}) = -(\mathbf{I} - \mathbf{K})\tilde{\mathbf{y}}$ and $\mathrm{Cov}[\mathbf{r}(\hat{\mathbf{z}})] = \mathbf{I} - \mathbf{K}$ of Section 22.2.1: (a) show that a critical measurement has a residual that is zero for every record, not only on average; (b) derive the deletion identity (22.9); (c) in record (a) the residual sensitivity matrix has $(\mathbf{I}-\mathbf{K})_{36} = -0.376$ between the supply and return rows, with diagonal entries 0.4505 and 0.4523; state the correlation of their normalized residuals and what it says about identifying either. ((a) $1 - K_{kk} = \mathbf{e}_k^{\mathsf{T}}(\mathbf{I} - \mathbf{K})\mathbf{e}_k = \|(\mathbf{I} - \mathbf{K})\mathbf{e}_k\|^2$ by idempotence, so a zero redundancy number gives $(\mathbf{I}-\mathbf{K})\mathbf{e}_k = \mathbf{0}$ and $r_k = -\mathbf{e}_k^{\mathsf{T}}(\mathbf{I}-\mathbf{K})\tilde{\mathbf{y}} = 0$ for every data vector: $\mathbf{e}_k$ lies in the column space of $\mathbf{J}$, and any error on row k is a change of state. (b) Deleting row k is the same as adding a free entry x on it with no prior: the misfit is $\min_x \|(\mathbf{I}-\mathbf{K})(\mathbf{e}_k x - \tilde{\mathbf{y}})\|^2 = 2\mathcal{L} - [\mathbf{e}_k^{\mathsf{T}}(\mathbf{I}-\mathbf{K})\tilde{\mathbf{y}}]^2/[\mathbf{e}_k^{\mathsf{T}}(\mathbf{I}-\mathbf{K})\mathbf{e}_k] = 2\mathcal{L} - r_k^2/(1 - K_{kk})$. (c) By equation (22.5), $-0.376/\sqrt{0.4505 \times 0.4523} = -0.83$: a bias on either sensor moves both normalized residuals, in opposite directions and nearly in proportion, so that the residual test separates them only weakly; at the limit of Section 22.6.1 the correlation reaches -1 and they form a critical pair.)

Exercise 22.4. *Derivation.* Derive the evidence of a bias drawing as a rank-one update of the base evidence for linear rows, equation (22.16), with the bias estimate and its spread, and evaluate it for the supply sensor of record (a) with the numbers of Exercise 22.1. Compare with Table 22.11. (With $b' = b/\sigma_b$ and $\mathbf{a} = (\sigma_b/\sigma_k)\mathbf{e}_k$, profiling the state gives $2\mathcal{L}_k(b') = 2\mathcal{L}_0 + 2(\sigma_b/\sigma_k)r_{0,k}b' + (1 + \psi_k)b'^2$, minimized at $\hat{b}' = -(\sigma_b/\sigma_k)r_{0,k}/(1 + \psi_k)$ with a drop of $\psi_k(r_k^{\mathrm{N}})^2/(1 + \psi_k)$; the Schur complement of the bias in the curvature is $1 + \psi_k$, so $\det \mathbf{\Lambda}'_k = (1 + \psi_k)\det \mathbf{\Lambda}_0$, and with no constant left in the normalized bookkeeping $\ln B_{k0} = \frac{1}{2}[\psi_k(r^{\mathrm{N}})^2/(1 + \psi_k) - \ln(1 + \psi_k)]$. Here $\psi = (2/0.3)^2 \times 0.4505 = 20.02$, so $\ln B = \frac{1}{2}[0.9524 \times 3.980 - \ln 21.02] = \frac{1}{2}(3.791 - 3.046) = 0.372$; by equation (22.17), $\hat{b} = 4 \times 0.3 \times 1.339/(0.09 + 4 \times 0.4505) = 0.849\,\mathrm{K}$ and $\mathrm{sd}(b \mid \mathbf{y}) = 2/\sqrt{21.02} = 0.436\,\mathrm{K}$, as the full fit. The ratio of Laplace evidences gives 0.227 and the exact integral 0.233: the linear form overstates by 0.14 nats because the curvature of the hall rows changes when the fitted supply temperature moves from 12.70 to 12.23°C.)

Exercise 22.5. *Application.* The chiller heat meter develops a gross error and reads 60.5 kW, 8 kW high, the other readings those of the clean record. Run both methods, and say why they return what they do. (Without bias the fit gives $2\mathcal{L} = 2.00$ on four degrees of freedom, tail $(1 + 1.00)\mathrm{e}^{-1.00} = 0.74$: no detection. It has moved the hall-2 load from 22.39 to 28.41 kW, the hall-2 conductance from 4.11 to 4.86 kW/K and the supply temperature from 12.15 to 11.84°C. The normalized residuals are +1.36 on the return sensor, −1.30 on the supply sensor and −1.27 on the meter, all below 2.77, and the meter is not even first. The meter's redundancy number is 0.21 and its minimum detectable bias $3 \times 2/\sqrt{0.21} = 13\,\mathrm{kW}$, above the error: the hall-2 load, which only the meter and the return sensor inform, absorbs most of it. The evidence puts the meter's drawing first but at $\ln B = +0.06$, with $\hat{b} = 0.96 \pm 1.82\,\mathrm{kW}$: the prior of 2 kW holds the bias near zero; with $\sigma_b = 10\,\mathrm{kW}$ the bias is $4.6 \pm 4.0\,\mathrm{kW}$ and $\ln B = -0.20$, the Occam factor now outweighing the fit. Neither method finds an error below the meter's minimum detectable bias; one of 16 kW gives $2\mathcal{L} = 9.93$, tail 0.042, and normalized residuals of −3.11 on

the meter, +3.00 on the return sensor and −2.86 on the supply sensor, so that the residual test flags the meter. The evidence ranks the return sensor first, $\ln B = 2.91$ and a family probability of 0.56, the supply sensor second at 2.19, and the meter third at 0.75, probability 0.06: by equation (22.16) the meter's drawing is capped by its $\psi = (2/2)^2 \times 0.20 = 0.20$, against 19.5 for the return sensor, so the same size of normalized residual buys the temperature sensors far more evidence. Here the two methods part, and it is the residual test that names the faulty reading: a bias prior of 2 kW, set for calibration drift, cannot reach a gross error of 16 kW.)

Exercise 22.6. *Application.* The budget allows one new instrument to make the hall-2 temperature sensor checkable: a second hall-2 temperature sensor of 0.5 K, or a flow meter on hall 2's coil which, with the water's temperature rise, measures the coil's heat to 1.0 kW. Choose by the redundancy number each gives the hall-2 reading, and explain the difference. (The coil meter measures $Q_2 = \dot{m}_2 c_p \Delta T_2$, a row on Q_2: the hall-2 reading's redundancy number goes from 0.078 to 0.082 and its minimum detectable bias stays at 4.4 K on the refit route (5.4 to 5.2 K at the fit linearization). The second sensor is a row on the same balance as the first: 0.52 and 2.1 K. Choose the sensor. The prediction of the hall-2 reading from the other rows has a variance of $2.97\,\mathrm{K}^2$, of which 2.67 is the conductance prior mapped through the rise of 5.45 K; the coil meter removes part of the remaining $0.3\,\mathrm{K}^2$ from the load, which the chiller meter and the return sensor already hold to 2.2 kW, while the bias still trades against g_2 along the direction (22.20). Only a second path to the hall-2 temperature, or to g_2 itself, removes that variance.)

Chapter 23

Fault Detection and Isolation

23.1 A plant that has changed

The two data halls of Chapter 17 were fitted, checked and propagated in Chapters 18 to 20 on one snapshot of their five readings, here called period 1. Some weeks later the maintenance engineer has a second record, period 2: ten days of readings in one steady operating condition, averaged day by day. Hall 1 now reads 19.80°C where it read 17.65°C; its load meter reads 29.5 kW, a little less than the 31.0 kW of period 1; the supply sensor, the hall-2 sensor and the chiller's heat meter are close to their old values. The first reading of the record is simple. In period 1 the load meter's 31.0 kW held hall 1 at $17.655 - 12.10 = 5.555$ K above the supply sensor, a conductance of $31.0/5.555 = 5.58$ kW/K; the same arithmetic on period 2 gives a rise of $19.80 - 12.20 = 7.60$ K and $29.5/7.60 = 3.88$ kW/K, a loss of 30 %: the coil has fouled, and the work order is to clean it.

The engineer who must sign the work order asks what that arithmetic rests on. It reads every sensor as exact, so it cannot say whether a rise of two kelvin in a record of noisy readings is a change at all. It takes the hall-1 sensor's word for the hall temperature, so a sensor that had drifted two kelvin high would give the same 3.88, and so would a supply warmer than its sensor reports or a load larger than its meter shows. And it would have been just as sure after one day. The question of this chapter is the one those doubts add up to,

how a plant that has changed is told from one that has not; how the component at fault is named among candidates generated from the plant graph, by the evidence of each; how large the fault is and how well that is known; how two fits of two periods are compared entry by entry to say which quantity moved; and what the length of the record buys.

What changes from Chapter 22, which freed a bias on one sensor of one record, is the evidence and the family: a second record is set against the posterior of the first, and the candidates are generated from the whole graph, its paths and its sources as well as its sensors.

Jargon: fault detection, fault isolation, fault identification, fault detection and isolation (FDI), model-based diagnosis, healthy model, candidate drawing, candidate family, Occam factor, posterior model probability, parameter-change detection, drift attribution, squared Mahalanobis distance, leave-one-out predictive check, Kullback–Leibler divergence

Fault detection decides whether the observed behavior is inconsistent with normal operation; *fault isolation* decides which component or fault is responsible; *fault identification* estimates the size of the fault. Detection and isolation together are *fault detection and isolation* (FDI), the standard name of the discipline; isolation by comparing models of the plant is *model-based diagnosis*. The *healthy model* is the model of the plant in normal operation against which a new record is tested; here it is the posterior of an earlier period entered as a prior. A *candidate drawing* (a fault hypothesis) is the healthy model with one quantity freed to represent one fault, and the *candidate family* is the finite list of such drawings, written down before the new record is examined. The *Occam factor* is the part of a drawing's evidence that penalizes the prior volume a freed quantity spreads over values the data do not use. The *posterior model probability* of a drawing is its probability given the record, over the family. *Parameter-change detection between two periods*, here called *drift attribution*, compares two fits of two periods entry by entry to say which entry of the state moved. The *squared Mahalanobis distance* of a vector from a point is the quadratic form of their difference in the inverse of a covariance matrix, a sum of squared standardized differences that accounts for correlation. The *leave-one-out predictive check* compares a reading with its prediction from a fit that did not use it. The *Kullback–Leibler divergence* of one density from another is the expected logarithm of their ratio under the first.

Key idea

A fault is diagnosed by asking which small change of the healthy model, the plant as an earlier period left it, explains the new record best. The changes are generated from the plant graph, one freed quantity each, and scored by their evidence, which rewards the fit and charges an Occam factor for the freedom taken; the ranking is a posterior over that family and holds only within it. Two changes that reach the readings through the same row and no other predict the same record, and no length of record separates them; only a new row does. Comparing the fits of two periods entry by entry locates the entry that moved without any list of causes, and it is blind in the same place.

23.1.1 From a reading that lies to a plant that has changed

Chapter 22 worked on one record. Its question was whether any one reading disagreed with the others and with the model, and its answer freed a bias on one sensor at a time and compared the drawings by their evidence. The plant itself was taken as described by the handbook priors. A plant in service is seen many times, and the question that maintenance asks of a new record is a different one: whether the plant still behaves as it did. That question needs two things the bad-data chapter did not. The first is a statement of how the plant behaved, a healthy model built from an earlier record rather than from the handbook; the second is a list of the ways it could have changed, which includes the sensors but also the physical paths and the sources. A fouled heat-removal path, a sensor that has drifted and a meter that reads low each change the readings in their own pattern, and the patterns, read through the plant's physical rows, are what tells them apart.

Chapter 2 put this question in the diagnosis category: hypothesis comparison run over a family of candidate faults generated from the plant graph, with a ranking as its answer (Section 2.5.3). Here that category is carried through on a fitted plant. The candidates are

drawings of the healthy model with one quantity freed, the evidence of each is computed from its fit, and the posterior probabilities over the family are the ranking.

23.1.2 Detection, isolation, identification and attribution

Detection, isolation and identification, the three tasks of the condition-monitoring and fault-diagnosis chapter of the companion volume *Reliability Engineering*, are three readings of one family of drawings: the chi-square tail of Chapter 18 on the healthy model's misfit, the posterior probabilities of the drawings, and the posterior of the freed quantity under the leading one. Drift attribution, the fourth task, reads two fits made with the handbook priors instead, and locates a change without a list of candidates and without naming a cause.

23.1.3 The plant and the two records (step 0)

The plant is the graph of Figure 17.1: two halls with heat loads Q_1, Q_2 in kW, heat-removal paths of conductance $G_i = G_{\mathrm{ref}}\mathrm{e}^{g_i}$ with $G_{\mathrm{ref}} = 1\,\mathrm{kW/K}$ pinned, one supply header at temperature T_s, and the chiller. The inference runs on the reduced state of Chapter 17, equation (17.7),

$$\mathbf{z} = (g_1,\ g_2,\ Q_1,\ Q_2,\ T_s)^{\mathsf{T}}, \qquad T_i(\mathbf{z}) = T_s + Q_i\,\mathrm{e}^{-g_i}, \qquad P(\mathbf{z}) = Q_1 + Q_2, \tag{23.1}$$

with g_i dimensionless, Q_i and the chiller's duty P in kW and the temperatures in °C; the hall temperatures and the duty are the points of the balance set that the free entries name. The five readings are the hall temperatures y_{T_1}, y_{T_2}, the supply temperature y_{T_s}, the chiller's heat meter y_P and the hall-1 load meter y_{Q_1}, with the standard deviations $\sigma_T = 0.5\,\mathrm{K}$, $\sigma_s = 0.3\,\mathrm{K}$, $\sigma_P = 2\,\mathrm{kW}$ and $\sigma_Q = 1.5\,\mathrm{kW}$ of one reading.

Table 23.1 lists the two records. Each line is one sensor or meter, with the standard deviation of one reading and of a ten-day average, the period-1 reading of Chapter 17, the period-2 value, and the value of the measured quantity from which the period-2 value was generated. Period 1 is one snapshot. Period 2 is illustrative: it was generated from the period-1 plant with hall 1's conductance lowered from 5.5 to 4.0 kW/K, hall 2's unchanged at $G_2 = 4.0\,\mathrm{kW/K}$, the loads and the supply temperature unchanged at $Q_1 = 30\,\mathrm{kW}$, $Q_2 = 24\,\mathrm{kW}$ and $T_s = 12°\mathrm{C}$, and no sensor or meter at fault, with errors of $+0.3$, -0.1 and $+0.2\,\mathrm{K}$, -1.0 and $-0.5\,\mathrm{kW}$ added in the order of the table. The period-2 values are the same at both record lengths: the chapter treats them once as the readings of a single day and once as the averages of ten days, so that the two analyses differ in precision only. A real ten-day average would lie closer to the truth than these values, two of which are about two ten-day standard deviations off.

Table 23.1: The two records of the two-hall plant. Period 1 is the single snapshot of Chapter 17 (its hall-1 value is 17.6545°C to four decimals). Period 2 is used once as one day and once as the average of ten daily readings, with standard deviation $\sigma/\sqrt{10}$. The last column is the value of the measured quantity in the plant that generated period 2.

Sensor or meter	sd, one day	sd, ten days	Period 1	Period 2	Generating value
hall-1 temperature y_{T_1}	0.5 K	0.158 K	17.655°C	19.80°C	19.50°C
hall-2 temperature y_{T_2}	0.5 K	0.158 K	17.700°C	17.90°C	18.00°C
supply temperature y_{T_s}	0.3 K	0.095 K	12.10°C	12.20°C	12.00°C
chiller heat meter y_P	2 kW	0.632 kW	52.5 kW	53.0 kW	54.0 kW
hall-1 load meter y_{Q_1}	1.5 kW	0.474 kW	31.0 kW	29.5 kW	30.0 kW

The record of period 1 was fitted in Chapter 18 with the handbook priors $g_i \sim \mathcal{N}(\ln 5,\ 0.3^2)$ and $T_s \sim \mathcal{N}(12,\ 1^2)$; its MAP state and spreads (Table 19.4) are

$$\begin{aligned} \hat{\mathbf{z}}_1 &= (1.7030,\ 1.3999,\ 30.688\,\text{kW},\ 22.231\,\text{kW},\ 12.111^\circ\text{C}), \\ \sqrt{\operatorname{diag} \boldsymbol{\Sigma}_1} &= (0.1067,\ 0.1370,\ 1.452\,\text{kW},\ 2.346\,\text{kW},\ 0.280\,\text{K}), \end{aligned} \tag{23.2}$$

that is $\hat{G}_1 = 5.490$ and $\hat{G}_2 = 4.055\,\text{kW/K}$, with a correlation of -0.029 between g_1 and g_2 (Table 19.3). Here and below a subscript 1 or 2 on a bold vector or matrix names the period: $\mathbf{y}_1$ and $\mathbf{y}_2$ are the two records, $\hat{\mathbf{z}}_1$ and $\boldsymbol{\Sigma}_1$ the MAP state and covariance of period 1. These are the given data of the chapter.

23.1.4 The questions (step 1)

Table 23.2 lists the five questions of the chapter. The first column states each in engineering language; the second the general form of its category, from Chapter 2; the third its mathematical form for the two halls; the fourth its category; the last the subsection where it is answered. The first is detection, a check of the healthy model read as a tail of its predictive; the second is isolation, the diagnosis category proper; the third is identification, an estimation within the leading drawings; the fourth is attribution, the posterior of the change between two fits; the fifth asks what the length of the record buys, a design question answered by the same computations at other precisions. Which reading to add next is not a row: it is the next measurement of step 7.

Table 23.2: The questions of fault detection and isolation on the two-hall plant, the general form of their category, their mathematical form, their category and where each is answered. $\mathbf{y}_1$, $\mathbf{y}_2$ are the two records; M_0 is the healthy model and $M_1, \dots, M_8$ the candidate drawings of Table 23.3; $\mathcal{L}_{M_0}$ is the objective of M_0; $\boldsymbol{\Delta} = \mathbf{z}_2 - \mathbf{z}_1$ is the change of the state between the periods, each fitted with the handbook priors; N is the number of days in the record of period 2.

Engineering question	General form	Mathematical form	Category	Answered in
Has the plant changed since period 1?	a tail of the predictive, equation (2.13)	$\Pr(\chi^2_3 > 2\mathcal{L}_{M_0}(\hat{\mathbf{z}}_{M_0}))$ under the healthy model; each reading against its leave-one-out predictive	check of the healthy model (detection)	23.3.1
Which component?	$\Pr(M_j \mid \mathbf{y}) \propto p(\mathbf{y} \mid M_j)\Pr(M_j)$, equation (2.15)	$\Pr(M_j \mid \mathbf{y}_2)$ for $j = 0, \dots, 8$ at equal prior odds, from the evidence of each drawing	diagnosis (isolation)	23.3.2
How large is the fault, and how well is it known?	$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y})$, equation (2.9)	$p(g_1 \mid \mathbf{y}_2, M_1)$ with the 95 % interval of G_1; $p(b_{T_1} \mid \mathbf{y}_2, M_5)$	estimation within the leading drawings (identification)	23.3.3
Which parameter moved between two periods? (without naming a fault)	$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y})$, equation (2.9)	$p(\boldsymbol{\Delta} \mid \mathbf{y}_1, \mathbf{y}_2)$: $Z_i = \hat{\Delta}_i / \sqrt{(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)_{ii}}$ and χ^2_Δ against χ^2_5	estimation of a change (attribution)	23.3.4
Could one day of readings have told?	$U(a)$, equation (2.16)	the answers to the first four questions as functions of N, at the observed averages	design (the length of the record)	23.3.5

23.1.5 The evidence as rows and the candidate family (step 2)

Step 2 attaches the two records to the graph as rows and extends the graph by the drawings that the questions compare. Each reading of period 2 is one row on the reduced state, the

average of its days with the standard deviation of one day divided by $\sqrt{N}$; the prior rows are those of the drawing.

The healthy model. The healthy model states what period 1 established about the plant, in a form that period 2 can be tested against. Two of its quantities are properties of the plant: the conductances of the two paths do not change unless something happens to the coils. The loads and the supply temperature are operating conditions: the loads follow the IT equipment and the supply follows the chiller's control. The healthy model therefore carries the period-1 posterior of the two log-conductances as their prior rows and gives the operating conditions the priors they had before any record, the flat prior on the loads and the set point on the supply temperature:

$$M_0: \quad g_1 \sim \mathcal{N}(g_1^{\mathrm{h}},\ \sigma_{\mathrm{h}1}^2), \quad g_2 \sim \mathcal{N}(g_2^{\mathrm{h}},\ \sigma_{\mathrm{h}2}^2), \quad T_s \sim \mathcal{N}(12,\ 1^2), \quad Q_1, Q_2 \text{ flat}, \tag{23.3}$$

with $g_1^{\mathrm{h}} = 1.7030$, $\sigma_{\mathrm{h}1} = 0.1067$, $g_2^{\mathrm{h}} = 1.3999$ and $\sigma_{\mathrm{h}2} = 0.1370$, the period-1 MAP values and marginal spreads of equation (23.2); the superscript and subscript h mark the healthy model. The symbols $\sigma_{\mathrm{h}1}$ and $\sigma_{\mathrm{h}2}$ are standard deviations of prior rows, as every σ of the book is; their values are the period-1 posterior spreads $\sqrt{(\mathbf{\Sigma}_1)_{11}}$ and $\sqrt{(\mathbf{\Sigma}_1)_{22}}$, the one place where a spread of one fit becomes the standard deviation of a row of the next. With the five period-2 readings and these three prior rows the healthy model has $n_r = 8$ rows over $n_z = 5$ entries, $\nu = 3$, the count of equation (18.3).

Candidates generated from the graph. A candidate drawing is the healthy model with one quantity freed. The plant graph says what can be freed, element by element, and the discipline of Chapter 2 is to write the list down before the period-2 record is examined, so that the record cannot suggest its own explanation. Three rules generate the candidates.

1. *Release a prior.* Every free quantity whose value the healthy model holds by a prior can have changed: a path's log-conductance (a fouled or a repaired coil), the supply temperature (a chiller that no longer holds its set point). The drawing keeps the prior's center and widens its spread.
2. *Add a bias.* Every reading can carry a bias: a sensor that has drifted, a meter that reads low. The drawing adds one entry to the state, attached to the reading's row, with a prior centered on zero (the bias entry of Chapter 22).
3. *Free a pinned quantity.* Every pinned quantity can be wrong. The two-hall plant pins only the reference conductance, which is a unit, so this rule generates nothing here; in a plant with a pinned ambient temperature or a pinned flow it would.

Applied to the two halls, the rules give eight drawings besides the healthy model. Rule 1 releases the two log-conductances (M_1, M_2) and the set point (M_8); rule 2 puts a bias on each of the five readings, the supply sensor (M_3), the chiller meter (M_4), the hall-1 sensor (M_5), the hall-2 sensor (M_6) and the hall-1 load meter (M_7); rule 3 gives none. Table 23.3 lists the family. Each line is one drawing, with the element of the graph it concerns, the quantity it frees, its prior and row, and its counts of rows, entries and redundancy. The last column marks M_0 and M_1 to M_4, whose evidence bookkeeping is worked in full: a path on each hall, the header's sensor, the chiller's meter. The answer to the isolation question is the ranking over all nine. The paths are released to a spread of $\sigma_{\mathrm{w}} = 1$ in the log-conductance, a factor of $\mathrm{e} = 2.7$ either way, which says that a fault may have moved the conductance anywhere a fouled or damaged coil could take it, and the set point to 3 K. The biases have the prior $\mathcal{N}(0,\ \sigma_b^2)$ with $\sigma_b = 2$ in the reading's units, that of Chapter 22. The chiller-meter bias is written as an offset b_P entering with a minus sign, so that the meter reads $Q_1 + Q_2 - b_P$, a meter that under-reads the heat it removes by b_P.

Figure 23.1 draws the family on the plant graph. The graph is that of Figure 17.1: green boxes are free nodes, white boxes sources, solid green arrows the physical paths of heat, dashed

Table 23.3: The candidate family generated by the three rules. Each drawing is the healthy model (23.3) with one quantity freed; the bias priors are $\mathcal{N}(0,\ 2^2)$ in the reading's units. n_r rows, n_z entries, $\nu = n_r - n_z$. Worked: the subset whose evidence bookkeeping is worked in full.

Drawing	Element of the graph	Freed quantity	Prior and row	n_r, n_z, ν	Worked
M_0	none (healthy)	none	equation (23.3)	8, 5, 3	yes
M_1	hall-1 path	g_1 (fouled path 1)	$g_1 \sim \mathcal{N}(1.7030,\ 1^2)$ in place of $\sigma_{\mathrm{h1}} = 0.1067$	8, 5, 3	yes
M_2	hall-2 path	g_2 (fouled path 2)	$g_2 \sim \mathcal{N}(1.3999,\ 1^2)$ in place of $\sigma_{\mathrm{h2}} = 0.1370$	8, 5, 3	yes
M_3	supply sensor	bias b [K]	the sensor reads $T_s + b$	9, 6, 3	yes
M_4	chiller heat meter	offset b_P [kW]	the meter reads $Q_1 + Q_2 - b_P$	9, 6, 3	yes
M_5	hall-1 sensor	bias b_{T_1} [K]	the sensor reads $T_1 + b_{T_1}$	9, 6, 3	–
M_6	hall-2 sensor	bias b_{T_2} [K]	the sensor reads $T_2 + b_{T_2}$	9, 6, 3	–
M_7	hall-1 load meter	bias b_{Q_1} [kW]	the meter reads $Q_1 + b_{Q_1}$	9, 6, 3	–
M_8	chiller set point	T_s	$T_s \sim \mathcal{N}(12,\ 3^2)$ in place of 1^2	8, 5, 3	–

boxes and arrows the readings. Each red outline marks the element a candidate drawing frees, with the drawing's name beside it: dashed for the worked subset, dotted for the other four. Every candidate sits on a different element of the graph.

Which row acts on which entry. The graph fixes where the Jacobian of the rows can be nonzero, marked $*$, before any number is computed. For the reduced state of equation (23.1) with the bias entry of M_5 appended as a sixth column,

$$\mathbf{J} = \begin{bmatrix} * & 0 & * & 0 & * & * \\ 0 & * & 0 & * & * & 0 \\ 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & * & * & 0 & 0 \\ 0 & 0 & * & 0 & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & * \end{bmatrix} \begin{matrix} \text{hall-1 reading} \\ \text{hall-2 reading} \\ \text{supply reading} \\ \text{chiller meter} \\ \text{load meter} \\ \text{prior, } g_1 \\ \text{prior, } g_2 \\ \text{prior, } T_s \\ \text{prior, } b_{T_1} \end{matrix}, \qquad \text{columns } g_1, g_2, Q_1, Q_2, T_s, b_{T_1}, \tag{23.4}$$

the first eight rows and five columns being the healthy model. Each other bias drawing moves the reading-row $*$ of the last column to the row of its own reading; a released prior changes the value of an entry of rows six to eight and not the pattern. The columns are the footprints of the candidates on the readings. The supply temperature reaches both hall rows and its own; the chiller offset reaches only the chiller row, which it shares with the two loads; and two columns, g_1 and b_{T_1}, reach the readings through the hall-1 row alone, each held otherwise only by its own prior row. Two drawings whose columns have the same footprint on the readings cannot be told apart by them.

The family is finite and its members are exclusive: each says that one thing, and only that thing, has changed since period 1. The healthy model M_0 belongs to the family, so that the statement that nothing has changed competes on the same footing as each fault. A fault outside the family, two faults at once, or a change that the drawings do not represent cannot be named by the ranking; Section 23.6.6 treats that limit.

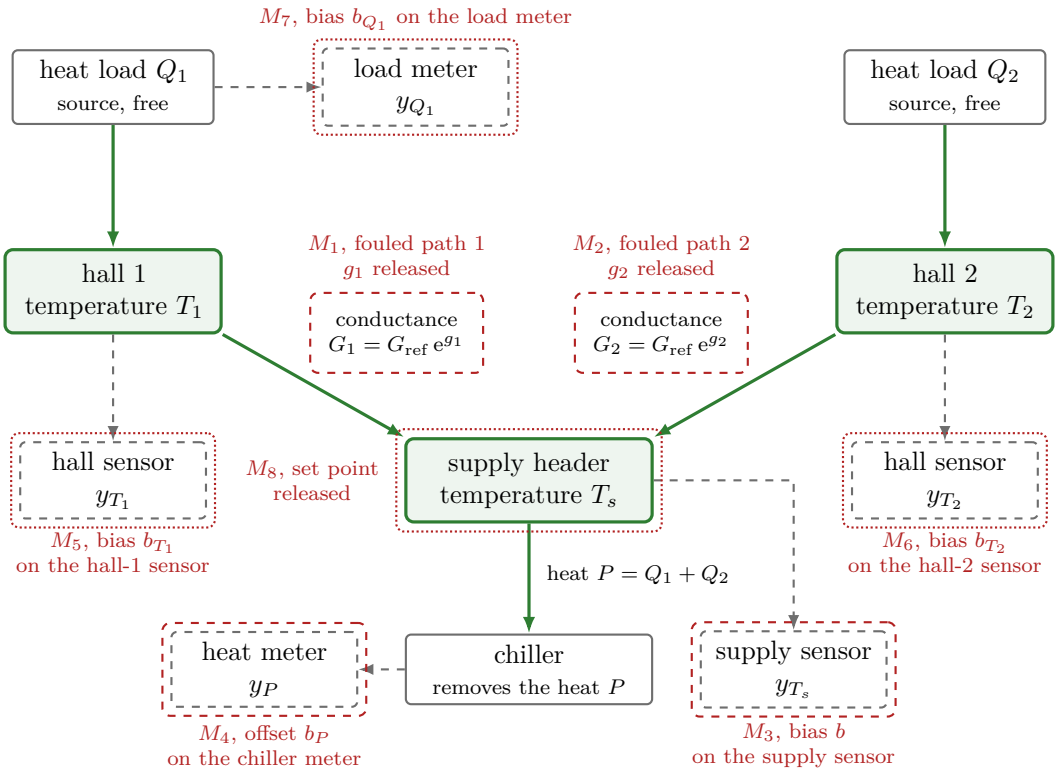


Figure 23.1: The candidate family on the plant graph of the two halls. Free: g_1, g_2, Q_1, Q_2 and T_s in every drawing, plus the freed bias or offset in M_3 to M_7. Pinned: $G_{\text{ref}} = 1\,\text{kW/K}$. Observed (dashed): the five readings of period 2. Red outlines: the element each drawing of Table 23.3 frees, dashed for the worked subset M_1 to M_4 and dotted for M_5 to M_8. The priors of the healthy model, equation (23.3), are rows of the evidence and are not drawn.

A record of several days as one reading. Let reading k be taken on N days, $y_{k1}, \dots, y_{kN}$, with the plant in the same steady state on each day, so that one state $\mathbf{z}$ holds for the whole period, and with independent errors $\mathcal{N}(0, \sigma_k^2)$. Write $h_k(\mathbf{z})$ for the value the model predicts for the reading and $\bar{y}_k = \frac{1}{N}\sum_j y_{kj}$ for the average. The likelihood of the N readings is

$$p(y_{k1}, \dots, y_{kN} \mid \mathbf{z}) = \prod_{j=1}^{N} \frac{1}{\sqrt{2\pi}\,\sigma_k} \exp\left[-\frac{\big(y_{kj} - h_k(\mathbf{z})\big)^2}{2\sigma_k^2}\right].$$

Adding and subtracting the average inside each square,

$$\sum_{j=1}^{N}\big(y_{kj} - h_k\big)^2 = \sum_{j=1}^{N}\big(y_{kj} - \bar{y}_k\big)^2 + 2\big(\bar{y}_k - h_k\big)\sum_{j=1}^{N}\big(y_{kj} - \bar{y}_k\big) + N\big(\bar{y}_k - h_k\big)^2,$$

and the middle sum is zero by the definition of the average. Therefore

$$p(y_{k1}, \dots, y_{kN} \mid \mathbf{z}) = F_k \exp\left[-\frac{\big(\bar{y}_k - h_k(\mathbf{z})\big)^2}{2\sigma_k^2/N}\right], \qquad F_k = \frac{1}{(\sqrt{2\pi}\,\sigma_k)^N} \exp\left[-\frac{\sum_j (y_{kj} - \bar{y}_k)^2}{2\sigma_k^2}\right]. \tag{23.5}$$

The factor F_k depends on the scatter of the daily readings about their average, not on the state, and it is the same under every drawing of the family, since every drawing is offered the same readings. It cancels from every posterior and from every ratio of evidences. What is left is the density of one reading $\bar{y}_k$ with standard deviation $\sigma_k/\sqrt{N}$: a record of N days enters the fit as its average, with the variance divided by N. That is how Table 23.1 was built.

The rows of a drawing. On the reduced state the rows of the healthy model are the eight rows of equation (17.15) with the averaged readings and the healthy prior rows,

$$\mathbf{r}_{M_0}(\mathbf{z}) = \begin{bmatrix} (T_s + Q_1\,\mathrm{e}^{-g_1} - \bar{y}_{T_1})\sqrt{N}/\sigma_T \\ (T_s + Q_2\,\mathrm{e}^{-g_2} - \bar{y}_{T_2})\sqrt{N}/\sigma_T \\ (T_s - \bar{y}_{T_s})\sqrt{N}/\sigma_s \\ (Q_1 + Q_2 - \bar{y}_P)\sqrt{N}/\sigma_P \\ (Q_1 - \bar{y}_{Q_1})\sqrt{N}/\sigma_Q \\ (g_1 - g_1^{\mathrm{h}})/\sigma_{\mathrm{h}1} \\ (g_2 - g_2^{\mathrm{h}})/\sigma_{\mathrm{h}2} \\ (T_s - 12)/1 \end{bmatrix}, \tag{23.6}$$

with $G_{\mathrm{ref}} = 1\,\mathrm{kW/K}$ absorbed and the temperature rows in kelvin before normalization. A released prior changes one standard deviation in rows six to eight: $\sigma_{\mathrm{h}1}$ to σ_{w} under M_1, $\sigma_{\mathrm{h}2}$ to σ_{w} under M_2, 1 to 3 K under M_8. A bias drawing appends the entry b to the state, adds it to the row of its reading (the hall-1 row becomes $(T_s + Q_1\mathrm{e}^{-g_1} + b_{T_1} - \bar{y}_{T_1})\sqrt{N}/\sigma_T$ under M_5; the chiller row carries $-b_P$ under M_4) and appends the ninth row b/σ_b. The handbook fits of attribution use the first five rows with the handbook prior rows $(g_i - \ln 5)/0.3$ and $(T_s - 12)/1$ in place of rows six to eight, one fit per period, with $N = 1$ for period 1. Every row is normalized, so the information matrix is $\mathbf{J}^{\mathsf{T}}\mathbf{J}$; written with the unnormalized differences and the weights $w_k = N/\sigma_k^2$ it is $\mathbf{J}^{\mathsf{T}}\mathbf{W}\mathbf{J}$, the same matrix.

The modeling choices. Five choices make the rows and the family, and each buys something and can fail. Each day of period 2 is taken at one steady state, the same on all the days: this is what lets the record enter as five averages, each with the variance of one day divided by the number of days (Section 23.1.5), and it fails if the loads or the supply moved from day to day, which Section 23.6.6 discusses. The healthy model is taken as the plant of period 1: this

is what makes a misfit a sign of change rather than of a poor handbook, and it misleads if period 1 itself carried a fault, which nothing in the two records tests. The family is the nine drawings of Table 23.3, written before period 2 was examined, each freeing one quantity, with the healthy model among them: this is what makes the posterior a normalized ranking that the record cannot tailor to itself, and it fails for a cause outside the list or for two causes at once, which the misfit of the leading drawing (Section 23.4.1) and Section 23.6.6 take up. A released prior keeps its center and takes the spread $\sigma_{\mathrm{w}} = 1$ (3 K for the set point), and a bias the prior $\mathcal{N}(0,\ \sigma_b^2)$ with $\sigma_b = 2$: this gives each drawing a definite Occam factor, and the ranking moves with these widths, which Section 23.4.2 varies; the release is symmetric, although fouling only lowers a conductance (Section 23.6.6). The evidence of each drawing is its Laplace evidence about its own MAP state: this reads each evidence from one fit, and it fails where the freed quantity's posterior is far from Gaussian, which importance sampling checks (Section 23.4.1).

23.2 The posterior of the state under each drawing (step 3)

The five questions of Table 23.2 rest on fits of one plant under several drawings, on the reduced state of Chapter 17. Four objects of each fit recur in the specialized forms: the MAP state $\hat{\mathbf{z}}_M$ of drawing M, the minimum of its objective $\mathcal{L}_M$; the information matrix $\mathbf{\Lambda}_M = \mathbf{J}^{\mathsf{T}}\mathbf{J}$ at it, from the Jacobian of the drawing's normalized rows, whose pattern is equation (23.4); the Laplace covariance $\mathbf{\Sigma}_M = \mathbf{\Lambda}_M^{-1}$; and the Laplace evidence of the drawing, read from the same three. The rows are those of step 2 (Section 23.1.5).

All five questions are read from the posteriors of the drawings: the posterior of the state that the rows of step 2 give under each drawing, with its Laplace evidence, and the handbook fits of the two periods. This section derives them and computes them for the two records; Section 23.3 reads every answer from them.

23.2.1 The posterior from the rows of each drawing

Every drawing is fitted on the reduced state of Chapter 17: the three balances are the limit $\sigma_p \to 0$ of their physics rows, the reduced Jacobian is the projection $\mathbf{J} = \mathbf{J}_{\mathrm{e}}\mathbf{N}$ of the rows of the evidence onto the null directions of the physics rows, equation (17.18), with the bold $\mathbf{N}$ the null-space basis of Chapter 17 and not the number of days N, and the reduced covariance is the free-entry marginal of the full one in that limit, equation (17.19). Section 23.4.4 checks the limit on the evidences and on the change between the periods.

The Jacobian. The rows of equation (23.6) are differentiated entry by entry; the only derivative that is not a constant is that of the product, $\partial(Q_i \mathrm{e}^{-g_i})/\partial g_i = -Q_i \mathrm{e}^{-g_i}$. The Jacobian of the healthy model is the 8×5 matrix

$$\mathbf{J}_{M_0}(\mathbf{z}) = \begin{bmatrix} -Q_1 \mathrm{e}^{-g_1}\sqrt{N}/\sigma_T & 0 & \mathrm{e}^{-g_1}\sqrt{N}/\sigma_T & 0 & \sqrt{N}/\sigma_T \\ 0 & -Q_2 \mathrm{e}^{-g_2}\sqrt{N}/\sigma_T & 0 & \mathrm{e}^{-g_2}\sqrt{N}/\sigma_T & \sqrt{N}/\sigma_T \\ 0 & 0 & 0 & 0 & \sqrt{N}/\sigma_s \\ 0 & 0 & \sqrt{N}/\sigma_P & \sqrt{N}/\sigma_P & 0 \\ 0 & 0 & \sqrt{N}/\sigma_Q & 0 & 0 \\ 1/\sigma_{\mathrm{h1}} & 0 & 0 & 0 & 0 \\ 0 & 1/\sigma_{\mathrm{h2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad (23.7)$$

in the columns $(g_1, g_2, Q_1, Q_2, T_s)$ and with the pattern of equation (23.4). The first entry of each hall row is minus the hall's temperature rise $T_i - T_s$ in units of the averaged sensor, the column of the conductance as the hall reading sees it. A bias drawing appends the column

$\sqrt{N}/\sigma_k$ in the row of its reading ($-\sqrt{N}/\sigma_P$ for the chiller offset) and $1/\sigma_b$ in the ninth row, a 9×6 matrix.

The fit. The normal equations of Chapter 17, equation (17.23), are solved at each iterate,

$$(\mathbf{J}^\mathsf{T}\mathbf{J})\,\mathbf{s} = -\mathbf{J}^\mathsf{T}\mathbf{r}, \qquad \mathbf{z}^{(j+1)} = \mathbf{z}^{(j)} + \mathbf{s}, \tag{23.8}$$

with the Gauss–Newton step written $\mathbf{s}$ here, since $\boldsymbol{\Delta}$ names the change between the periods, and with $\mathbf{J}$ and $\mathbf{r}$ of the drawing evaluated at $\mathbf{z}^{(j)}$, until the gradient $\mathbf{J}^\mathsf{T}\mathbf{r}$ vanishes, at the MAP state $\hat{\mathbf{z}}_M$. The information matrix and the Laplace covariance there are

$$\boldsymbol{\Lambda}_M = \mathbf{J}(\hat{\mathbf{z}}_M)^\mathsf{T}\mathbf{J}(\hat{\mathbf{z}}_M), \qquad \boldsymbol{\Sigma}_M = \boldsymbol{\Lambda}_M^{-1}. \tag{23.9}$$

The two hall rows are curved, through the products $Q_i \mathrm{e}^{-g_i}$, and the Hessian of the objective adds to $\boldsymbol{\Lambda}_M$ the curvature term $\sum_k r_k\, \partial^2 r_k / \partial \mathbf{z}\, \partial \mathbf{z}^\mathsf{T}$, the curvature of the balance set as the readings see it (Section 17.2.2); it is proportional to the hall rows' residuals at the MAP state, and the fits and the evidences of this chapter use the Gauss–Newton $\boldsymbol{\Lambda}_M$. The answers are read off the fit: an entry $\hat{z}_i$ with $\sqrt{\Sigma_{ii}}$, and a derived quantity f through its gradient $\mathbf{g}$ at $\hat{\mathbf{z}}_M$, with spread $\sqrt{\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}_M\mathbf{g}}$. The return temperature of the loop, $T_r = T_s + (Q_1 + Q_2)/(\dot{m}c_p)$ with the flow $\dot{m}c_p = 10\,\mathrm{kW/K}$ pinned (Chapter 21), has no sensor and the gradient $\mathbf{g}_r = (0,\ 0,\ 0.1\,\mathrm{K/kW},\ 0.1\,\mathrm{K/kW},\ 1)$.

The evidence of a drawing. The evidence of drawing M is the probability of the record under it,

$$p(\mathbf{y} \mid M) = \int p(\mathbf{y} \mid \mathbf{z}, M)\, p(\mathbf{z} \mid M)\, \mathrm{d}\mathbf{z}.$$

With the rows written normalized, the product under the integral is $C_M \exp[-\mathcal{L}_M(\mathbf{z})]$, where $C_M = c_Q \prod_k (\sqrt{2\pi}\,\sigma_k)^{-1}$ collects the normalizer of every row of M, reading or prior, and c_Q is the constant of the flat prior on the two loads, as in Chapter 22. The Laplace evidence of Chapter 4, equation (4.19), replaces $\mathcal{L}_M$ by its second-order expansion at $\hat{\mathbf{z}}_M$, equation (19.7), with the information matrix $\boldsymbol{\Lambda}_M$ of the drawing in place of the Hessian, and the remaining Gaussian integral is equation (19.9):

$$\int \exp[-\mathcal{L}_M(\mathbf{z})]\, \mathrm{d}\mathbf{z} \approx \exp[-\mathcal{L}_M(\hat{\mathbf{z}}_M)]\, (2\pi)^{n_z/2}\, (\det \boldsymbol{\Lambda}_M)^{-1/2}.$$

On the reduced state the physics rows have left no normalizer: in the limit of exact balances their constants cancel against the widths they give the posterior, as equation (4.19) shows for one hall. Taking logarithms,

$$\ln p(\mathbf{y} \mid M) \approx -\mathcal{L}_M(\hat{\mathbf{z}}_M) - \tfrac{1}{2} \ln \det \boldsymbol{\Lambda}_M + \tfrac{n_z}{2} \ln 2\pi - \sum_{k \in M} \ln(\sqrt{2\pi}\,\sigma_k) + \ln c_Q. \tag{23.10}$$

Chapter 22 used this formula for a bias drawing, and Chapter 24 derives the evidence in general and states when its Laplace form holds. For two drawings of one family the reading rows are the same rows with the same standard deviations, and both carry the flat prior on the loads: those terms cancel from the ratio. What remains is the misfit, the curvature, the count of entries and the normalizers of the prior rows, and the two kinds of drawing in Table 23.3 differ in the last two.

23.2.2 The posterior of the two halls under each drawing

The plant's inputs are the two records of Table 23.1, period 2 at the precision of one day and of ten days; the period-1 posterior of equation (23.2); the release spreads $\sigma_\mathrm{w} = 1$ and $3\,\mathrm{K}$ and the bias spread $\sigma_b = 2$.

Table 23.4 runs the normal equations (23.8) for the healthy model on the ten-day record, from the period-1 MAP state $\hat{\mathbf{z}}_1$. Each line is an iteration; the columns give the five entries of the state and twice the objective. The first step moves almost all the way, lowering g_1 by 0.36 to meet the warm hall 1, and the misfit falls from 187.2 to 9.887 in three steps.

Table 23.4: Gauss–Newton iterations of equation (23.8) for the healthy model M_0 on the ten-day record of period 2, rows of equation (23.6), from $\hat{\mathbf{z}}_1$ of equation (23.2).

Iteration	g_1 [–]	g_2 [–]	Q_1 [kW]	Q_2 [kW]	T_s [°C]	$2\mathcal{L}_{M_0}$ [–]
0	1.70298	1.39993	30.6882	22.2307	12.11145	187.158
1	1.34585	1.41400	29.7376	23.2489	12.24706	13.193
2	1.37928	1.41143	29.7211	23.2684	12.23002	9.889
3	1.38037	1.41162	29.7206	23.2687	12.23102	9.887
4	1.38039	1.41162	29.7206	23.2687	12.23106	9.887

At the MAP state $\hat{\mathbf{z}}_{M_0} = (1.3804,\ 1.4116,\ 29.721\,\text{kW},\ 23.269\,\text{kW},\ 12.231°\text{C})$ the Jacobian (23.7) and the Laplace covariance (23.9) are

$$\mathbf{J}_{M_0}(\hat{\mathbf{z}}_{M_0}) = \begin{bmatrix} -47.271 & 0 & 1.591 & 0 & 6.325 \\ 0 & -35.871 & 0 & 1.542 & 6.325 \\ 0 & 0 & 0 & 0 & 10.541 \\ 0 & 0 & 1.581 & 1.581 & 0 \\ 0 & 0 & 2.108 & 0 & 0 \\ 9.373 & 0 & 0 & 0 & 0 \\ 0 & 7.300 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{\Sigma}_{M_0} = \begin{bmatrix} 0.00080 & -0.00010 & 0.00699 & -0.00690 & 0.00112 \\ -0.00010 & 0.00198 & -0.00859 & 0.02397 & 0.00142 \\ 0.00699 & -0.00859 & 0.21593 & -0.20806 & -0.00001 \\ -0.00690 & 0.02397 & -0.20806 & 0.58610 & -0.00129 \\ 0.00112 & 0.00142 & -0.00001 & -0.00129 & 0.00868 \end{bmatrix}.$$

One entry in full: the g_1 entry of the hall-1 row is $-(T_1 - T_s)\sqrt{10}/\sigma_T = -(19.705 - 12.231) \times 6.3246 = -47.27$. The spreads $\sqrt{\operatorname{diag} \mathbf{\Sigma}_{M_0}}$ are $(0.0283,\ 0.0444,\ 0.465\,\text{kW},\ 0.766\,\text{kW},\ 0.093\,\text{K})$, and the largest correlations are $\operatorname{corr}(g_2, Q_2) = 0.704$, $\operatorname{corr}(Q_1, Q_2) = -0.585$ and $\operatorname{corr}(g_1, Q_1) = 0.532$. The other drawings, the one-day record and the handbook fits of the two periods converge in the same way. Each is this construction on its own state and rows: a widened drawing changes the spread of one prior row, a bias drawing appends its entry with the column and the ninth row of equation (23.7), six entries in all, and the handbook fit of a period carries the handbook priors of Table 17.2. Each gives its $\hat{\mathbf{z}}_M$, $\mathbf{\Sigma}_M$ by equation (23.9) and its Laplace evidence by equation (23.10), and every answer of step 4 is computed from them; their numbers are given with the answers that use them.

23.3 The answers (step 4)

The five questions of Table 23.2 are answered in its order: detection, isolation, identification, attribution and the length of the record. Each is answered from the posteriors of Section 23.2: its engineering question, the general form of its category, the form it takes for the two halls, the quantities of the posterior it uses, and the plant's numbers. The questions name equal prior odds and, for the record length, the period-2 averages at $N = 1$ to 1000 days. The reported route for every evidence is the Laplace route, from the fit of each drawing; importance sampling

of the evidence integral, which is exact to its Monte Carlo error, and the two exact one-entry identities give the values set beside it in the same answer, with the gap.

23.3.1 The misfit of the healthy model and the readings one at a time

Engineering question. Has the plant changed since period 1? The operators ask whether period 2 is consistent with the plant as period 1 described it, and, if it is not, which readings disagree.

General form. A check of a model against a record is the predictive of Chapter 2,

$$p(y_{\text{new}} \mid \mathbf{y}) = \int p(y_{\text{new}} \mid \mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

equation (2.13), read as a tail: the probability that a record drawn from it would disagree with the model at least as much as the record in hand.

Specialized form. The two checks are tails of equation (2.13) under the healthy model. For the first, y_{new} is a replicate of the whole record, read on its misfit, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}_{M_0}, \boldsymbol{\Sigma}_{M_0})$; for the second, $y_{\text{new}} = \bar{y}_k$, one averaged reading, $p(y_{\text{new}} \mid \mathbf{z}) = \mathcal{N}(f_k(\mathbf{z}),\ \sigma_k^2/N)$, and $p(\mathbf{z} \mid \mathbf{y})$ the same construction on the record without that reading, $\mathcal{N}(\hat{\mathbf{z}}_{-k}, \boldsymbol{\Sigma}_{-k})$. One choice enters here. The reference distribution of the misfit is the chi-square of Chapter 18: this gives the tail in closed form, exactly for linear rows with Gaussian errors of the stated standard deviations, and it fails if those standard deviations are misstated, which the daily scatter would test (Section 23.4.4).

The first check is the misfit of M_0 on the period-2 record. If its rows and their standard deviations are right, twice its minimized objective is approximately chi-square with $\nu = n_r - n_z = 3$ degrees of freedom, equation (18.24), and the tail is in closed form, equation (18.27):

$$\Pr(\chi^2_3 > x) = \sqrt{\frac{2x}{\pi}}\, \mathrm{e}^{-x/2} + 2[1 - \Phi(\sqrt{x})], \qquad x = 2\mathcal{L}_{M_0}(\hat{\mathbf{z}}_{M_0}).$$

A value far in the upper tail says that the record and the healthy model disagree. It detects; it does not locate.

The second check takes the readings one at a time. Reading k is compared with its predictive from the healthy model fitted without it, the device of equation (21.21): with $\hat{\mathbf{z}}_{-k}$ and $\boldsymbol{\Sigma}_{-k} = (\mathbf{J}_{-k}^{\mathsf{T}}\mathbf{J}_{-k})^{-1}$ the MAP state and covariance of M_0 with row k removed, $f_k(\mathbf{z})$ the quantity the reading measures and $\mathbf{g}_k$ its gradient at $\hat{\mathbf{z}}_{-k}$, the predictive is normal with mean $f_k(\hat{\mathbf{z}}_{-k})$ and spread

$$\mathrm{sd}(y_k \mid \mathbf{y}_-) = \sqrt{\sigma_k^2/N + \mathbf{g}_k^{\mathsf{T}}\boldsymbol{\Sigma}_{-k}\mathbf{g}_k},$$

with $\mathbf{y}_-$ the record without reading k, and the standardized departure $d_k = (\bar{y}_k - f_k(\hat{\mathbf{z}}_{-k}))/\mathrm{sd}(y_k \mid \mathbf{y}_-)$ of equation (21.22) is standard normal under the model. The reading was not used, so its departure is not absorbed by the entry it moves, the weakness of the residual of the full fit (Section 18.3.2); for linear rows d_k is minus the normalized residual, equation (21.24), and with curved rows the two differ. The pattern of the d_k is a first hint of where a change lies. Two rows of one chain, a reading and the prior row of the only entry it sees besides the shared ones, share their normalized residual and cannot be split by it (Section 18.3.2).

Building blocks. The fit of M_0 on the averaged rows, $\hat{\mathbf{z}}_{M_0}$, $\mathcal{L}_{M_0}$ and $\mathbf{J}$ (Section 23.2.1); the tail of equation (18.27) and Φ; for each reading the fit without its row, the gradient $\mathbf{g}_k$ and $\boldsymbol{\Sigma}_{-k}$.

Detection. Table 23.5 collects the detection checks. Its first rows give the misfit of the healthy model at both record lengths, the share of it carried by the g_1 prior row, its tail and the fitted $\hat{G}_1$; its last rows give the leave-one-out departure d_k of each reading, equation (21.22).

Table 23.5: Detection: the healthy model M_0 fitted to period 2 at the precision of one day and of ten days. The departures d_k are (reading minus leave-one-out prediction) over the held-out predictive spread $\mathrm{sd}(y_k \mid \mathbf{y}_-)$.

Quantity	One day	Ten days
rows n_r, entries n_z, redundancy ν	8, 5, 3	8, 5, 3
$2\mathcal{L}_{M_0}(\hat{\mathbf{z}}_{M_0})$	6.04	9.89
from the prior row of g_1	3.26	9.14
tail $\Pr(\chi^2_3 > 2\mathcal{L}_{M_0})$ (95 % point 7.81)	0.110	0.020
fitted $\hat{G}_1$ under M_0 [kW/K]	4.53	3.98
d_k, hall-1 sensor	+2.61	+3.66
d_k, hall-2 sensor	+0.02	−0.09
d_k, supply sensor	−1.38	−1.86
d_k, chiller heat meter	−0.02	+0.09
d_k, hall-1 load meter	−1.63	−2.03

With one day of readings the healthy model fits with $2\mathcal{L}_{M_0} = 6.04$ on three degrees of freedom. By equation (18.27), with $\sqrt{6.04} = 2.458$, $\Phi(2.458) = 0.99301$ and $\mathrm{e}^{-3.021} = 0.04875$,

$$\Pr(\chi^2_3 > 6.04) = \sqrt{\tfrac{2\times 6.04}{\pi}} \times 0.04875 + 2 \times 0.00699 = 1.961 \times 0.04875 + 0.01397 = 0.110 :$$

a misfit this large or larger would occur in one record in nine from an unchanged plant, so one day says that the plant may have changed and does not say that it has. With ten days the statistic is 9.89 and the tail 0.020: the healthy model is rejected at the 2 % level. The statistic did not grow tenfold. At one day the fit splits the conflict between the hall-1 reading and the g_1 prior, lowering G_1 to 4.53 kW/K; at ten days the reading is stiff, the fit lowers G_1 to 3.98, and 9.14 of the 9.89 sit on the prior row of g_1, which the fit now misses by 3.0 of its spreads.

Figure 23.2 sets the readings beside the healthy model's predictions. The left panel holds the three temperatures and the right panel the two meters, each on the horizontal axis; the vertical axis is the reading in its units. For each reading the open black circle is the period-1 value with its $\pm 2\sigma$ bar, the filled red circle the period-2 value with its ten-day $\pm 2\sigma/\sqrt{10}$ bar, and the gray box the healthy model's leave-one-out prediction of that reading with its ± 2 standard deviation band, the thin line inside it the predicted value. The hall-1 reading lies 3.66 predictive standard deviations above its predicted value, 1.0 K (another 1.66 standard deviations) above the upper edge of its box; the hall-2 sensor and the chiller meter sit in the middle of theirs. The supply sensor and the hall-1 load meter sit at the lower edges of their boxes. Without the supply sensor, the healthy model explains the warm hall 1 in part by a warmer supply, and predicts 13.06°C; without the load meter, in part by a larger load, and predicts 34.1 kW. Each departure is the trace of one explanation of the same warm hall, and the readings that pin the supply temperature and the load are what make those explanations testable. The pattern points at hall 1; it does not say whether the path, the supply or the sensor is at fault. The normalized residuals of the healthy fit say the same in their own way: at ten days the hall-1 row and the g_1 prior row share the largest, −3.14, the two rows of one chain.

23.3.2 The posterior over the candidate family

Engineering question. Which component? Given that the plant may have changed, the operators ask which element of the graph, a path, a source, a sensor or a meter, changed, among the candidates the graph generates.

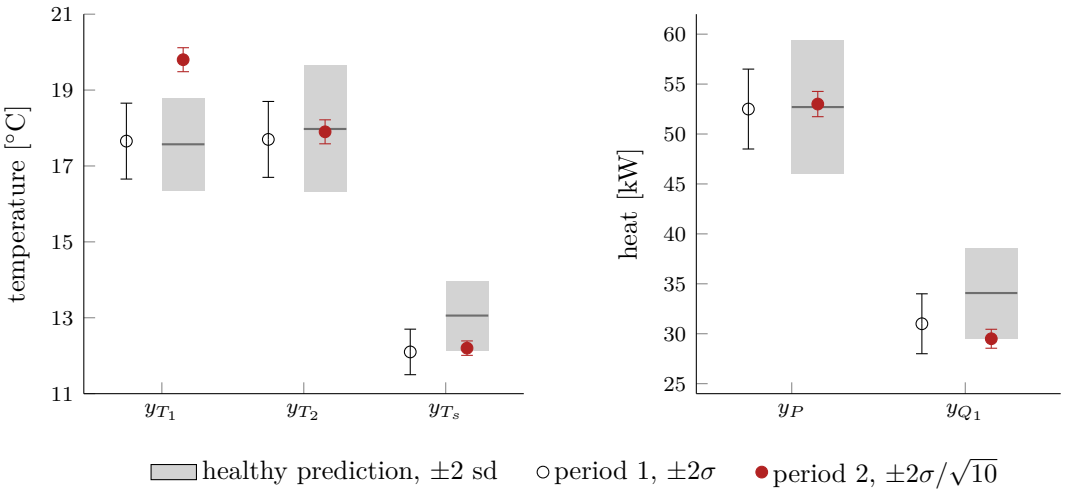


Figure 23.2: The readings of the two periods against the healthy model. Open circles: period 1 with $\pm 2\sigma$ of one reading; filled circles: period 2 with $\pm 2\sigma/\sqrt{10}$ of a ten-day average; gray boxes: the leave-one-out prediction of each period-2 reading under M_0, equation (21.21), ±2 predictive standard deviations, with the predicted value as the line inside. The departures are those of Table 23.5.

General form. A diagnosis ranks a family of candidate faults by their posterior probabilities,

$$\Pr(M_j \mid \mathbf{y}) \propto p(\mathbf{y} \mid M_j)\, \Pr(M_j),$$

equation (2.15), with the evidence $p(\mathbf{y} \mid M_j)$ of equation (2.26).

Specialized form. The ranking is equation (2.15) with the family $\mathcal{F}$ the drawings of Table 23.3, $\mathbf{y} = \mathbf{y}_2$ the period-2 record, and each evidence $p(\mathbf{y} \mid M_j)$ the integral of the drawing's rows over its state, evaluated by the Laplace form (23.10) at the drawing's posterior of step 3, $\hat{\mathbf{z}}_{M_j}$ and $\mathbf{\Lambda}_{M_j}$. One choice enters here. The drawings have equal prior odds: this lets the record alone speak, and it ignores a maintenance history of fouled coils or drifting sensors, which enters through $\Pr(M_j)$ (Exercise 23.1).

With the drawing as an unknown taking the values in a family $\mathcal{F}$, equation (2.46) gives its posterior. Dividing the numerator and the denominator by $p(\mathbf{y}_2 \mid M_0)$ and writing $\ln B_{j0}$ for the log Bayes factor of M_j against the healthy model,

$$\Pr(M_j \mid \mathbf{y}_2) = \frac{\Pr(M_j)\, \mathrm{e}^{\ln B_{j0}}}{\sum_{k \in \mathcal{F}} \Pr(M_k)\, \mathrm{e}^{\ln B_{k0}}}, \qquad \ln B_{00} = 0. \tag{23.11}$$

At equal prior odds, $\Pr(M_j) = 1/|\mathcal{F}|$ with $|\mathcal{F}|$ the number of drawings, the prior probabilities cancel and the posterior is the normalized exponential of the log Bayes factors. The denominator runs over the listed drawings only, so the probabilities sum to one over the family: they compare the listed drawings with one another and say nothing about a drawing that is not listed. Over the full family of nine the ranking is the answer to the isolation question; over the worked subset of five it shows how a ranking forms.

A widened prior. M_1 has the rows and entries of M_0; only the standard deviation of the g_1 prior row changes, from $\sigma_{\mathrm{h}1}$ to σ_{w}, and with it that row's normalizer, from $-\ln(\sqrt{2\pi}\,\sigma_{\mathrm{h}1})$ to $-\ln(\sqrt{2\pi}\,\sigma_{\mathrm{w}})$. The difference of equation (23.10) between the two drawings is

$$\ln B_{10} = -\left[\mathcal{L}_{M_1}(\hat{\mathbf{z}}_{M_1}) - \mathcal{L}_{M_0}(\hat{\mathbf{z}}_{M_0})\right] - \tfrac{1}{2} \ln \frac{\det \mathbf{\Lambda}_{M_1}}{\det \mathbf{\Lambda}_{M_0}} + \ln \frac{\sigma_{\mathrm{h}1}}{\sigma_{\mathrm{w}}}, \tag{23.12}$$

with $B_{10} = p(\mathbf{y}_2 \mid M_1)/p(\mathbf{y}_2 \mid M_0)$: a misfit term, a curvature term and a constant. The constant is the ratio of the two normalizers. It is negative whenever the prior is widened, the price of the freedom, and only the misfit term can pay it.

An added entry. M_3 to M_7 each add one entry and its prior row, and their log Bayes factor is that of the bias entry, equation (22.13): the misfit and curvature terms of equation (23.12) with the constant $-\ln \sigma_b$ in place of the normalizer ratio, the $\sqrt{2\pi}$ of the new row cancelling against the $\frac{1}{2}\ln 2\pi$ of the new entry.

What the curvature term carries. The curvature term reads $\det \mathbf{\Lambda}_M$, the volume the drawing's posterior occupies. Widening the prior of an entry that precise readings already hold barely changes that volume, since the readings, not the prior, set the entry's spread; the term is then close to zero and the misfit relief and the constant decide. A bias on a reading that reaches several entries, such as the supply sensor, spreads the posterior along a direction the other rows hold weakly, and the term is negative. A bias whose column shares its only reading row with entries that nothing else pins, as the chiller offset shares the chiller row with the load Q_2, which has no prior and is otherwise seen only through the hall-2 row, keeps its posterior close to its prior; with its estimate near zero its log Bayes factor is the Savage–Dickey ratio, the ratio of its posterior to its prior spread, and the drawing neither gains nor loses much.

The Occam factor in closed form. The structure of equation (23.12) is clearest when the other rows are reduced to what they say about the freed entry. Integrate the other entries out of the likelihood under their priors, which the two drawings share; for linear rows what remains is a Gaussian likelihood of g_1 alone, with estimate $\hat{g}_1^{\mathrm{L}}$ and variance v_{L} (L for that likelihood), times a factor that does not depend on g_1 and is common to both drawings. Under a prior $\mathcal{N}(g_1^{\mathrm{h}}, s^2)$ the estimate is the sum of the true g_1 and an independent error, so its marginal density is $\mathcal{N}(\hat{g}_1^{\mathrm{L}};\ g_1^{\mathrm{h}},\ s^2 + v_{\mathrm{L}})$. The ratio of the two marginals is the Bayes factor,

$$\ln B_{10} = \underbrace{\tfrac{1}{2}\ln \frac{\sigma_{\mathrm{h}1}^2 + v_{\mathrm{L}}}{\sigma_{\mathrm{w}}^2 + v_{\mathrm{L}}}}_{\text{Occam factor}} + \underbrace{\tfrac{1}{2}(\hat{g}_1^{\mathrm{L}} - g_1^{\mathrm{h}})^2 \left[\frac{1}{\sigma_{\mathrm{h}1}^2 + v_{\mathrm{L}}} - \frac{1}{\sigma_{\mathrm{w}}^2 + v_{\mathrm{L}}}\right]}_{\text{relief of the misfit}}, \tag{23.13}$$

exact for linear rows with v_{L} so defined; for the plant, $\hat{g}_1^{\mathrm{L}}$ and v_{L} are approximated by the mean and variance of g_1 under M_1, whose wide prior adds almost nothing to them. The first term is the Occam factor of the widening, close to $\ln(\sigma_{\mathrm{h}1}/\sigma_{\mathrm{w}})$ when the readings are precise. The second is half the squared conflict between the readings and the healthy model, measured in the healthy spread, less its value in the widened one. Two properties follow. Setting the derivative with respect to σ_{w}^2 to zero, $-\frac{1}{2}/(\sigma_{\mathrm{w}}^2 + v_{\mathrm{L}}) + \frac{1}{2}(\hat{g}_1^{\mathrm{L}} - g_1^{\mathrm{h}})^2/(\sigma_{\mathrm{w}}^2 + v_{\mathrm{L}})^2 = 0$, the evidence is largest for $\sigma_{\mathrm{w}}^2 + v_{\mathrm{L}} = (\hat{g}_1^{\mathrm{L}} - g_1^{\mathrm{h}})^2$, a widening matched to the size of the change, and it falls like $-\ln \sigma_{\mathrm{w}}$ beyond: the ranking depends on the stated width of the freed prior. When the entry needs no change, $\hat{g}_1^{\mathrm{L}} = g_1^{\mathrm{h}}$, only the Occam factor remains and the drawing loses.

Two drawings with one footprint. The pattern of equation (23.4) says which pairs of drawings the evidence can separate. The columns of g_1 under M_1 and of b_{T_1} under M_5 reach the readings through the hall-1 row alone, $(T_s + Q_1 \mathrm{e}^{-g_1} + b_{T_1} - \bar{y}_{T_1})\sqrt{N}/\sigma_T$, so for every value of one there is a value of the other that leaves every reading row unchanged. The two drawings then predict the same record and differ only in their prior rows and their constants, and their evidence ratio is set by those: by how far each must move its freed quantity in units of that quantity's prior, and by the price each pays for its freedom. The reading rows do not decide between them, whatever their precision.

A bias on the shared header. The supply temperature's column reaches both hall rows and its own. A supply sensor that reads low therefore makes every temperature rise smaller than it looks, raises both fitted conductances and puts the true supply above the set point. When only one hall has warmed, that explanation is partial: it spreads the conflict on the warm hall's conductance over the other conductance and the set-point prior instead of removing it, so its misfit relief is smaller than that of the fouled path. It predicts the readings almost as the fouled path does and disagrees with it on quantities no sensor reads, the true supply temperature and the return temperature of the loop, which is where a new reading can separate the two (Section 23.6.2). Releasing the set point instead frees an entry that the supply sensor reads directly, and it relieves a misfit only where that sensor and the set point disagree.

Two exact identities. The Savage–Dickey ratio of Chapter 7, equation (7.68), applies to an added entry, since M_3 with $b = 0$ is M_0: $B_{03} = p(b = 0 \mid \mathbf{y}_2, M_3)/p(b = 0 \mid M_3)$. A widened prior is not nested at a point, but the same argument gives the ratio as an expectation. Writing p_0 and p_1 for the two priors of g_1, the other priors being common,

$$p(\mathbf{y} \mid M_0) = \int p(\mathbf{y} \mid \mathbf{z})\, p_1(\mathbf{z})\, \frac{p_0(g_1)}{p_1(g_1)}\, \mathrm{d}\mathbf{z} = p(\mathbf{y} \mid M_1) \int p(\mathbf{z} \mid \mathbf{y}, M_1)\, \frac{p_0(g_1)}{p_1(g_1)}\, \mathrm{d}\mathbf{z},$$

so $B_{01} = \mathbb{E}\big[p_0(g_1)/p_1(g_1) \mid \mathbf{y}, M_1\big]$: the posterior expectation, under the widened drawing, of the ratio of the two priors. Both identities are exact; each needs the posterior of one entry, and they are only as good as the density used for it. Evaluated with the Laplace marginal of that entry, they give values set beside the reported ones below.

Building blocks. The fit of each drawing, $\hat{\mathbf{z}}_M$, $\mathcal{L}_M(\hat{\mathbf{z}}_M)$ and $\ln\det\boldsymbol{\Lambda}_M$ (Section 23.2.1); the constants $\ln(\sigma_{\mathrm{h}i}/\sigma_{\mathrm{w}})$ and $-\ln\sigma_b$; the mean and variance of g_1 under M_1 for the closed form; the normalized exponential of equation (23.11). For the values beside them: an importance-sampling estimate of each evidence integral, 200 000 draws from a Student-t density with five degrees of freedom centered at $\hat{\mathbf{z}}_M$ and scaled to $1.5^2\boldsymbol{\Sigma}_M$, with a Monte Carlo standard error of 0.004 nats on a log Bayes factor; and the two identities with the Laplace marginal of the freed entry.

The evidence bookkeeping on the worked subset. Table 23.6 works the evidence of the subset M_0 to M_4 term by term at the two record lengths; its probabilities are over those five drawings only. Each line is one drawing; the columns give twice its minimized objective, the three terms of its log Bayes factor against M_0 by equations (23.12) and (22.13), their sum $\ln B_{j0}$, and the posterior probability at equal prior odds, equation (23.11). The three terms are the misfit term $-[\mathcal{L}_{M_j} - \mathcal{L}_{M_0}]$, the curvature term $-\frac{1}{2}\ln(\det\boldsymbol{\Lambda}_{M_j}/\det\boldsymbol{\Lambda}_{M_0})$ and the constant.

The bookkeeping, one line each. At ten days the fouled path of hall 1 relieves the misfit by half of $9.887 - 0.173$, that is 4.857, changes the curvature by 0.014 and pays $\ln(\sigma_{\mathrm{h}1}/\sigma_{\mathrm{w}}) = \ln(0.1067/1) = -2.238$:

$$\ln B_{10} = 4.857 + 0.014 - 2.238 = 2.633.$$

The supply-sensor bias relieves 1.465, loses 0.562 in curvature and pays $-\ln 2$: $1.465 - 0.562 - 0.693 = 0.210$. Freeing g_2 gives $0.004 + 0.055 - 1.988 = -1.929$, the constant being $\ln(0.1370/1)$, within 0.006 of the Occam factor $\frac{1}{2}\ln[(\sigma_{\mathrm{h}2}^2 + v_{\mathrm{L}})/(1 + v_{\mathrm{L}})] = -1.935$ of equation (23.13) with $v_{\mathrm{L}} = 0.046^2$. The chiller-meter offset has the posterior $\hat{b}_P = -0.08 \pm 1.72\,\mathrm{kW}$ and the log Bayes factor $\ln(1.719/2) = -0.151$, its Savage–Dickey ratio at $\hat{b}_P \approx 0$. The closed form of equation (23.13), with $\hat{g}_1^{\mathrm{L}} - g_1^{\mathrm{h}} = 1.3563 - 1.7030 = -0.3467$ and $v_{\mathrm{L}} = 0.0291^2 = 0.000847$, gives

$$\ln B_{10} \approx \tfrac{1}{2}\ln\frac{0.012229}{1.000847} + \tfrac{1}{2} \times 0.12017 \times \Big(\frac{1}{0.012229} - \frac{1}{1.000847}\Big) = -2.202 + 4.853 = 2.65,$$

against 2.633 from the full fits.

Table 23.6: The evidence bookkeeping of the worked subset on period 2. Log Bayes factors in nats against the healthy model, split into the three terms of equations (23.12) and (22.13); posterior probabilities at equal prior odds over the five drawings of the subset.

Drawing	$2\mathcal{L}_{M_j}$	Misfit	Curvature	Constant	$\ln B_{j0}$	$\Pr(M_j \mid \mathbf{y}_2)$
One day						
M_0 healthy	6.04	–	–	–	0	0.190
M_1 fouled path 1	0.16	2.940	0.171	−2.238	+0.87	0.456
M_2 fouled path 2	6.04	0.000	0.393	−1.988	−1.59	0.039
M_3 supply-sensor bias	4.38	0.829	−0.436	−0.693	−0.30	0.141
M_4 chiller-meter offset	6.04	0.000	0.606	−0.693	−0.09	0.174
Ten days						
M_0 healthy	9.89	–	–	–	0	0.058
M_1 fouled path 1	0.17	4.857	0.014	−2.238	+2.63	0.811
M_2 fouled path 2	9.88	0.004	0.055	−1.988	−1.93	0.008
M_3 supply-sensor bias	6.96	1.465	−0.562	−0.693	+0.21	0.072
M_4 chiller-meter offset	9.88	0.001	0.541	−0.693	−0.15	0.050

The values beside them. Importance sampling of the evidence integrals gives, at ten days, $\ln B_{10} = 2.634$ against the reported 2.633, 0.210 reported against 0.227 for the supply bias and −0.150 against −0.151 for the chiller offset; at one day 0.852 against 0.873, −0.302 against −0.300 and −0.091 against −0.087. The largest gap is 0.021 nats, at one day for the fouled path, above the Monte Carlo error and so the error of the Laplace route itself; no ranking changes. The two identities, evaluated with the Laplace marginal, give 2.658 and 1.059 for the fouled path at ten days and one day (prior-ratio expectation), and 0.138 and −0.362 for the supply bias and −0.151 and −0.087 for the chiller offset (Savage–Dickey). They agree where the marginal is read near its center, the chiller offset whose estimate is near zero, and err where it is read in its tail: by 0.19 nats for the fouled path at one day, where the healthy prior's region lies 1.6 posterior standard deviations from the center of the M_1 posterior.

Figure 23.3 draws the four log Bayes factors. The horizontal axis lists the drawings and the vertical axis is $\ln B_{j0}$ in nats, zero being the healthy model; for each drawing the gray bar is the one-day record and the blue bar the ten-day record, with its value above or below it. One day of readings ranks the fouled hall-1 path first, +0.87, odds of 2.4 to one against the healthy model and a posterior probability of 0.46. Ten days raise it to +2.63, odds of 13.9 to one, and its probability to 0.81; the other drawings change by 0.51 nats or less, the fouled hall-2 path falling further.

The supply-sensor bias and the true supply temperature. Table 23.7 sets three ten-day fits side by side. Each line is a row of the model or a derived quantity; the columns give the period-2 reading or prior value and the fitted value under the healthy model, the fouled hall-1 path and the supply-sensor bias, with the standardized residual of the row in parentheses. The last lines give each drawing's $2\mathcal{L}$ and two quantities that no sensor reads: the true supply temperature and the return temperature of the loop.

Under M_3 the supply sensor reads 0.81 K low and the true supply is 13.02°C. Hall 1's rise of $19.73 - 13.02 = 6.71$ K at 29.7 kW gives $G_1 = 4.43$, and hall 2's rise of 4.92 K at 23.1 kW gives $G_2 = 4.70$ kW/K. The prior rows of g_1, g_2 and the set point stand at −2.02, +1.08 and +1.02 spreads (−3.02, +0.09 and +0.23 under M_0) and carry 6.3 of its $2\mathcal{L} = 6.96$. M_1 and M_3 predict the five readings within 0.07 K and 0.21 kW of each other. The true supply is 12.20 under M_1 and 13.02°C under M_3, and the return 17.50 and 18.30°C.

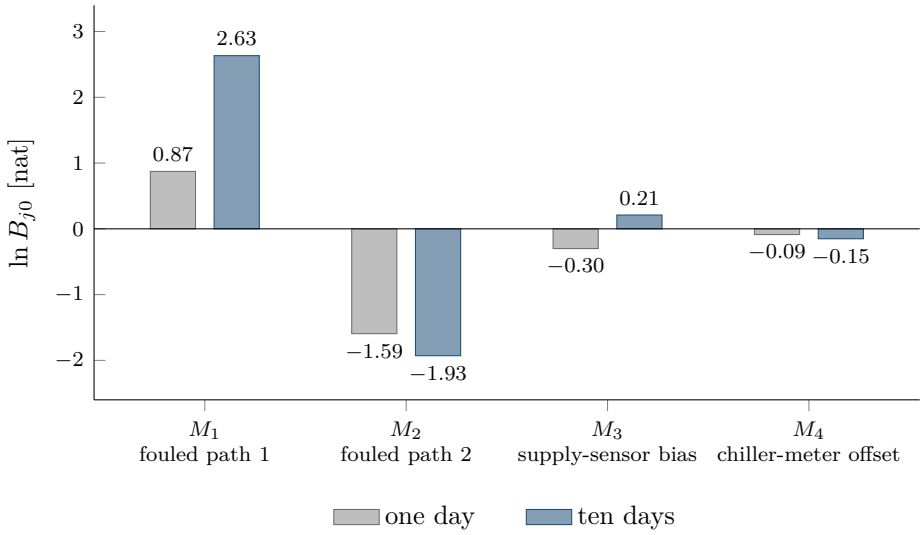


Figure 23.3: Log Bayes factors of the four drawings of the worked subset against the healthy model, Table 23.6, for the one-day and the ten-day records of period 2. Positive values favor the drawing.

Table 23.7: The fits of M_0, M_1 and M_3 to the ten-day record. Fitted values with standardized residuals in parentheses. The return temperature $T_r = T_s + (Q_1 + Q_2)/(\dot{m}c_p)$, with $\dot{m}c_p = 10\,\text{kW/K}$ pinned (Chapter 21), has no sensor; its band is $\pm\sqrt{\mathbf{g}_r^{\mathsf{T}}\boldsymbol{\Sigma}_M\mathbf{g}_r}$.

Row or quantity	Reading or prior	M_0 healthy	M_1 fouled path 1	M_3 supply bias
hall-1 sensor [°C]	19.80	19.705 (−0.60)	19.799 (−0.01)	19.730 (−0.45)
hall-2 sensor [°C]	17.90	17.903 (+0.02)	17.903 (+0.02)	17.940 (+0.25)
supply sensor [°C]	12.20	12.231 (+0.33)	12.197 (−0.03)	12.202 (+0.02)
chiller meter [kW]	53.0	52.99 (−0.02)	52.99 (−0.02)	52.86 (−0.21)
hall-1 load meter [kW]	29.5	29.72 (+0.47)	29.51 (+0.02)	29.72 (+0.46)
prior of g_1, as G_1 [kW/K]	5.49	3.976 (−3.02)	3.882 (−0.35)	4.427 (−2.02)
prior of g_2, as G_2 [kW/K]	4.06	4.103 (+0.09)	4.114 (+0.11)	4.700 (+1.08)
set point of T_s [°C]	12.0	12.231 (+0.23)	12.197 (+0.20)	13.016 (+1.02)
bias prior [K]	0	–	–	−0.814 (−0.41)
$2\mathcal{L}$ [–]	–	9.89	0.17	6.96
true supply T_s [°C]	–	12.23	12.20	13.02
return T_r [°C]	no sensor	17.53 ± 0.11	17.50 ± 0.11	18.30 ± 0.44

The ranking over the full family. Table 23.8 ranks the nine drawings of Table 23.3, the answer to the isolation question. Each line is one drawing, with its freed quantity's estimate and posterior spread where it has one, $2\mathcal{L}$, $\ln B_{j0}$ and the posterior probability at equal prior odds over the nine.

Table 23.8: The ranking over the full family, at equal prior odds over nine drawings. Estimates of the freed quantity ± its posterior spread $\sqrt{\Sigma_{66}}$: the temperature biases in K, the meter biases in kW, and for M_8 the fitted supply temperature in °C with $\sqrt{\Sigma_{T_sT_s}}$.

	One day				Ten days			
Drawing	Estimate	$2\mathcal{L}$	$\ln B_{j0}$	Prob.	Estimate	$2\mathcal{L}$	$\ln B_{j0}$	Prob.
M_0 healthy	–	6.04	0	0.079	–	9.89	0	0.020
M_1 fouled path 1	–	0.16	0.87	0.189	–	0.17	2.63	0.276
M_2 fouled path 2	–	6.04	−1.59	0.016	–	9.88	−1.93	0.003
M_3 supply bias	-0.77 ± 0.59	4.38	−0.30	0.058	-0.81 ± 0.45	6.96	0.21	0.024
M_4 chiller offset	0.01 ± 1.83	6.04	−0.09	0.072	-0.08 ± 1.72	9.88	−0.15	0.017
M_5 hall-1 sensor bias	1.89 ± 0.80	1.10	1.66	0.415	2.03 ± 0.60	1.19	3.42	0.603
M_6 hall-2 sensor bias	0.02 ± 0.98	6.04	−0.71	0.039	-0.06 ± 0.77	9.88	−0.97	0.008
M_7 load-meter bias	-1.55 ± 1.70	5.20	0.25	0.102	-2.14 ± 1.52	7.79	0.77	0.043
M_8 set point released	12.42 ± 0.27	5.90	−0.99	0.029	12.23 ± 0.09	9.84	−1.07	0.007

The hall-1 sensor bias leads the full family, with $\ln B = 3.42$ against the fouled path's 2.63, and takes 0.60 of the probability to the path's 0.28; the two together hold 0.88. Under M_5 the conductance is 5.32 against the healthy 5.49 kW/K, and every other fitted entry matches M_1's of Table 23.7 within 0.02 kW and 0.005 K, the shared footprint of Section 23.3.2. The two misfits are 0.17 and 1.19; the bias moves 2.03 K under a prior of spread 2 K, the fouled path 0.35 in the log-conductance under a prior of spread 1. Their ratio, $e^{3.42-2.63} = 2.2$, is the prior odds of a fouled coil over a 2 K sensor fault above which the path ranks first. At ten days the hall-2 bias gives $\ln B = -0.97$ and the load-meter bias +0.77, with $\hat{b}_{Q_1} = -2.14 \pm 1.52$ kW; the released set point gives −0.99 at one day and −1.07 at ten. Beside the leader's reported $\ln B = 3.416$ at ten days, importance sampling gives 3.427, and at one day 1.649 against 1.660; the Savage–Dickey identity with the Laplace marginal gives 4.562 and 1.850, off by 1.1 nats at ten days because zero lies $2.03/0.60 = 3.4$ standard deviations from the center of that marginal. The order of the two leaders is the same by the sampled evidences.

23.3.3 The size of the fault

Engineering question. How large is the fault, and how well is it known? Under each leading drawing the operators ask for the size of the change it represents, with its band.

General form. An estimation question asks for the posterior distribution of a quantity,

$$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) \le f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

equation (2.9), summarized by a location and the central 95 % interval of equation (2.11).

Specialized form. The question is equations (2.9) and (2.11) with $f(\mathbf{z}) = g_1$ under M_1 and $f(\mathbf{z}) = b_{T_1}$ under M_5, and $p(\mathbf{z} \mid \mathbf{y})$ the drawing's posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}_{M_1}, \mathbf{\Sigma}_{M_1})$ or $\mathcal{N}(\hat{\mathbf{z}}_{M_5}, \mathbf{\Sigma}_{M_5})$. Two choices enter here. The size is read within each leading drawing, as if that drawing were true: this is what gives a fault a size at all, and it makes the size conditional on the ranking, so that two leading drawings report two sizes of two different faults. The marginal of the freed entry is the Laplace normal of the drawing's fit: this gives the

band in closed form, and it would mislead if that marginal were skewed, which the agreement of the Laplace and sampled evidences bears on only indirectly (Section 23.6.6).

Under M_1 the quantity is an entry, $f(\mathbf{z}) = g_1$, read off as $\hat{g}_1$ with $\sqrt{\Sigma_{g_1 g_1}}$ from $\mathbf{\Sigma}_{M_1}$; the conductance with its 95 % interval is $\hat{G}_1 = \mathrm{e}^{\hat{g}_1}$ and $\exp(\hat{g}_1 \pm 1.960\sqrt{\Sigma_{g_1 g_1}})$, an interval for G_1 because the exponential is increasing and the quantiles of an increasing function are the images of the quantiles. Under M_5 the quantity is the sixth entry, $f(\mathbf{z}) = b_{T_1}$, read off as $\hat{b}_{T_1} \pm \sqrt{\Sigma_{66}}$ from $\mathbf{\Sigma}_{M_5}$. How firmly the conductance fell between the periods is a different statement: M_1 releases g_1 to the spread σ_w and keeps almost nothing of what period 1 knew of it, so that statement belongs to the comparison of the two fits, Section 23.3.4.

Building blocks. $\hat{g}_1$ and $\mathbf{\Sigma}_{M_1}$; $\hat{b}_{T_1}$ and $\mathbf{\Sigma}_{M_5}$ (Section 23.2.1); the normal quantile 1.960.

The size of the fault. Each of the two leading drawings has its own fault size. Table 23.9 gives the fouled path's. Each line is one record length; the columns give the fitted hall-1 conductance, the posterior spread of its logarithm, the 95 % interval of the conductance and its ratio to the period-1 value 5.490 kW/K.

Table 23.9: The size of the fault on the hall-1 path under M_1. Intervals are $\exp(\hat{g}_1 \pm 1.960\sqrt{\Sigma_{g_1 g_1}})$. The plant that generated period 2 had $G_1 = 4.0$ kW/K, 0.727 of its period-1 value 5.5.

Record	$\hat{G}_1$ [kW/K]	$\mathrm{sd}(g_1 \mid \mathbf{y}_2, M_1)$ [–]	95 % interval [kW/K]	$\hat{G}_1/5.490$ [–]
one day	3.886	0.091	3.25 to 4.64	0.708
ten days	3.882	0.029	3.67 to 4.11	0.707

Under the fouled path the hall-1 conductance has fallen from 5.49 to 3.88 kW/K, a loss of 29 %. With ten days, $\hat{g}_1 = \ln 3.882 = 1.3563$ and $1.960 \times 0.0291 = 0.0570$, so the interval is $\mathrm{e}^{1.2993}$ to $\mathrm{e}^{1.4134}$, 3.67 to 4.11 kW/K, which contains the generating value of 4.0; one day gives the same point with an interval three times as wide. The fault size is known to about 3 % after ten days. Under the hall-1 sensor bias the fault size is the bias, $\hat{b}_{T_1} = 2.03 \pm 0.60$ K at ten days (1.89 ± 0.80 at one day), with the conductance near its healthy value. Two other fits give values of G_1 that are not answers: the healthy model puts G_1 at 3.98 kW/K, a compromise with a prior it misses by three of that prior's spreads, and the supply-sensor drawing at 4.43, which belongs to a different explanation.

23.3.4 The change between the two periods

Engineering question. Which parameter moved between two periods? The question asks for the entry of the state that changed, without naming a fault and without a list of candidates.

General form. The change is a derived quantity of the states of the two periods, and the question is an estimation of it, the form (2.9) read entry by entry and, for the whole change at once, as the posterior probability of a credible region, equation (2.11) in several dimensions.

Specialized form. The question is equations (2.9) and (2.11) with the state the stacked pair $\mathbf{x} = (\mathbf{z}_1, \mathbf{z}_2)$ below, the derived quantity $f(\mathbf{x}) = \mathbf{\Delta} = \mathbf{z}_2 - \mathbf{z}_1$, entry by entry and as a whole, and $p(\mathbf{x} \mid \mathbf{y}_1, \mathbf{y}_2)$ the product of the two handbook fits of step 3, the Laplace normals $\mathcal{N}(\hat{\mathbf{z}}_p, \mathbf{\Sigma}_p)$ of the two periods. Three choices enter here. Each period has its own state and its own copy of the handbook prior of Table 17.2, independent between the periods: this is what makes the joint posterior factorize, and it overstates the spread of the difference if the two periods are one plant whose state was drawn once from the handbook, which Exercise 23.3 derives and Section 23.4.3 evaluates. The period-2 fit uses the handbook priors, not the healthy model:

the healthy model of isolation would count period 1 twice here. And each fit is its Laplace normal: this makes the change Gaussian and its joint statistic chi-square, which simulated pairs of records check (Section 23.4.3).

Stack the two states, $\mathbf{x} = (\mathbf{z}_1, \mathbf{z}_2)$, ten entries. The readings of the two periods are independent given the states, and Bayes' theorem for the pair is

$$p(\mathbf{z}_1, \mathbf{z}_2 \mid \mathbf{y}_1, \mathbf{y}_2) \propto p(\mathbf{y}_1 \mid \mathbf{z}_1)\, p(\mathbf{y}_2 \mid \mathbf{z}_2)\, p_{\mathrm{h}}(\mathbf{z}_1)\, p_{\mathrm{h}}(\mathbf{z}_2) = \left[p(\mathbf{y}_1 \mid \mathbf{z}_1)\, p_{\mathrm{h}}(\mathbf{z}_1)\right]\left[p(\mathbf{y}_2 \mid \mathbf{z}_2)\, p_{\mathrm{h}}(\mathbf{z}_2)\right],$$

with p_{h} the handbook prior. The rows of period 1 touch only the first five entries and those of period 2 only the last five, so the stacked Jacobian is block diagonal, its information matrix is $\mathrm{diag}(\mathbf{\Lambda}_1, \mathbf{\Lambda}_2)$, and the Laplace covariance of the pair is $\mathrm{diag}(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2)$: the two fits are independent given both records. The change $\mathbf{\Delta} = \mathbf{z}_2 - \mathbf{z}_1 = \mathbf{A}\mathbf{x}$, with $\mathbf{A} = [\,-\mathbf{I}_5 \;\; \mathbf{I}_5\,]$, is linear in the stacked state, its gradient matrix is $\mathbf{A}$, and the band of a derived quantity, $\mathbf{A}\,\mathrm{diag}(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2)\mathbf{A}^{\mathsf{T}}$, gives

$$\mathbf{\Delta} \mid \mathbf{y}_1, \mathbf{y}_2 \approx \mathcal{N}(\hat{\mathbf{\Delta}},\ \mathbf{\Sigma}_1 + \mathbf{\Sigma}_2), \qquad \hat{\mathbf{\Delta}} = \hat{\mathbf{z}}_2 - \hat{\mathbf{z}}_1, \qquad Z_i = \frac{\hat{\Delta}_i}{\sqrt{(\mathbf{\Sigma}_1 + \mathbf{\Sigma}_2)_{ii}}}. \tag{23.14}$$

Entry by entry, Z_i is the standardized change of entry i, and the posterior probability that the entry moved the other way is $\Phi(-|Z_i|)$.

The joint statistic. Five standardized changes are five looks, and the largest of them is inflated by being the largest; a Bonferroni bound, the number of looks times the two-sided tail of the largest, corrects for that and ignores that the looks are correlated through the covariances. The joint statistic is the squared Mahalanobis distance of no change from the estimated change,

$$\chi^2_\Delta = \hat{\mathbf{\Delta}}^{\mathsf{T}} (\mathbf{\Sigma}_1 + \mathbf{\Sigma}_2)^{-1} \hat{\mathbf{\Delta}}. \tag{23.15}$$

Write $\mathbf{C} = \mathbf{\Sigma}_1 + \mathbf{\Sigma}_2$ and $\mathbf{u} = \mathbf{C}^{-1/2}(\mathbf{\Delta} - \hat{\mathbf{\Delta}})$; under equation (23.14) $\mathbf{u}$ has independent standard normal entries, so $(\mathbf{\Delta} - \hat{\mathbf{\Delta}})^{\mathsf{T}}\mathbf{C}^{-1}(\mathbf{\Delta} - \hat{\mathbf{\Delta}}) = \mathbf{u}^{\mathsf{T}}\mathbf{u} \sim \chi^2_5$. The point $\mathbf{\Delta} = \mathbf{0}$ lies on the ellipsoid on which this quadratic form equals χ^2_Δ, and the posterior probability inside that ellipsoid is $\Pr(\chi^2_5 < \chi^2_\Delta)$. The tail $\Pr(\chi^2_5 > \chi^2_\Delta)$ is therefore the posterior probability outside the smallest credible ellipsoid that contains no change; a small tail says that no change lies outside nearly all the posterior mass of the change. Read as a test of no change, the same number is approximately the probability that two periods of an unchanged plant would differ this much.

The reach and limits of attribution. Attribution gives the where: the entry that moved, its sign and size, with a band, and no list of candidates. It lacks a cause. A fouled hall-1 path and a hall-1 sensor that reads high both move the fitted g_1, through the one row of equation (23.4) that they share, and attribution reports the same change for both. It can only report a change in an entry of the state, so a fault outside the state, such as a sensor bias, appears as a change of whichever entry the biased reading moves. It also lacks a prior on the change beyond the one implied by two independent handbook priors, so it locates a change without saying how probable a change was; the comparison of drawings does both, at the price of a list.

Building blocks. The handbook fits of the two periods, $\hat{\mathbf{z}}_p$ and $\mathbf{\Sigma}_p$ (Section 23.2.1); $\mathbf{C}$ and its inverse; Φ and the tail of χ^2_5.

Drift attribution entry by entry. Table 23.10 lists the comparison of the two handbook fits by equations (23.14) and (23.15) for the ten-day record. Each line is one entry of the state, with its period-1 and period-2 estimates and spreads, the change and its spread, the standardized change Z_i, and the same Z_i for the one-day record.

Table 23.10: Drift attribution between the two periods, both fitted with the handbook priors. Period 2 at ten days unless marked. Spreads are $\sqrt{\Sigma_{ii}}$ of each fit; the spread of the change is $\sqrt{(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)_{ii}}$. Units: dimensionless for g_i, kW for the loads, K for T_s.

Entry	Period 1	Period 2	Change	Z_i, ten days	Z_i, one day
g_1	1.7030 ± 0.1067	1.3581 ± 0.0290	-0.345 ± 0.111	-3.12	-2.38
g_2	1.3999 ± 0.1370	1.4205 ± 0.0461	$+0.021 \pm 0.145$	$+0.14$	$+0.26$
Q_1	30.69 ± 1.45	29.50 ± 0.47	-1.19 ± 1.53	-0.78	-0.56
Q_2	22.23 ± 2.35	23.53 ± 0.79	$+1.30 \pm 2.47$	$+0.53$	$+0.46$
T_s	12.111 ± 0.280	12.205 ± 0.094	$+0.093 \pm 0.296$	$+0.32$	$+0.32$
joint χ^2_Δ on 5 degrees of freedom				14.47	8.85
tail $\Pr(\chi^2_5 > \chi^2_\Delta)$				0.013	0.115

The hall-1 log-conductance changed by -0.345 ± 0.111, from the spreads $\sqrt{0.1067^2 + 0.0290^2} = 0.111$, so $Z_1 = -0.345/0.111 = -3.12$, a fall of the conductance to $\mathrm{e}^{-0.345} = 0.708$ of its period-1 value. The posterior probability that g_1 rose instead is $\Phi(-3.12) = 0.0009$. The other four entries moved by less than one of their spreads. Jointly, $\chi^2_\Delta = 14.47$ on five degrees of freedom, and no change lies outside the credible ellipsoid of probability $1 - 0.013 = 0.987$. A Bonferroni bound on the largest of the five, five times the two-sided tail of 3.12, is $5 \times 0.0018 = 0.009$. With one day, $Z_1 = -2.38$ is still the largest, the joint $\chi^2_\Delta = 8.85$ has tail 0.115 and the Bonferroni bound is 0.088.

Figure 23.4 draws the standardized changes. The vertical axis lists the five entries and the horizontal axis is Z_i; each open circle is the one-day record and each filled circle the ten-day record, with a bar of ± 2, the width of a 95 % band in these units, and the vertical line marks no change. Only the bars of g_1 exclude zero, at both record lengths; the ten-day bar of g_1 lies wholly beyond -1.

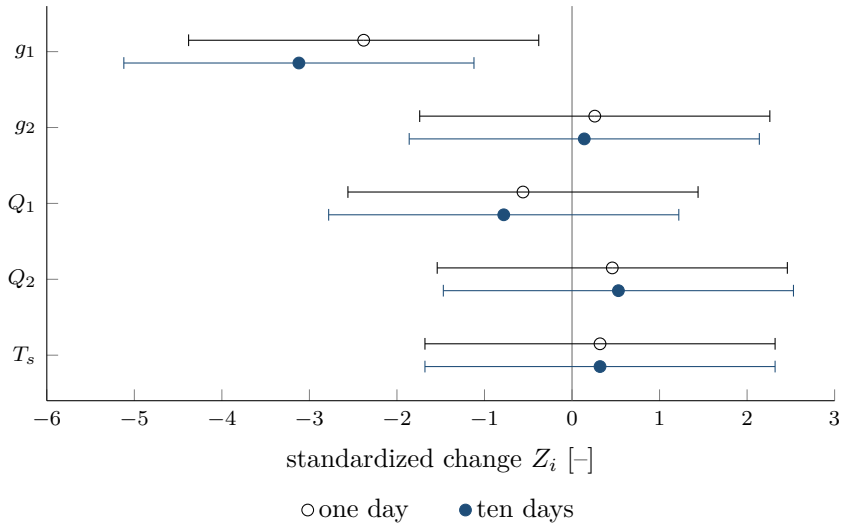


Figure 23.4: Drift attribution: the standardized change Z_i of each entry between the two periods, equation (23.14), with bars of ± 2. Joint statistic $\chi^2_\Delta = 8.85$ (tail 0.115) for one day and 14.47 (tail 0.013) for ten days, Table 23.10.

23.3.5 The length of the record

Engineering question. Could one day of readings have told? The operators ask what the same analysis would have said after the first day, and what a second record longer than ten

days would add.

General form. A design question values a record before it is taken by what it would do to the posterior,

$$U(a) = \mathbb{E}_{\tilde{y}_a|\mathbf{y}}\Big[\, u\big(p(\mathbf{z} \mid \mathbf{y}, \tilde{y}_a)\big)\Big],$$

equation (2.16), here with the candidate a a record of N days and u the answers to the first four questions; what a record is worth in units of the decision it informs is left to a later chapter on the value of information.

Specialized form. The question is equation (2.16) with the action a a period-2 record of N days, u the answers to the first four questions, and $p(\mathbf{z} \mid \mathbf{y}, \tilde{y}_a)$ the posterior of step 3 of each drawing and of the period-2 handbook fit, refitted with the reading rows at the standard deviations $\sigma_k/\sqrt{N}$. One choice enters here. The outcome of a record of N days is held at the observed averages instead of averaged over its predictive: this separates what the record's precision buys from what its noise does, and it overstates the scatter of a long record, since a real average of many days would sit closer to the plant's values than the period-2 values do (Section 23.6.6). Each answer is then a function of N through the standard deviations $\sigma_k/\sqrt{N}$ of the reading rows alone: at each N every drawing and the period-2 handbook fit are refitted on the same rows with those standard deviations, and $\hat{\mathbf{z}}_M$, $\mathbf{\Sigma}_M$ and $\ln\det\mathbf{\Lambda}_M$ of the new fit enter the same formulas, equations (23.10), (23.11) and (23.14), with $\hat{\mathbf{z}}_1$ and $\mathbf{\Sigma}_1$ of period 1 held fixed.

The misfit under averaging. At a fixed state the normalized reading row of a record of N days is $r_k = (h_k - \bar{y}_k)\sqrt{N}/\sigma_k$, with $h_k(\mathbf{z})$ the value the model predicts for the reading and $\bar{y}_k$ the average, so its square is $N(h_k - \bar{y}_k)^2/\sigma_k^2$. If the model is right at the true state, $\bar{y}_k - h_k(\mathbf{z}^{\text{true}})$ has variance σ_k^2/N and $\mathbb{E}[r_k^2] = 1$ for every N: the reference does not move. If the plant has changed so that the healthy model's prediction misses the reading by a fixed amount δ_k, then $r_k^2 \to N\delta_k^2/\sigma_k^2$, which grows in proportion to N. A persistent discrepancy therefore gains weight linearly with the length of the record against a reference that stays where it is.

The limit of detection. The minimized statistic grows more slowly, because the prior rows do not scale with N. As N grows the reading rows stiffen, the fit moves the state toward what the readings alone imply, and the misfit migrates into the prior rows. The two-hall plant has five reading rows for five entries and a design of full rank (Section 19.3.2), so in the limit the readings fix the state at the readings-only solution $\mathbf{z}_{\text{r}}$, the square system of Section 18.4.6, and

$$\lim_{N\to\infty} 2\mathcal{L}_{M_0}(\hat{\mathbf{z}}_{M_0}) = \sum_{k\,\in\,\text{prior rows}} r_k(\mathbf{z}_{\text{r}})^2, \tag{23.16}$$

and the detection statistic saturates at the conflict between the readings-only state and the healthy model, measured in the healthy model's own spreads. A longer second period cannot detect a change that the healthy model is too vague to resolve; a more precise healthy model, from a longer first period, can.

The limits of isolation and attribution. The same reasoning bounds the other answers. In equation (23.13) the variance v_{L} of the readings' estimate of g_1 falls as $1/N$, and $\ln B_{10}$ tends to the finite limit

$$\ln B_{10} \to \ln\frac{\sigma_{\text{h}1}}{\sigma_{\text{w}}} + \tfrac{1}{2}\big(\hat{g}_1^{\text{L}} - g_1^{\text{h}}\big)^2\big(\sigma_{\text{h}1}^{-2} - \sigma_{\text{w}}^{-2}\big),$$

set by the healthy model's spread, as the detection statistic was. In equation (23.14) the period-2 entries of $\mathbf{\Sigma}_2$ fall as $1/N$ wherever the readings determine the entry, so $Z_1 \to (g_{1,\text{r}} - \hat{z}_{1,1})/\sqrt{(\mathbf{\Sigma}_1)_{11}}$,

with $g_{1,\mathrm{r}}$ the log-conductance of the readings-only state and $\hat{z}_{1,1}$ the first entry of $\hat{\mathbf{z}}_1$: the change of g_1 from period 1 to the readings-only state in units of period 1's spread. And the evidence ratio of two drawings with one footprint, M_5 against M_1, is decided by their prior rows (Section 23.3.2), so it should hardly move with N.

Building blocks. The fits of the drawings and of the two periods at $N = 1$, 3, 10, 30, 100 and 1000 days with the period-2 averages; the readings-only state $\mathbf{z}_{\mathrm{r}}$; the prior rows at it.

The value of a record length. Table 23.11 repeats detection, isolation and attribution for records of $N = 1$ to 1000 days, with the same averaged values. Each line is one record length; the columns give the healthy model's misfit and its tail, the log Bayes factors of the fouled path and the supply bias, the probability of the fouled path in the worked subset, the log Bayes factor of the hall-1 sensor bias against the fouled path, and the standardized change of g_1 with the joint statistic and its tail.

Table 23.11: Detection, isolation and attribution against the length N of the period-2 record in days, with the period-1 record fixed at one snapshot. $\Pr(M_1)$ is over the worked subset; $\ln B_{51}$ is the hall-1 sensor bias against the fouled path, which fixes their ratio in the full family.

N	$2\mathcal{L}_{M_0}$	tail	$\ln B_{10}$	$\ln B_{30}$	$\Pr(M_1)$	$\ln B_{51}$	Z_1	χ^2_Δ	tail of χ^2_Δ
1	6.04	0.110	0.87	−0.30	0.456	0.79	−2.38	8.85	0.115
3	8.50	0.037	1.98	+0.05	0.701	0.78	−2.87	12.44	0.029
10	9.89	0.020	2.63	+0.21	0.811	0.78	−3.12	14.47	0.013
30	10.36	0.016	2.86	+0.25	0.842	0.78	−3.20	15.18	0.0096
100	10.54	0.014	2.95	+0.27	0.852	0.78	−3.24	15.44	0.0086
1000	10.61	0.014	2.98	+0.27	0.856	0.78	−3.25	15.55	0.0083

At one day the tail is 0.11, the fouled path has probability 0.46 in the subset and the change of g_1 is 2.4 spreads with a joint tail of 0.12. Three days give a tail of 0.037 and the fouled path 0.70 in the subset; ten days give 0.81 and a joint tail of 0.013. From thirty to 1000 days $\ln B_{10}$ gains 0.12 nats, Z_1 moves by 0.05, and the two tails fall from 0.016 to 0.014 and from 0.0096 to 0.0083. The readings-only state of period 2 is $T_s = 12.20°\mathrm{C}$, $Q_1 = 29.5\,\mathrm{kW}$, $Q_2 = 53.0 - 29.5 = 23.5\,\mathrm{kW}$, $G_1 = 29.5/(19.80 - 12.20) = 3.882$ and $G_2 = 23.5/(17.90 - 12.20) = 4.123\,\mathrm{kW/K}$, and equation (23.16) gives

$$\left(\frac{\ln 3.882 - 1.7030}{0.1067}\right)^2 + \left(\frac{\ln 4.123 - 1.3999}{0.1370}\right)^2 + \left(\frac{12.20 - 12}{1}\right)^2 = 3.250^2 + 0.122^2 + 0.20^2 = 10.62,$$

with tail 0.014. The isolation limit is $\ln B_{10} \to \ln(0.1067) + \frac{1}{2} \times 0.3468^2 \times (87.86 - 1) = -2.238 + 5.222 = 2.98$, and the attribution limit $Z_1 \to (1.3562 - 1.7030)/0.1067 = -3.25$, all three set by the period-1 spread 0.107 of g_1. The column of $\ln B_{51}$ stays at 0.78–0.79 from one day to 1000.

23.4 The checks (step 5)

What can fail is the Laplace route where it stands in for the evidence integral, the widths stated for the released priors, the Gaussian reading of the change between the periods, and the exact balances that the reduced state takes as a limit.

23.4.1 The Laplace evidences against importance sampling

Table 23.12 compares three routes to each log Bayes factor. The first column is the Laplace value of Tables 23.6 and 23.8; the second an importance-sampling estimate of each evidence

integral, with 200 000 draws from a Student-t density with five degrees of freedom centered at the MAP state and scaled to $1.5^2\boldsymbol{\Sigma}_M$, whose effective sample sizes were 78 000 to 87 000, so that the Monte Carlo standard error of each log evidence, $(1/\mathrm{ESS} - 1/N)^{1/2}$ with ESS the effective sample size and N the number of draws, is 0.003 nats and that of a log Bayes factor 0.004; the third the one-entry identity of Section 23.3.2, the Savage–Dickey ratio for an added entry and the posterior expectation of the prior ratio for a widened prior, each evaluated with the Laplace marginal of that entry.

Table 23.12: Checks of the log Bayes factors against M_0: Laplace, importance sampling (IS) and the one-entry identities of Section 23.3.2 evaluated with the Laplace marginal (prior-ratio expectation for M_1, Savage–Dickey for the others).

	One day			Ten days		
Drawing	Laplace	IS	Identity	Laplace	IS	Identity
M_1 fouled path 1	0.873	0.852	1.059	2.633	2.634	2.658
M_3 supply bias	−0.300	−0.302	−0.362	0.210	0.227	0.138
M_4 chiller offset	−0.087	−0.091	−0.087	−0.151	−0.150	−0.151
M_5 hall-1 sensor bias	1.660	1.649	1.850	3.416	3.427	4.562

The Laplace values agree with importance sampling to 0.021 nats or better in every case, and the ranking is the same by both; the larger differences exceed the Monte Carlo error and are the error of the Laplace approximation itself. The one-entry identities are exact, but evaluated with a Gaussian marginal they fail where that marginal is read far from its center. Savage–Dickey reads the posterior density of the bias at zero: for the chiller offset zero is the center and the identity is exact; for the hall-1 bias at ten days zero lies $2.03/0.60 = 3.4$ standard deviations out, and the identity is off by 1.1 nats. The prior-ratio expectation for M_1 weights the region where the healthy prior overlaps the M_1 posterior, within one of that posterior's standard deviations of its center at ten days and 1.6 of them out at one day, where it errs by 0.19. This is the warning of Section 7.4.4: a Laplace normal read at a point away from its center is not the exact identity. The Laplace evidence of equation (23.10), which uses the curvature at the MAP state of each drawing, is the reliable one here, and it is the route reported.

The misfit of the leading drawings. Under M_1 at ten days, $2\mathcal{L}_{M_1} = 0.17$ on three degrees of freedom, tail 0.98, and no standardized residual exceeds 0.35 (the prior of g_1, on its widened spread); under M_5, $2\mathcal{L}_{M_5} = 1.19$, tail 0.76.

23.4.2 Sensitivity to the released priors

The spread $\sigma_\mathrm{w} = 1$ of the released priors is a statement, and the ranking must be read with it. Figure 23.5 plots $\ln B_{10}$ of the fouled path against σ_w, and Table 23.13 gives, for five spreads from 0.25 to 4, what the figure does not: the log Bayes factor of the other path drawing, the posterior probability of M_1 in the worked subset and its fitted $\hat{G}_1$, at both record lengths. Each line of the table is one spread.

In the figure the horizontal axis is σ_w on a logarithmic scale and the vertical axis $\ln B_{10}$; the gray curve is the one-day record and the blue curve the ten-day record, with the spreads 0.5, 1 and 2 marked. The dotted line marks the healthy spread $\sigma_{\mathrm{h}1} = 0.107$, at which M_1 is M_0 and both curves start from zero, and the dashed line marks the matched width $|\hat{g}_1^{\mathrm{L}} - g_1^{\mathrm{h}}| = 0.35$, at which equation (23.13) puts the maximum. Both curves rise to that maximum and then fall like $-\ln \sigma_\mathrm{w}$, the Occam factor taking over. At ten days the fouled path stays first in the subset over the whole range, its probability between 0.54 and 0.87; the size of the fault does not move at all. At one day it falls below the healthy model for σ_w beyond 2.5: a vague enough release

Table 23.13: The worked subset against the spread σ_w of the released log-conductance priors; $\ln B_{10}$ is plotted in Figure 23.5. M_3 and M_4 do not depend on the spread ($\ln B = -0.30$ and -0.09 at one day, $+0.21$ and -0.15 at ten days). $\hat{G}_1$ in kW/K.

	One day			Ten days		
σ_w [–]	$\ln B_{20}$	$\Pr(M_1)$	$\hat{G}_1$	$\ln B_{20}$	$\Pr(M_1)$	$\hat{G}_1$
0.25	−0.36	0.556	4.040	−0.56	0.862	3.899
0.5	−0.93	0.567	3.919	−1.24	0.873	3.886
1	−1.59	0.456	3.886	−1.93	0.811	3.882
2	−2.28	0.313	3.878	−2.62	0.697	3.881
4	−2.97	0.190	3.876	−3.31	0.540	3.881

of the prior makes the fault drawing lose to no fault, which is how the evidence penalizes a hypothesis that predicts everything.

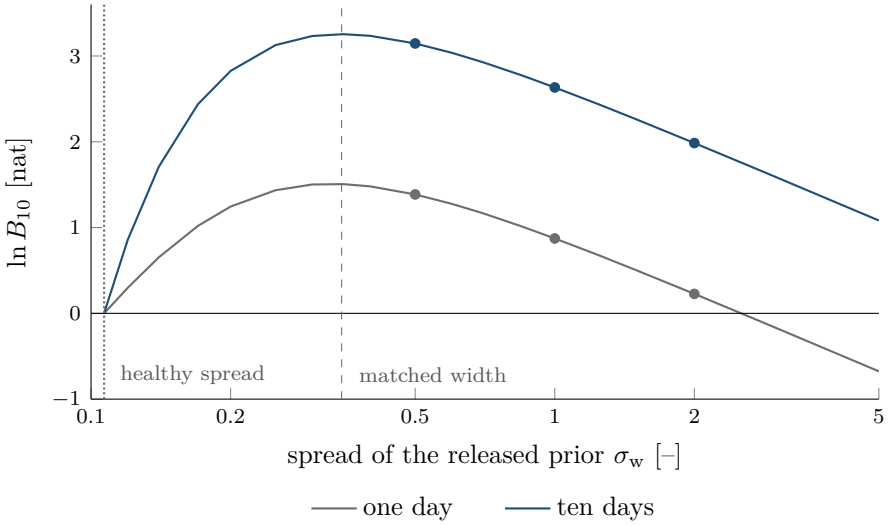


Figure 23.5: The log Bayes factor of the fouled hall-1 path against the healthy model as a function of the spread σ_w of its released prior, by the Laplace evidence of equation (23.12); marks at $\sigma_w = 0.5$, 1, 2 (Table 23.13). Dotted: the healthy spread 0.107; dashed: the matched width 0.35 of equation (23.13).

The leading pair of the full family moves with the stated widths and is not decided by them: with the bias prior at 1, 2 or 4 K the hall-1 sensor bias has $\ln B = 2.78$, 3.42 or 3.17 at ten days, and the fouled path's runs from 1.30 to 3.15 over the release widths of Table 23.13. The two ranges overlap, so a modest change of the stated widths reorders the pair; no stated width lets the record separate it. The scale of each release is a modeling statement and is reported with the ranking.

Common mistake

Widening a released prior to be safe. A vaguer release does not make a fault drawing more general at no cost: its Occam factor falls like $-\ln \sigma_w$, and at one day a spread of 4 ranks the fouled path below the healthy model. The release is a statement of how large a fault is plausible, and it is set from engineering knowledge of the fault, not as large as possible.

23.4.3 The covariance of the difference

Two thousand pairs of records were drawn from the period-1 plant, unchanged, one day for the first period and ten days for the second, and each pair was fitted with the handbook priors. The simulated standard deviations of the change were 0.105, 0.122, 1.50 kW, 2.36 kW and 0.282 K for the five entries, against $\sqrt{(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)_{ii}} = 0.111$, 0.145, 1.53, 2.47 and 0.296 from the fits. The posterior formula is slightly conservative for this purpose because the simulation holds the true state fixed. At a fixed state a fit with linear rows has sampling covariance $\boldsymbol{\Sigma}_p - \boldsymbol{\Sigma}_p \boldsymbol{\Lambda}_{\text{prior}} \boldsymbol{\Sigma}_p$ for period p, with $\boldsymbol{\Lambda}_{\text{prior}}$ the precision of the handbook prior: the shrinkage toward the prior takes out the prior's share. The sum over the two periods, evaluated at the two fits, gives 0.103, 0.130, 1.49, 2.35 and 0.279, close to the simulated values; the remaining differences, largest for g_2, come from evaluating the curvature at the fitted rather than the true state. The mean of χ^2_Δ over the pairs was 4.69 (standard error about 0.07) against the 5 of χ^2_5, and 1.05 % of the pairs (standard error 0.23 %) exceeded the observed 14.47, against the 1.29 % of the chi-square tail.

A separate question is a shared prior: treating the state as drawn once from the handbook for both periods adds the cross term of Exercise 23.3, which moves the hall-1 change to $Z_1 = -3.15$ and the joint statistic to $\chi^2_\Delta = 14.62$ (tail 0.012), a correction of one percent here. One systematic effect showed in the simulation: the mean change of g_2 was -0.035, a quarter of its spread, with no change in the plant. The two periods have unequal precision, so the handbook prior pulls the one-day fit harder than the ten-day fit, and an entry held partly by its prior appears to drift when the record length changes.

23.4.4 The exact balances as a limit

Every fit of this chapter runs on the reduced state, the limit $\sigma_p \to 0$ of the three balances. Table 23.14 repeats the leading answers on the full state of Chapter 17, the eight entries $(g_1, g_2, Q_1, Q_2, T_s, T_1, T_2, P)$ with the two hall balances and the header balance as physics rows of tolerance σ_p in kelvin and $\sigma_{p,P}$ in kilowatts of the same number, the evidence of each drawing carrying the physics rows' normalizers. Each line is one tolerance; the columns are the log Bayes factors of the fouled path, the hall-1 sensor bias and the supply-sensor bias at ten days, and the standardized change of g_1 with the joint statistic.

Table 23.14: The ten-day answers on the full state with the balances as physics rows of tolerance σ_p (σ_p in K for the halls, $\sigma_{p,P}$ in kW of the same number for the header), against the reduced state.

Tolerance	$\ln B_{10}$	$\ln B_{50}$	$\ln B_{30}$	Z_1	χ^2_Δ
$\sigma_p = 0.1$	2.5648	3.3477	0.1908	−3.062	13.68
$\sigma_p = 0.01$	2.6325	3.4152	0.2102	−3.119	14.46
$\sigma_p = 0.001$	2.6331	3.4159	0.2104	−3.119	14.47
reduced state	2.6331	3.4159	0.2104	−3.119	14.47

At $\sigma_p = 0.1$ a balance may fail by a tenth of a kelvin, and the fouled path loses 0.07 nats, because a hall that may miss its balance absorbs part of the hall-1 rise without a change of g_1. From $\sigma_p = 0.01$ every printed value agrees with the reduced state to within 0.001 nats in the evidences and 0.01 in the joint statistic, and at 0.001 to every printed digit: the normalizers of the physics rows cancel against the widths they give the posterior, as equation (4.19) states for one hall, and the reduced state is the limit, not a substitution.

The standard deviations σ_k of the readings have one more check that the record cannot supply. The factor F_k of equation (23.5) holds the scatter of the daily readings about their average, and that scatter, against σ_k, would test the stated standard deviations and the steady state of the period; the illustrative record gives the averages only, so the check is left to the site's daily logs (Section 23.6.6).

23.5 The report (step 6)

The first reading of period 2 does not survive as a work order. The record does show that the plant has changed, at the 2 % level with ten days, and that the change is in hall 1; it does not show that the coil fouled. A hall-1 sensor reading 2 K high writes the same record, and over the nine drawings that the graph generates it takes 0.60 of the probability against 0.28 for the fouled coil, a ratio of 2.2 that the two drawings' priors set and that stays the same at every length of record. The 0.81 that the fouled coil takes among the four drawings worked first is a share of a family that left the sensor out. The answers below are given in the units in which the questions were asked.

23.5.1 The answers and their bands

Table 23.15 gives the answers. Each row is a question of Table 23.2, in its order; the second column gives the answer from the posteriors of the state, the third the value beside it from importance sampling of the evidence integral where one exists, and the fourth the evidence and the band behind it.

Table 23.15: The answers to the questions of the chapter, for the ten-day record unless stated: answers from the posteriors of the state (Laplace evidences), with the importance-sampling value beside them where one exists.

Question	Answer (state route)	Exact	Evidence and band
Has the plant changed since period 1?	the healthy model is rejected at the 2 % level with ten days; not with one	–	$2\mathcal{L}_{M_0} = 9.89$ on 3 degrees of freedom, tail 0.020 (one day 6.04, tail 0.110); hall-1 reading 3.7 predictive standard deviations above the healthy prediction
Which component?	hall 1: its path 0.28 or its sensor 0.60, not separable by these readings ($\ln B = 2.633$ and 3.416)	$\ln B = 2.634$ and 3.427	full family of nine at equal prior odds; $\ln B_{51} = 0.78$–0.79 at every record length from 1 to 1000 days; the pair reorders within the release and bias widths examined; Table 23.8
How large is the fault, and how well is it known?	under the fouled path, $G_1 = 3.88\,\mathrm{kW/K}$, 95 % interval 3.67 to 4.11, against 5.49 in period 1, a loss of 29 %; under the sensor bias, $b_{T_1} = 2.03 \pm 0.60\,\mathrm{K}$	–	M_1 and M_5 fits, Table 23.9
Which parameter moved between two periods?	the hall-1 log-conductance, by -0.345 ± 0.111, $Z = -3.1$; no other entry by more than one spread	–	$\chi^2_\Delta = 14.47$ on 5 degrees of freedom, tail 0.013; $\Pr(g_1 \text{ rose}) = 0.0009$; Table 23.10
Could one day of readings have told?	it would have suggested, not shown: three days reject the healthy model at 5 %, thirty reach the limit set by the baseline; no length separates path from sensor	–	tail 0.11, $Z_1 = -2.4$ with joint tail 0.12 at one day; limits 10.62 (tail 0.014), 2.98 and -3.25; Table 23.11

The report line. Ten days of readings in period 2 reject the healthy model of period 1 at the 2 % level. Of the nine drawings that the plant graph generates, the fault is isolated to

hall 1: its heat-removal path, fouled to 3.88 kW/K (between 3.67 and 4.11 at 95 %, 29 % below its period-1 value), with probability 0.28, or its temperature sensor, reading 2.0 K high, with probability 0.60 at equal prior odds. These readings cannot separate the two at any record length; their evidence ratio, 2.2 in favor of the sensor at the stated priors, is set by those priors. A comparison of the fits of the two periods, made without any list of candidates, finds the same entry and the same change, 3.1 standard deviations from zero, and cannot separate them either.

23.5.2 Isolation against attribution

The reach and limits of the two methods are those derived in Sections 23.3.2 and 23.3.4. On this record the two methods agree on the entry (g_1) and on the size of the change (29 %), and both are blind in the same place: attribution reports a fall of g_1 whether the path fouled or the sensor drifted, and the ranking splits the two by their priors. They disagree in other records. Exercise 23.6 works a drift that is entirely a supply-sensor bias: attribution reports that the supply temperature fell, which it did not, while the family, which contains the bias, ranks the bias first. Attribution is the screening method, cheap across many periods and entries; isolation is the diagnosis, once the change has been located and the candidates can be written down.

> **Principle 23.1.** *Attribution says where the state moved; isolation says which listed cause moved it. Neither can separate two causes that change the same rows in the same way, and the redundancy of the design, not the length of the record, decides which causes those are.*

Common mistake

Reading the leading drawing's probability as the probability of the fault. The fouled path takes 0.81 in the worked subset, a family that leaves out the hall-1 sensor; in the full family the same record gives it 0.28. A posterior model probability is conditional on the family, and it is reported with the family's list, the priors of the freed quantities and the prior odds.

Common mistake

Reading the entry that moved as the cause. Attribution found that the hall-1 log-conductance fell by 29 %; it would have found the same fall for a hall-1 sensor reading 2 K high, and the coil should not be cleaned on that evidence alone. The entry that moved is where to look, and the cause is a comparison of drawings, or a new row.

23.6 The next measurement (step 7)

The report leaves open the question the work order turns on: the hall-1 path or the hall-1 sensor, at $\ln B_{51} = 0.78$. Within the worked subset a second pair is open as well, the fouled path against the supply-sensor bias, at $\ln B_{13} = 2.633 - 0.210 = 2.42$, and it is kept because it shows the use of a temperature that no sensor reads.

23.6.1 The candidate readings

Three candidates are weighed, each read for one day or averaged over ten days. A second temperature sensor in hall 1, of $\sigma_a = 0.5\,\mathrm{K}$, reads the true T_1: its row is the hall-1 row of equation (23.6) without any bias, so under M_5 it does not share the first sensor's bias and under M_1 it repeats the first sensor's reading. A sensor on the return header, of $\sigma_a = 0.3\,\mathrm{K}$, reads $T_r = T_s + (Q_1 + Q_2)/(\dot{m}c_p)$ with the gradient $\mathbf{g}_r = (0,\ 0,\ 0.1\,\mathrm{K/kW},\ 0.1\,\mathrm{K/kW},\ 1)$ of Section 23.2.1. A second supply sensor, of $\sigma_a = 0.3\,\mathrm{K}$, reads the true T_s, with the gradient $(0,\ 0,\ 0,\ 0,\ 1)$. Averaged over ten days each standard deviation is divided by $\sqrt{10}$.

23.6.2 The expected change of an evidence ratio

The open question is a ranking of two drawings, not a width, so a candidate is valued by the evidence it is expected to add rather than by the width the preposterior covariance $\mathbf{\Sigma}_a$ of equation (2.17) would remove from an entry; the design form (2.16) is taken with u the change of the log evidence ratio. The evidence of drawing M for the record extended by a candidate reading y_a factorizes by the product rule, $p(\mathbf{y}_2, y_a \mid M) = p(\mathbf{y}_2 \mid M)\, p(y_a \mid \mathbf{y}_2, M)$, so for two drawings M_j and M_l

$$\ln B_{jl}(\mathbf{y}_2, y_a) = \ln B_{jl}(\mathbf{y}_2) + \ln \frac{p(y_a \mid \mathbf{y}_2, M_j)}{p(y_a \mid \mathbf{y}_2, M_l)},$$

and the change is the log ratio of the two drawings' predictive densities at the reading. If M_j is true, y_a is drawn from its predictive, and the expected change is the Kullback–Leibler divergence of the M_l predictive from the M_j predictive, which is never negative: on average a new reading moves the evidence toward the true drawing, although a particular reading can move it the other way. Each predictive is Gaussian by the device of equation (21.21), $\mathcal{N}(\mu_j,\ \mathrm{sd}(y_a \mid \mathbf{y}_2, M_j)^2)$ with

$$\mu_j = f_a(\hat{\mathbf{z}}_{M_j}), \qquad \mathrm{sd}(y_a \mid \mathbf{y}_2, M_j)^2 = \mathbf{g}_a^{\mathsf{T}} \mathbf{\Sigma}_{M_j} \mathbf{g}_a + \sigma_a^2,$$

where f_a is the quantity the reading measures, $\mathbf{g}_a$ its gradient at $\hat{\mathbf{z}}_{M_j}$, $\mathbf{\Sigma}_{M_j}$ the drawing's covariance and σ_a the reading's standard deviation; the predictive spread is the reading's own error added to the band of the drawing. The expectation of the log ratio of two Gaussian densities is

$$\mathbb{E}\big[\Delta \ln B_{jl} \mid M_j\big] = \tfrac{1}{2}\left[\ln \frac{\mathrm{sd}(y_a \mid \mathbf{y}_2, M_l)^2}{\mathrm{sd}(y_a \mid \mathbf{y}_2, M_j)^2} + \frac{\mathrm{sd}(y_a \mid \mathbf{y}_2, M_j)^2 + (\mu_j - \mu_l)^2}{\mathrm{sd}(y_a \mid \mathbf{y}_2, M_l)^2} - 1\right], \tag{23.17}$$

and the same with the indices exchanged if M_l is true. The candidate reading that separates two drawings is the one whose predicted values differ most in units of their spreads; equation (23.17) puts the difference in nats of evidence. It is the preposterior of Chapter 2 applied to an evidence ratio instead of a posterior spread, and it needs the same two objects as $\mathbf{\Sigma}_a$: the drawing's covariance and the candidate's gradient.

23.6.3 The ranking of the readings

Table 23.16 evaluates the candidates for each pair at the ten-day state. Each line is one candidate reading with its standard deviation, the pair (M_j, M_l) it addresses, the two drawings' predictive means μ and spreads $\mathrm{sd}(y_a \mid \mathbf{y}_2, M)$ of that reading, their separation $|\mu_j - \mu_l| / [\mathrm{sd}(y_a \mid \mathbf{y}_2, M_j)^2 + \mathrm{sd}(y_a \mid \mathbf{y}_2, M_l)^2]^{1/2}$, and the expected change of the log Bayes factor by equation (23.17) under each drawing of the pair.

The leading pair is separated by a second view of hall 1: a second temperature sensor reads 19.80°C if the path is fouled and 17.76°C if the first sensor is 2 K high. Ten days of it are expected to add 6.3 nats if the path is fouled, turning the present 2.2 to one against the path into about 250 to one for it, and 44 nats the other way if the sensor is biased. The return sensor adds nothing to this pair: the fouled path and the hall-1 sensor bias predict the same supply and the same return.

For the subset's pair and the return sensor read over ten days the arithmetic of equation (23.17) is, with $\mathrm{sd}(y_a \mid \mathbf{y}_2, M_1)^2 = 0.111^2 + 0.095^2 = 0.0213$, $\mathrm{sd}(y_a \mid \mathbf{y}_2, M_3)^2 = 0.439^2 + 0.095^2 = 0.2017$ and $\mu_1 - \mu_3 = -0.806$,

$$\mathbb{E}\big[\Delta \ln B_{13} \mid M_1\big] = \tfrac{1}{2}\left[\ln \frac{0.2017}{0.0213} + \frac{0.0213 + 0.6496}{0.2017} - 1\right] = \tfrac{1}{2}\,(2.248 + 3.326 - 1) = 2.29.$$

In words: if hall 1's path is fouled, ten days of a return sensor of 0.3 K are expected to add 2.3 nats to the case against the supply bias, taking its odds from 1 in 11 to about 1 in 110; if the

Table 23.16: Candidate readings and the expected change of the evidence ratio of the two drawings named, at the ten-day state. Predictive means ± spreads $\mathrm{sd}(y_a \mid \mathbf{y}_2, M)$ in °C; the spread includes the reading's own σ_a. The last two columns are equation (23.17) with the first and the second drawing of the pair true.

Candidate reading	Pair	σ_a [K]	First	Second	Separation	If first	If second
second hall-1 sensor, one day	M_1, M_5	0.50	19.80 ± 0.52	17.76 ± 0.77	2.20	3.67	7.78
second hall-1 sensor, ten days	M_1, M_5	0.158	19.80 ± 0.22	17.76 ± 0.60	3.18	6.32	43.9
return sensor, ten days	M_1, M_5	0.095	17.50 ± 0.15	17.50 ± 0.15	0.02	0.00	0.00
return sensor, one day	M_1, M_3	0.30	17.50 ± 0.32	18.30 ± 0.53	1.30	1.34	3.55
return sensor, ten days	M_1, M_3	0.095	17.50 ± 0.15	18.30 ± 0.45	1.71	2.29	18.3
second supply sensor, one day	M_1, M_3	0.30	12.20 ± 0.31	13.02 ± 0.53	1.32	1.38	3.80
second supply sensor, ten days	M_1, M_3	0.095	12.20 ± 0.13	13.02 ± 0.45	1.74	2.40	22.8

supply sensor is in fact biased, they would reverse the ranking by some 18 nats. A check by refits, drawing 400 return readings from the M_1 predictive and refitting both drawings with the added row, gave an average gain of 2.32 ± 0.04 for the ten-day sensor and 1.47 ± 0.05 for the one-day reading (Monte Carlo standard errors), against 2.29 and 1.34 from the Gaussian formula: the formula holds for the ten-day sensor and understates the one-day reading by 0.13, where the reading's predictive is less Gaussian; the gain was negative for 3 % and 9 % of the draws. A second supply sensor does the same work, since the two drawings differ in the true supply temperature and the return temperature inherits that difference.

Figure 23.6 draws this pair's separation on the return temperature. The horizontal axis is T_r and the vertical axis a probability density; the solid curves are the posteriors of the return temperature under M_1 (blue) and M_3 (red), and the dashed curves the predictive densities of a one-day return reading of 0.3 K under each. The two posteriors overlap only in the upper tail of the first: under the fouled path the return is known to 0.11 K around 17.50°C, under the supply bias it is near 18.30°C with a spread four times as wide. One reading blurs both by its own 0.3 K and the predictive curves overlap more, which is why a single day of it is worth 1.3 nats and ten days 2.3.

23.6.4 The recommendation

The next measurement is a second temperature sensor in hall 1, read for ten days in the same operating condition. It settles the open question at an expected 6.3 nats if the path is fouled and 44 if the first sensor is biased, and it gives the hall-1 reading the partner that Chapter 18 recommended for the hall-2 reading. A sensor on the return header would in addition retire the supply-bias drawing, at 2.3 nats, and check the supply sensor for every later record; it does nothing for the leading pair. The coil is not cleaned before the second sensor reports. If it confirms the fouled path, the coil is cleaned and the cleaned plant becomes the healthy model of the next period; if it confirms the biased sensor, the sensor is recalibrated and period 2 is refitted without its bias.

23.6.5 What the analysis showed

The first reading of period 2 said that the hall-1 coil had fouled to 3.88 kW/K and should be cleaned; the analysis keeps the number and withdraws the conclusion. Ten days of readings do show that this plant changed and that the change is in hall 1, and if the coil fouled, it fouled to

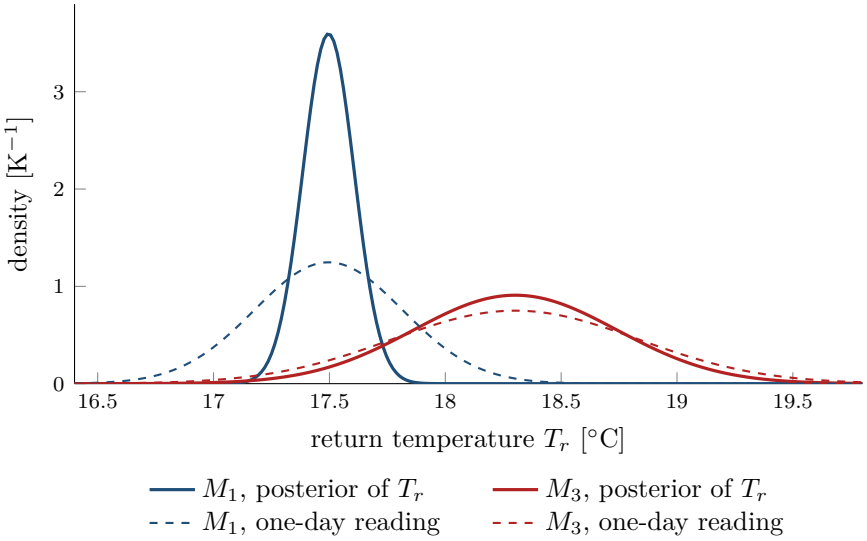


Figure 23.6: The return temperature under the two leading drawings of the worked subset at the ten-day state. Solid: posterior densities $\mathcal{N}(17.50,\ 0.111^2)$ under M_1 and $\mathcal{N}(18.30,\ 0.439^2)$ under M_3; dashed: predictive densities of a one-day return reading with $\sigma_a = 0.3\,\mathrm{K}$, $\mathcal{N}(17.50,\ 0.320^2)$ and $\mathcal{N}(18.30,\ 0.532^2)$ (Table 23.16).

that value, known to about 3 %. But the hall-1 sensor is the only reading that sees the hall-1 conductance, so a coil that lost 29 % and a sensor that reads two kelvin high write the same record, and the record's verdict between them is the ratio of their priors, two to one in favor of the sensor at the stated widths. That blindness is a property of the five sensors, not of the ten days: neither a longer record nor the comparison of the two periods' fits removes it, and a second view of hall 1 does, at little cost. The analysis also says what the record could not have done sooner and could not do later. One day would only have suggested a change; beyond about thirty days nothing improves, because the baseline is a single snapshot, and the next healthy model of the two halls should be built from a longer first period. The engineer carries back to the site a hold on the work order, a second hall-1 sensor for ten days, and a note to keep the daily readings, not only their averages, so that the next analysis can check the sensors' stated spreads.

23.6.6 Limits of the model

The numbers of this chapter rest on the family, the healthy model and the rows; the checks test them only in part.

Causes outside the family. The posterior of equation (23.11) is normalized over the family, and a cause outside it is assigned to whichever listed drawing mimics it best. Two checks guard against that. The first is the misfit of the leading drawing: a cause that no drawing represents leaves a misfit in every fit, and a leading drawing with $2\mathcal{L}$ far in the tail of its chi-square says that none of the drawings is the right one; the two leading drawings here leave none. The second check is the one the first cannot do. The hall-1 sensor bias is outside the worked subset and leaves no trace in the residuals of M_1, because it explains the record through the same row as the fouled path; a residual check catches causes that leave misfit, not causes confounded with a listed one. The remedy is to generate the family by the rules in full, as Section 23.1.5 did, and to read the pattern of equation (23.4) for pairs of drawings whose freed quantities reach the readings through one row only: those pairs are not separable by any record of the present readings, and only a new row separates them.

Two faults at once. A family of single changes does not contain them. A drawing that frees both quantities competes with the single-cause drawings and pays two Occam factors, so it wins only when neither single change explains the record. Exercise 23.5 works a fouled path together with a step of the hall-1 load, which the family handles without a two-cause drawing because the loads are free in every drawing: an operating change is not a fault, and the load meter reads it.

The healthy model and the releases. The healthy model takes period 1 as fault-free; a fault already present then would be part of the baseline. It carries the marginals of the period-1 conductances and drops their correlation of -0.03; with a larger correlation the joint of the two would be carried. The released prior is symmetric, while fouling only lowers a conductance; a half-normal release of the same spread would at most double the drawing's prior density where the posterior lies, raising $\ln B_{10}$ by at most $\ln 2 = 0.69$, and it would not change the confound with the sensor bias. The drawing M_4 represents a chiller-meter fault, not a derated chiller. A chiller that cannot carry the load shows as a supply above its set point, the drawing M_8, which the supply sensor reads directly and the record does not need ($\ln B = -1.07$). Equal prior odds are a convention: a plant with a history of coil fouling, or of sensor drift, should enter it through $\Pr(M_j)$, and Exercise 23.1 shows the change. The ranking also depends on the stated spreads σ_{w} and σ_b, and is reported with them.

The rows and the record. Each day of a record is taken at one steady state; a period whose loads change from day to day needs one state per day, with the conductances shared, and its average is not a sufficient summary. The stated standard deviations of the readings are taken as known, and the daily scatter that would test them was not recorded. The period-2 values were used at every record length so that only the precision changed; a real average of many days would sit closer to the plant's values, so the long-record rows of Table 23.11 show what precision buys, not what a long record would read. In the comparison of two periods, unequal record lengths shrink the two fits unequally toward the handbook and create a small apparent drift in the entries the priors hold, a quarter of a spread for g_2 here, and the five standardized changes are five looks, which the joint statistic accounts for and the largest $|Z_i|$ alone does not. The Laplace evidences agreed with importance sampling here, and the fault size is read from the Laplace marginal; a drawing whose freed quantity is poorly determined, or a posterior with a long tail, needs both checked again.

The evidence has served here, as in the previous chapter, to rank drawings of one plant that differ in a freed quantity; the next chapter derives it in general and applies it to drawings that differ in the structure of the graph itself.

23.7 What to retain

Definition 23.1 (Fault detection, isolation and attribution on a fitted plant)**.** A record of N days enters a fit as its averages, each with standard deviation $\sigma_k/\sqrt{N}$. The healthy model carries an earlier period's posterior of the plant's properties as prior rows, whose standard deviations are that period's spreads; detection is the chi-square tail of its misfit on the new record, which saturates as N grows at the conflict between the readings-only state and the healthy model. A candidate drawing frees one quantity of the healthy model, by releasing a prior, adding a bias to a reading or freeing a pinned quantity, and the family is written before the record is examined; the pattern of the Jacobian shows which drawings share a footprint on the readings. The evidence of each drawing is the Laplace formula $-\mathcal{L}_M - \frac{1}{2}\ln\det\boldsymbol{\Lambda}_M + \frac{n_z}{2}\ln 2\pi - \sum_k \ln(\sqrt{2\pi}\sigma_k)$, in which a widened prior contributes $\ln(\sigma_{\mathrm{h}}/\sigma_{\mathrm{w}})$ and an added entry $-\ln\sigma_b$ against the healthy model; the posterior probabilities over the family are the isolation, conditional on the family, and the

freed quantity's posterior under each leading drawing is the identification. Attribution compares two fits made with independent priors: the change is approximately $\mathcal{N}(\hat{\mathbf{z}}_2 - \hat{\mathbf{z}}_1,\ \boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)$, read entry by entry as Z_i and jointly as χ^2_Δ against $\chi^2_{n_z}$. A new reading's expected effect on an evidence ratio is the Kullback–Leibler divergence of the two drawings' predictive densities of that reading, with spreads $\mathrm{sd}(y_a \mid \mathbf{y}_2, M_j)$.

Three conclusions carry beyond the two halls. A fault diagnosis is a comparison of small changes of the healthy model, each generated from the plant graph and each charged by the evidence for the freedom it takes: the fouled hall-1 path wins against the healthy model because it removes a conflict of three healthy spreads at the price of 2.24 nats, and the fouled hall-2 path loses because it pays the price for nothing. The length of a record is bounded in what it can buy by the precision of the baseline it is compared with, and the confounds of the design are not bought off by any length: a fault on a quantity that a single reading sees cannot be told from a fault of that reading, and the remedy is a row, chosen by the expected change of the evidence ratio it would bring. And the two methods are complements: attribution locates the change without a list and without a cause, isolation names a cause from a list and is only as complete as the list. Appendix A collects these questions with those of the other chapters.

Table 23.17 lists the terms of the chapter with their expressions, units and the subsection that defines each.

Table 23.17: The terms of fault detection, isolation and drift attribution and where each is defined.

Term	Expression	Units	Meaning	Defined in
Record of N days	$\bar{y}_k$ with $\sigma_k/\sqrt{N}$, eq. (23.5)	reading	the average is the record	23.1.5
Healthy model	period-1 marginals as prior rows, eq. (23.3)	–	the baseline	23.1.5
Candidate drawing	one quantity released, biased or freed	–	one fault	23.1.5
Footprint	a candidate's column of $\mathbf{J}$ on the reading rows, eq. (23.4)	–	what the readings see of it	23.1.5
Leave-one-out departure	d_k, eq. (21.22)	–	reading-wise detection	23.3.1
Posterior over the family	eq. (23.11)	–	isolation	23.3.2
Laplace evidence	eq. (23.10)	–	with row constants	23.2.1
Widened-prior constant	$\ln(\sigma_\mathrm{h}/\sigma_\mathrm{w})$, eq. (23.12)	–	price of a release	23.2.1
Added-entry constant	$-\ln\sigma_b$, eq. (22.13)	–	price of an entry	23.2.1
Occam factor	$\frac{1}{2}\ln[(\sigma_\mathrm{h}^2+v_\mathrm{L})/(\sigma_\mathrm{w}^2+v_\mathrm{L})]$, eq. (23.13)	–	charge for freedom	23.2.1
Confounded drawings	freed quantities seen through one row	–	not separable by length	23.2.1
Fault size	$\exp(\hat{g}_1 \pm 1.960\sqrt{\Sigma_{g_1 g_1}})$	kW/K	identification	23.3.3
Change between periods	$\mathcal{N}(\hat{\mathbf{z}}_2-\hat{\mathbf{z}}_1,\ \boldsymbol{\Sigma}_1+\boldsymbol{\Sigma}_2)$, eq. (23.14)	entry	parameter-change detection (attribution)	23.3.4
Standardized change	$Z_i=\hat{\Delta}_i/\sqrt{(\boldsymbol{\Sigma}_1+\boldsymbol{\Sigma}_2)_{ii}}$	–	which entry moved	23.3.4
Joint statistic	$\chi^2_\Delta=\hat{\boldsymbol{\Delta}}^\mathsf{T}(\boldsymbol{\Sigma}_1+\boldsymbol{\Sigma}_2)^{-1}\hat{\boldsymbol{\Delta}}$, eq. (23.15)	–	whether anything moved	23.3.4
Detection limit	$\sum_{\text{prior rows}} r_k(\mathbf{z}_\mathrm{r})^2$, eq. (23.16)	–	set by the baseline	23.3.5
Predictive spread of a reading	$\mathrm{sd}(y_a \mid \mathbf{y}_2, M_j)^2=\mathbf{g}_a^\mathsf{T}\boldsymbol{\Sigma}_{M_j}\mathbf{g}_a+\sigma_a^2$	reading	under drawing M_j	23.6.2
Expected evidence change	Kullback–Leibler divergence, eq. (23.17)	nat	value of a reading for a pair	23.6.2

23.8 Exercises

Exercise 23.1. *Formulation.* (a) At ten days the log Bayes factors of the worked subset against the healthy model are +2.633, −1.929, +0.210 and −0.151 for M_1 to M_4. State the posterior probabilities of the five drawings at equal prior odds, and with prior probabilities 0.6 for M_0 and 0.1 for each fault. (b) A change of two entries is $\hat{\boldsymbol{\Delta}}=(-0.30,\ +0.10)$ with spreads 0.11 and 0.14 and correlation 0.3. State the standardized changes, χ^2_Δ and its tail on two degrees of freedom, and the same without the correlation. ((a) $\mathrm{e}^{\ln B_{j0}}=1$, 13.92, 0.145, 1.234, 0.860, sum 17.16, so 0.058, 0.811, 0.008, 0.072, 0.050; with the unequal priors the weights are 0.6, 1.392, 0.0145, 0.1234, 0.0860, sum 2.216, so 0.271, 0.628, 0.007, 0.056, 0.039: a prior that favors the healthy plant moves the fouled path from 0.81 to 0.63. (b) $Z=-2.73$, $+0.71$; with $\mathbf{C}=[0.0121,\ 0.00462;\ 0.00462,\ 0.0196]$, $\det\mathbf{C}=2.158\times10^{-4}$ and $\chi^2_\Delta=(0.0196\times0.09+2\times0.00462\times0.03+0.0121\times0.01)/2.158\times10^{-4}=10.02$; the tail of χ^2_2 is $\mathrm{e}^{-\chi^2_\Delta/2}=0.0067$. Without the correlation $\chi^2_\Delta=2.73^2+0.71^2=7.95$, tail 0.019: changes of

opposite sign in positively correlated entries are less likely together, and the correlation raises the statistic.)

Exercise 23.2. *Derivation.* From the Laplace evidence with every row's normalizer, derive the log Bayes factor against the healthy model of a drawing that widens one prior from σ_{h} to σ_{w}, and of a drawing that adds one entry with a prior $\mathcal{N}(0, \sigma_b^2)$ attached to one reading. Show also that the first ratio equals $B_{01} = \mathbb{E}[p_0(g_1)/p_1(g_1) \mid \mathbf{y}, M_1]$ exactly. (Equation (23.10): $\ln p(\mathbf{y} \mid M) \approx -\mathcal{L}_M - \frac{1}{2}\ln\det\boldsymbol{\Lambda}_M + \frac{n_z}{2}\ln 2\pi - \sum_k \ln(\sqrt{2\pi}\sigma_k) + \ln c_Q$; the reading rows and c_Q cancel. Widened: same n_z, one row's normalizer changes, $-\ln(\sqrt{2\pi}\sigma_{\mathrm{w}}) + \ln(\sqrt{2\pi}\sigma_{\mathrm{h}}) = \ln(\sigma_{\mathrm{h}}/\sigma_{\mathrm{w}})$. Added: $+\frac{1}{2}\ln 2\pi$ for the entry and $-\ln(\sqrt{2\pi}\sigma_b)$ for its row, net $-\ln\sigma_b$. Identity: $p(\mathbf{y} \mid M_0) = \int p(\mathbf{y} \mid \mathbf{z})\, p_1(\mathbf{z})\, [p_0/p_1]\, \mathrm{d}\mathbf{z} = p(\mathbf{y} \mid M_1) \int p(\mathbf{z} \mid \mathbf{y}, M_1)\, [p_0/p_1]\, \mathrm{d}\mathbf{z}$.)

Exercise 23.3. *Derivation.* Write $\mathrm{Cov}(\hat{\mathbf{z}}_2 - \hat{\mathbf{z}}_1)$ in general, and show that the cross terms vanish when the periods have independent priors and independent readings. For linear rows, show that if the two periods share one true state drawn from the handbook prior with precision $\boldsymbol{\Lambda}_{\mathrm{prior}}$, the estimation errors $\boldsymbol{\delta}_p = \hat{\mathbf{z}}_p - \mathbf{z}$ of the periods $p = 1, 2$ satisfy $\mathrm{Cov}(\boldsymbol{\delta}_1, \boldsymbol{\delta}_2) = \boldsymbol{\Sigma}_1\boldsymbol{\Lambda}_{\mathrm{prior}}\boldsymbol{\Sigma}_2$, and state the effect on Z_i. ($\mathrm{Cov} = \boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2 - \mathrm{Cov}(\mathbf{z}_1, \mathbf{z}_2) - \mathrm{Cov}(\mathbf{z}_2, \mathbf{z}_1)$, and the block-diagonal posterior of Section 23.3.4 makes the cross terms zero. Linear rows: $\hat{\mathbf{z}}_p = \boldsymbol{\Sigma}_p(\boldsymbol{\Lambda}_{\mathrm{prior}}\boldsymbol{\mu}_0 + \mathbf{J}_p^{\mathsf{T}}\mathbf{W}_p\mathbf{y}_p)$, $\boldsymbol{\Sigma}_p^{-1} = \boldsymbol{\Lambda}_{\mathrm{prior}} + \mathbf{J}_p^{\mathsf{T}}\mathbf{W}_p\mathbf{J}_p$, $\mathbf{y}_p = \mathbf{J}_p\mathbf{z} + \boldsymbol{\varepsilon}_p$, so $\boldsymbol{\delta}_p = \boldsymbol{\Sigma}_p\boldsymbol{\Lambda}_{\mathrm{prior}}(\boldsymbol{\mu}_0 - \mathbf{z}) + \boldsymbol{\Sigma}_p\mathbf{J}_p^{\mathsf{T}}\mathbf{W}_p\boldsymbol{\varepsilon}_p$; with $\mathrm{Cov}(\mathbf{z}) = \boldsymbol{\Lambda}_{\mathrm{prior}}^{-1}$ on the entries with priors and independent errors, $\mathrm{Cov}(\boldsymbol{\delta}_1, \boldsymbol{\delta}_2) = \boldsymbol{\Sigma}_1\boldsymbol{\Lambda}_{\mathrm{prior}}\boldsymbol{\Sigma}_2$, and $\mathrm{Var}(\hat{\boldsymbol{\Delta}}) = \boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2 - \boldsymbol{\Sigma}_1\boldsymbol{\Lambda}_{\mathrm{prior}}\boldsymbol{\Sigma}_2 - \boldsymbol{\Sigma}_2\boldsymbol{\Lambda}_{\mathrm{prior}}\boldsymbol{\Sigma}_1$ has a smaller diagonal: Z_i computed with $\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2$ are conservative. On the plant Z_1 moves from -3.12 to -3.15 and χ^2_Δ from 14.47 to 14.62.)

Exercise 23.4. *Derivation.* Show that a drawing outside the family can never be ranked by equation (23.11), and how the misfit of the leading drawing reveals a cause that leaves misfit. Then show that two drawings whose freed quantities reach the readings through one row only cannot be told apart by any misfit, using the fouled hall-1 path and the hall-1 sensor bias. (The denominator of equation (23.11) sums over the listed drawings, so an unlisted cause has no term and its probability is shared among the listed ones in proportion to their evidences. If none mimics it, every fit leaves a misfit, and $2\mathcal{L}$ of the leading drawing lies in the upper tail of χ^2_ν; here $2\mathcal{L}_{M_1} = 0.17$, tail 0.98. The hall-1 row is $T_s + Q_1\mathrm{e}^{-g_1} + b_{T_1} - y_{T_1}$, and g_1 and b_{T_1} enter no other reading row, the shared footprint of equation (23.4), so for every value of one there is a value of the other with the same predicted readings: the two drawings differ only in their prior rows ($2\mathcal{L} = 0.17$ and 1.19), and their evidence ratio, $\ln B_{51} = 0.78$–0.79, hardly moves with the record length, Table 23.11.)

Exercise 23.5. *Application.* In another period 2 the hall-1 path fouls as before ($G_1 = 4.0\,\mathrm{kW/K}$) and the hall-1 load steps from 30 to 36 kW, with the errors of Table 23.1: the ten-day readings are 21.30°C, 17.90°C, 12.20°C, 59.0 kW and 35.5 kW. Rank the five drawings, run the attribution, and say why the single-cause family suffices and what the attribution adds. (The healthy model gives $2\mathcal{L}_{M_0} = 9.81$, tail 0.020; $\ln B_{j0} = +2.59$, -1.93, -0.13, -0.15 and $\Pr(M_1) = 0.82$, with $\hat{G}_1 = 3.90\,\mathrm{kW/K}$; one day gives $+1.20$ and 0.54. Attribution: $Z = -3.11$ on g_1 and $+3.15$ on Q_1, a change of $+4.81 \pm 1.53\,\mathrm{kW}$, with $\chi^2_\Delta = 50.9$, tail 9×10^{-10}. The load step needs no drawing, because the loads are free in every drawing and the load meter reads the step; it is an operating change, not a fault. Attribution reports two moved entries and cannot say that one is a fault and the other a change of duty; the family can, because only one of them was ever held by a prior.)

Exercise 23.6. *Application.* In another period 2 nothing in the plant changes, but the supply sensor drifts to read 0.8 K low: the ten-day readings are 17.755°C, 17.90°C, 11.40°C, 53.0 kW and 29.5 kW. Run detection, the ranking of the five drawings and the attribution, and state where the two methods disagree. (The healthy model gives $2\mathcal{L}_{M_0} = 3.23$, tail 0.36: no detection,

because the supply set-point prior of $\pm 1\,\mathrm{K}$ accommodates a supply of 11.4°C. The ranking puts the supply bias first but weakly, $\ln B_{30} = +0.18$ with $\hat{b} = -0.79 \pm 0.44\,\mathrm{K}$ and probability 0.33, against 0.27 for the healthy model and 0.26 for the meter offset. Attribution reports the largest change in the supply temperature, $-0.70 \pm 0.30\,\mathrm{K}$, $Z = -2.37$, and a lesser fall of g_1, $Z = -1.49$, with $\chi^2_\Delta = 6.52$, tail 0.26. Attribution says the supply got colder, which it did not: the only reading of T_s is the biased one, and the entry follows it. Isolation, whose family contains the bias, names the right cause without being sure of it, and a return sensor, which sees the true supply through the loop balance, would settle it.)

Chapter 24

Model Selection and Bayes Factors

24.1 Questions about the drawing

The reliability engineer who signs the data center's availability report has been taking three statements from the drawings. The fits of Chapters 18 to 23 put hall 2 of the two-hall plant on the shared chilled-water loop; Chapter 13 reported the cooling of the equipment hall over the design file's two structures, three of its four CRAC units needed with weight 0.7 and two with weight 0.3; and the circulation pump of Chapter 6 is planned at the equipment handbook's rate of $0.001\,\mathrm{h}^{-1}$, although its own record holds four failures in 2000 h. Two new pieces of paper now sit on the desk: a commissioning note from a later extension, which says that hall 2 was moved to a branch of its own, and the hall's daily log of 400 days, which records how many units were down at each morning check and whether the hall lost cooling that day. The first reading is that every drawing fits, so each stands as drawn. The shared-loop fit passed its chi-square check, $2\mathcal{L} = 0.740$ against χ^2_3 with a tail of 0.864; the pump's central 95 % interval, $[0.73,\ 3.89] \times 10^{-3}\,\mathrm{h}^{-1}$, contains the handbook rate; and the design file's mixture gives the hall's cooling a probability of 0.321 (0.384 exactly) of meeting its target unavailability of 10^{-3}, the number the report would carry.

A drawing that fits says only that it could have produced the record. The commissioning note proposes a second drawing of the loop, and the chi-square check of the first cannot say whether the second fits as well or better; an interval that contains the handbook rate says that the rate is plausible, not that the pump is the machine the handbook describes; and the design file's 0.7 is a prior, which a record of losses can confirm or overturn. The question of this chapter is

which drawing of the plant the record supports, and how strongly; how the evidence of a drawing, the normalizing constant of its posterior, is derived and computed; how a reported quantity is formed when the drawing stays uncertain; and which reading would settle a structure question before it is installed.

It is answered with the method of the earlier chapters, with one change: the unknown is the drawing itself, a discrete unknown over which each drawing's state is integrated out, and the computation is the evidence of each drawing rather than the posterior inside one.

Jargon: model selection, drawing, evidence (marginal likelihood), Bayes factor, nat, prior and posterior odds, posterior model probability, Occam factor, nested model, point null, Savage–Dickey ratio, conjugate rows, nuisance quantity, Laplace evidence, improper prior, Lindley's paradox, mixture over drawings, model averaging, between-drawing variance, Kullback–Leibler divergence

Model selection is the choice among candidate models of the same data; here each candidate is a *drawing*, a plant graph with its rows and priors. The *evidence* or *marginal likelihood* $p(\mathbf{y} \mid M)$ of a drawing M is the probability of the record under it with its unknowns integrated over their prior. The *Bayes factor* is the ratio of two evidences; its natural logarithm is measured in *nats*. *Prior odds* and *posterior odds* are ratios of the probabilities of two drawings before and after the record; a *posterior model probability* is the probability of one drawing given the record. The *Occam factor* is the ratio of the posterior width of a freed quantity to its prior width, the factor by which the evidence charges for the freedom. A drawing is *nested* in another when it is the other with one quantity pinned at a value, a *point null*; the *Savage–Dickey ratio* gives the Bayes factor of the two as a ratio of a posterior to a prior density at that value. *Conjugate rows* are those whose prior and likelihood combine into a posterior of the prior's family, so that the evidence is a closed form. A *nuisance quantity* is an unknown not asked about, integrated out. The *Laplace evidence* is the evidence computed from the Gaussian approximation of the posterior at the fitted state. An *improper prior* is a constant density that does not integrate to one. *Lindley's paradox* is the standard name for the behavior of a Bayes factor under a widening prior: it moves toward the point null without bound, even for records that a significance test rejects. A *mixture over drawings*, or *model averaging*, reports a quantity as the probability-weighted average of its value under each drawing; the *between-drawing variance* is the part of its variance that comes from the uncertain drawing. The *Kullback–Leibler divergence* of one distribution from another is the expected logarithm of their density ratio under the first; it is the expected evidence a reading adds for the drawing that is true.

Key idea

A structure question asks which graph the plant is, and it is answered by the ratio of the evidences of the graphs: the probability each drawing gave the record before the record was seen, its state integrated over its prior. The evidence rewards the drawing that predicted the readings and charges a freed quantity by its Occam factor, the ratio of its posterior to its prior width. A reading that two drawings predict alike adds nothing to the ratio, however well both fit it; the reading that settles the question is the one whose predictions differ most across the drawings, and that is known before the reading is installed. When the drawing stays open, a reported quantity is the mixture of its values under the drawings with their posterior probabilities as weights, and its band carries the doubt about the drawing.

24.1.1 Drawings that differ in the graph

The earlier comparisons of the book held the graph fixed and changed a distribution or freed one quantity on it. The Weibull model of Chapter 7 differs from the exponential in one shape entry; the bias drawings of Chapter 22 add one entry to one reading row; the fault drawings of Chapter 23 widen the prior of one conductance or add an offset to one meter. Chapter 15 compared a shared cause against independence, and Chapter 13 left an open structure, 3-out-of-4 or 2-out-of-4, averaged with weights from a design file because no evidence on the structure was at hand. The drawings of this chapter differ in the structure of the plant. Whether hall 2 draws its chilled water from the shared supply header or from a branch of its own changes which node its balance is written against and which heat the return header carries, so it changes two rows

and adds a node. Whether the hall's cooling needs three units or two changes the structure function of Chapter 13, the map from the units' states to the service. Whether the pump's rate is the handbook's or its own changes a free entry of the state into a pinned one. In every case the question is *which graph*, and no posterior inside one graph can answer it, because each posterior is conditional on its own graph being right.

The questions are of a kind the field asks constantly and seldom formulates: is it really parallel; is this branch really on that loop; is this the plate rating or a different machine. The engineering instinct is to fit each drawing and look at which fits better. The instinct is half right. A better fit is part of the answer, and the other part is the price of the freedom a drawing used to achieve it: a drawing that frees a quantity can fit any record that quantity can absorb, and a fair comparison must charge it for that.

24.1.2 The plants and their drawings (step 0)

The three plants are already in the book; what is new is that each has two or more candidate drawings, and the drawings are the unknown.

The loop. The first plant is the two-hall plant of Chapter 17, with the return-header temperature sensor that Chapter 22 added. The return header carries the water of the halls that the shared supply header feeds, with the capacity rate $\dot{m}c_p = 10\,\mathrm{kW/K}$ read from its flow meter and the properties of water, pinned as in Chapter 21. The design drawings say that both halls are on the shared loop; the commissioning note says that hall 2 was moved to a heat exchanger off the chiller that delivers water at a supply temperature of its own and whose water returns to the chiller without passing the shared header.

Figure 24.1 draws the two drawings side by side. In each panel the green boxes are free nodes: the two halls with their temperatures, the supply header with the free supply temperature T_s and, in panel (b), a second supply node with its own free temperature T_{s2}. The white box at the bottom is the chiller, the source of chilled water, and the green box at the top is the return header, whose temperature T_r is fixed by the heat it receives. The solid green arrows are the water paths, and the open stubs marked "to chiller" carry the warmed water back; in panel (b) hall 2's water returns without passing the header. The red dashed box is the return sensor, the reading the two drawings predict differently. The other five readings are attached as in Figure 17.1 and are not drawn.

The cooling. The second plant is the hall of Chapter 13, cooled by four CRAC units. The design file holds the hall to need three of them at its design load, a 3-out-of-4 drawing, while the present partial load may need only two, a 2-out-of-4 drawing. In the graph of that chapter the two drawings share the four unit nodes and differ in the structure function that joins them to the service. Under each drawing the cooling is lost on a day with j units down whenever $j \geq m$, with m the smallest number of units down that defeats it, five less the number of units the drawing needs: $m = 2$ for 3-out-of-4, $m = 3$ for 2-out-of-4, $m = 1$ for 4-out-of-4 and $m = 4$ for 1-out-of-4, in which the hall keeps its cooling until every unit is down. The drawings are M_3 (3-out-of-4) and M_2 (2-out-of-4), with M_4 and the 1-out-of-4 drawing completing the family. The structure also admits losses that it does not explain, a controls fault, a power dip, a door left open: every drawing carries a node for other causes, which defeats the cooling on any day.

The pump. The third plant is the circulation pump of Chapter 6. Two drawings compete: M_1, the pump's rate free, and M_0, the pump's rate pinned at the handbook value $\lambda_\mathrm{h} = 0.001\,\mathrm{h}^{-1}$. In the graph of Figure 6.1(a) the difference is whether the pump node carries a free entry or a pinned one; M_0 is nested in M_1 at a point.

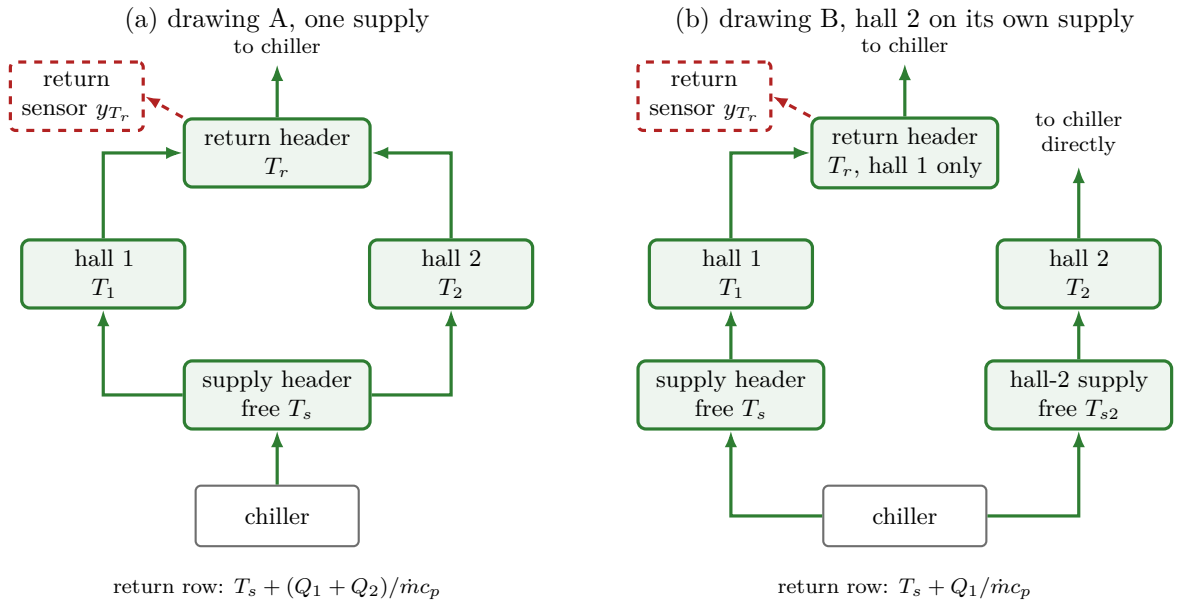


Figure 24.1: The two drawings of the two-hall plant. Free (green): the hall temperatures, fixed by the balances, and the state entries g_1, g_2, Q_1, Q_2 and T_s; in drawing B also the supply temperature T_{s2} of hall 2's own branch. Pinned: the capacity rate $\dot{m}c_p = 10\,\text{kW/K}$ of the water through the return header. Observed: the return sensor (red, dashed), whose row is written under each panel, and the five readings of Figure 17.1, not drawn. The chiller's heat meter reads $Q_1 + Q_2$ in both drawings.

24.1.3 The records and the states

The two records of the loop. The comparison of the loop is made on two records. *Record A* is the clean record of Chapter 22: the five readings of Table 17.1, generated under drawing A, and the return reading $y_{T_r} = 17.50°\text{C}$ with error standard deviation $\sigma_{T_r} = 0.3\,\text{K}$. *Record B* was generated under drawing B with hall 2 fed at 9°C: with $G_2 = 4.0\,\text{kW/K}$ and $Q_2 = 24\,\text{kW}$ the hall sits at $9 + 24/4 = 15.0°\text{C}$, and its sensor, with an error of $-0.2\,\text{K}$, reads $y_{T_2} = 14.80°\text{C}$; the return header carries only hall 1's 30 kW and reads $12.0 + 30/10 + 0.1 = 15.10°\text{C}$. The other four readings are those of record A. Both records are illustrative. The priors are those of Table 17.2, and drawing B adds one: the supply temperature of hall 2's branch has the prior of a second set point, $T_{s2} \sim \mathcal{N}(12°\text{C},\ 1\,\text{K}^2)$, as for T_s.

The loss record. Figure 24.2 shows the hall's record of 400 days. The five columns are the number of units down at the check, $j = 0$ to 4, with the number of days n_j in each; the upper row of cells counts the days on which the hall lost cooling, k_j, shaded red when nonzero, and the lower row the days on which it kept it. The two dashed vertical lines are the loss rules of the two drawings: under 3-out-of-4 the cooling is lost whenever two or more units are down, the line between $j = 1$ and $j = 2$; under 2-out-of-4 it is lost whenever three or more are down, the line between $j = 2$ and $j = 3$. A loss is a high-temperature alarm that required load to be shed. Summed over the days, the units-down column holds $40 + 2 \times 7 + 3 \times 1 = 57$ units found down in $4 \times 400 = 1600$ unit-checks.

The log shows losses on days the structure explains none, two with no unit down and one with one down. Every drawing therefore carries a nuisance quantity, the daily probability ν_o of a loss from other causes, with the prior $\nu_\text{o} \sim \text{Beta}(1, 99)$, of mean 0.01, one such day in a hundred. On a day with $j < m$ the hall loses cooling with probability ν_o; on a day with $j \geq m$ it loses it with probability one. A second, softer record differs in one column: two of the seven

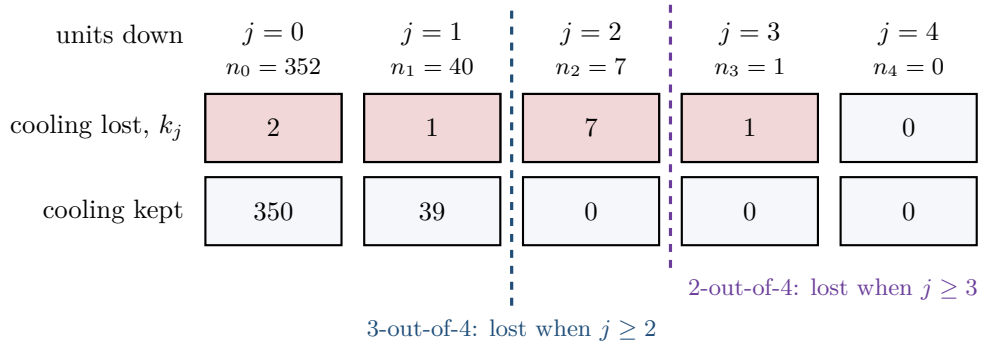


Figure 24.2: The daily loss record of the hall, 400 days: days with j CRAC units down at the check (n_j, top), days on which the hall lost cooling (k_j, upper cells, red when nonzero) and days on which it kept it (lower cells). Dashed lines: the loss rule of the 3-out-of-4 drawing (blue) and of the 2-out-of-4 drawing (purple). In the softer record two of the seven two-down days have no loss ($k_2 = 5$). The record is illustrative.

two-down days passed without a loss, $k_2 = 5$ in place of 7. Under 3-out-of-4 with a sure loss on every two-down day this record is impossible, and a drawing that is to survive it needs a second nuisance probability, the probability ν_{m} that a day the structure says is lost passes without a recorded loss, because the hall rode through on its thermal mass or the load happened to be low; its prior is $\nu_{\mathrm{m}} \sim \mathrm{Beta}(1, 19)$, of mean 0.05. On a day with $j \geq m$ the hall then loses cooling with probability $1 - \nu_{\mathrm{m}}$.

The pump's record. The pump has $k = 4$ failures in $T_{\mathrm{tot}} = 2000$ operating hours and the engineering prior $\lambda \sim \mathrm{Gamma}(2,\ 1000\,\mathrm{h})$ of equation (6.18).

The states. Each drawing carries the state the book carries for its quantities. The loop runs on the reduced state of Chapter 17, equation (17.7), five entries under drawing A and six under drawing B,

$$\mathbf{z}_{\mathrm{A}} = \begin{bmatrix} g_1 & g_2 & Q_1 & Q_2 & T_s \end{bmatrix}^{\mathsf{T}}, \qquad \mathbf{z}_{\mathrm{B}} = \begin{bmatrix} g_1 & g_2 & Q_1 & Q_2 & T_s & T_{s2} \end{bmatrix}^{\mathsf{T}}, \tag{24.1}$$

the hall temperatures and the return temperature being functions of them on the balance set. The probability ν_{o} of the cooling is carried as its logit $\xi_{\mathrm{o}} = \ln(\nu_{\mathrm{o}}/(1 - \nu_{\mathrm{o}}))$, the one number that is the state of every cooling drawing, with ξ_{m} added on the softer record. The pump's rate is carried as the log-rate $\omega = \ln(\lambda/\lambda_{\mathrm{ref}})$ of equation (6.11) with $\lambda_{\mathrm{ref}} = 10^{-3}\,\mathrm{h}^{-1}$, the one number that is the state of M_1; the handbook drawing is $\omega = 0$ and has no free entry. Where the general forms write $\mathbf{z}$ for these states, $\mathbf{z}$ is the one number ξ_{o} or ω, and every quantity of the fit is a number: each row's term $\mathcal{L}_k$, its slope $\mathcal{L}_k'$ and its curvature $\mathcal{L}_k''$ in that number (primes are derivatives with respect to it), the information $\Lambda = \mathcal{L}''$ at the minimum and the posterior variance $1/\Lambda$, as in Chapter 5.

Table 24.1 collects the three cases: each row is a case, and the columns give the drawings compared, what differs between them in the graph, the record and the quantities integrated out of each evidence.

24.1.4 The questions over the drawings (step 1)

Step 1 writes each engineering question over the drawings. Table 24.2 lists the five: the first column states each in the engineering language in which it is asked; the second the general form of its category in the notation of Chapter 2, with M a drawing and $p(\mathbf{y} \mid M)$ its evidence; the

Table 24.1: The three cases of the chapter: the drawings compared, the difference in the graph, the record and the quantities integrated out of each evidence.

Case	Drawings	Difference in the graph	Record	Integrated out
Loop	A: hall 2 on the shared supply; B: hall 2 on its own supply	B adds the node T_{s2} with a prior row, rewrites hall 2's row and the return row	six readings; record A ($y_{T_r} = 17.50$°C) or record B ($y_{T_2} = 14.80$, $y_{T_r} = 15.10$°C)	the state, five or six entries
Cooling	M_3: 3-out-of-4; M_2: 2-out-of-4; M_4: 4-out-of-4	the structure function of the hall	400 days of units down and losses; the softer record with $k_2 = 5$	ν_{o}; in the softer record also ν_{m}
Pump	M_1: rate free; M_0: rate pinned at $\lambda_{\mathrm{h}} = 0.001\,\mathrm{h}^{-1}$	a free entry ω against a pinned one, $\omega = 0$	$k = 4$ failures in 2000 h	λ under M_1

third its mathematical form for the case; the fourth its category; and the last the subsection where it is answered.

Table 24.2: The questions of the chapter, the general form of their category, their mathematical form, their category and where each is answered. $\mathbf{y}$ is the record of the case; M a drawing; U_{c} the unavailability of the hall's cooling; $U(a)$ the design objective of equation (2.16), not an unavailability; $\tilde{y}_a$ a candidate reading not yet installed; d' the drawing other than d.

Engineering question	General form	Mathematical form	Category	Answered in
Is hall 2 really on the shared loop?	$B_{AB} = p(\mathbf{y} \mid M_A)/p(\mathbf{y} \mid M_B)$	$p(\mathbf{y} \mid M_{\mathrm{A}})/p(\mathbf{y} \mid M_{\mathrm{B}})$ on records A and B	hypothesis comparison (structure)	24.3.1
Is it really parallel?	$B_{AB} = p(\mathbf{y} \mid M_A)/p(\mathbf{y} \mid M_B)$	$p(\mathbf{y} \mid M_3)/p(\mathbf{y} \mid M_2)$, the loss column given the units down	hypothesis comparison (structure)	24.3.1
Does N+1 hold: does the hall's cooling need three of its four units (N+1), or two (N+2)?				
Does the pump's record contradict the handbook rate?	$B_{AB} = p(\mathbf{y} \mid M_A)/p(\mathbf{y} \mid M_B)$	$p(\mathbf{y} \mid \lambda \text{ free})/p(\mathbf{y} \mid \lambda = \lambda_{\mathrm{h}})$	hypothesis comparison (a point null)	24.3.1
What availability is reported when the drawing is uncertain?	$\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]$, $\mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y})$	$\sum_d \Pr(M_d \mid \mathbf{y})\,\mathbb{E}[U_{\mathrm{c}} \mid M_d, \mathbf{y}]$ and its variance	estimation over an open structure	24.3.2
Which reading settles the loop question, before it is installed?	$U(a)$, the design form	$\sum_d \Pr(M_d \mid \mathbf{y})\,\mathbb{E}\big[\ln\{p(\tilde{y}_a \mid \mathbf{y}, M_d)/p(\tilde{y}_a \mid \mathbf{y}, M_{d'})\} \mid M_d\big]$	design	24.3.3

The comparisons. The first three questions are hypothesis comparisons. Their object is not a posterior inside either drawing but the ratio of the drawings' evidences. The second is the engineer's "is it really parallel?" in the form it takes for a redundant structure: the 2-out-of-4 drawing is the more parallel of the two, since it tolerates two units down, and the question is whether the hall really has that tolerance. It is also the question whether N+1 holds as designed: 3-out-of-4 is the N+1 arrangement of Chapter 13, and 2-out-of-4 would make the same four units N+2. The third compares a free rate with a pinned one, a drawing nested in the other at a point.

The mixture. The fourth question asks for a reported quantity when the drawing is not settled. By the law of total expectation over the drawing, the answer is the average of the answers under each drawing weighted by the drawings' posterior probabilities, with a variance that has a part within each drawing and a part between them, as Chapter 13 derived with weights that were priors.

The design. The fifth question is asked before the reading exists: which candidate reading the drawings of the loop predict differently enough to separate them. Its object is a preposterior, the evidence the reading is expected to add, averaged over its outcomes. How much each answer moves with the prior of the quantity a drawing frees is not a question of its own; it is the band of each comparison, measured in the checks (Section 24.4.4) and reported beside each answer. A further measurement for the questions once answered is the next measurement, Section 24.6.

24.1.5 The evidence as rows (step 2)

Step 2 attaches each record to each drawing as rows.

The modeling choices. Each case enters with the choices that define its drawings, and each choice buys something and can fail.

The loop. The family is written before the record is looked at, drawings A and B with their rows and priors: this is what makes the ratio a comparison of two stated descriptions of the plant, and it fails if the plant is a third drawing that neither describes, which the residual test of each drawing (Section 24.4.3) and Exercise 24.5 take up. Both drawings run on the reduced state of Chapter 17, the balances held exactly as the limit $\sigma_p \to 0$ of the full state's physics rows, with the reduced Jacobian $\mathbf{J} = \mathbf{J}_\mathrm{e}\mathbf{N}$ of equation (17.18), the rows of the evidence projected on the null directions $\mathbf{N}$ of the physics rows, and the covariance $(\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}$, the free-entry block of the full state's in that limit, equation (17.19): this lets each evidence be computed on five or six entries, and because an evidence keeps the constants that a posterior drops, Section 24.4.2 checks that it has the same limit. The reading errors are independent normals with the standard deviations of Table 17.1: this makes each evidence a near-Gaussian integral, and a biased sensor would break it, the subject of Chapter 22. Drawing B's new entry carries a proper prior, the second set point $\mathcal{N}(12^\circ\mathrm{C},\ 1\,\mathrm{K}^2)$, stated before the record: the ratio charges for that prior's width and would be meaningless under a flat one, and Section 24.4.4 measures how far the answer moves with the width. The route reported is the Laplace evidence of each fit, which a nonlinear plant admits where no closed form exists; it fails if the posterior is far from normal in some direction, and Section 24.4.1 sets importance sampling and a grid on the new entry beside it.

The cooling. The record is taken as what was observed, the two columns day by day, units down and losses: this keeps the comparison honest when the drawings group the days differently, since a binomial coefficient would cancel from the ratio only if both drawings saw the same groups. The units-down column has the same probability under every structure, since it depends on the units and not on what the hall needs: this lets it cancel from every ratio, leaving it to the unit unavailability (Section 24.2.6), and it fails if a common cause takes units down together with the cooling, which couples the two columns (Section 24.4.5). The days are independent given the drawing and the nuisance probability: this makes the loss column a product over days, and a run of losses from one lingering cause would break it (Section 24.6.5). The loss rule is deterministic at or above the threshold, with other causes below it at the rate $\nu_\mathrm{o} \sim \mathrm{Beta}(1, 99)$ and, on the softer record, misses at the rate $\nu_\mathrm{m} \sim \mathrm{Beta}(1, 19)$: the deterministic rule is what lets a record rule a drawing out, the nuisance priors set the price of each exception, and Section 24.4.4 varies them. The route reported is the Laplace evidence on the logit state; the exact Beta integral stands beside it, and Section 24.4.1 checks that by quadrature.

The pump. Failures arrive at a constant rate, as in Chapter 6: this is what makes the record a count in an exposure, and an aging pump would break it (Chapter 12). The free drawing's prior is the engineering prior Gamma(2, 1000 h), stated before the record: the factor charges for its width, and Section 24.4.4 varies its shape with its mean held. The handbook is a point, $\omega = 0$: this makes M_0 nested in M_1, and a handbook stated as an interval is a different drawing (Exercise 24.6). The route reported is the Laplace evidence in ω, which integrates the whole posterior and does not read the density at the pinned value, which may lie in a tail; the exact conjugate value stands beside it.

Under drawing A the loop has the eight rows of Chapter 17 on the reduced state, equation (17.15), and the return row: hall 2's balance is written against the shared supply, $T_2 = T_s + Q_2 \mathrm{e}^{-g_2}/G_{\mathrm{ref}}$, and the return header carries both halls' heat,

$$r^{\mathrm{A}}_{\mathrm{ret}}(\mathbf{z}) = \frac{T_s + (Q_1 + Q_2)/\dot{m}c_p - y_{T_r}}{\sigma_{T_r}}. \tag{24.2}$$

Under drawing B hall 2's reading row is written against the second supply, $(T_{s2} + Q_2 \mathrm{e}^{-g_2}/G_{\mathrm{ref}} - y_{T_2})/\sigma_T$, the return header carries hall 1's heat alone,

$$r^{\mathrm{B}}_{\mathrm{ret}}(\mathbf{z}) = \frac{T_s + Q_1/\dot{m}c_p - y_{T_r}}{\sigma_{T_r}}, \tag{24.3}$$

the chiller's heat meter still reads the heat of both halls, $P = Q_1 + Q_2$, and the prior of the second set point adds the row $(T_{s2} - 12^{\circ}\mathrm{C})/1\,\mathrm{K}$. Drawing A has nine rows (six readings and three priors) over five entries; drawing B has ten rows (the same six readings and four priors) over six entries. Both have four rows of redundancy. Neither drawing is nested in the other: drawing B with $T_{s2} = T_s$ is not drawing A, because its return row still omits hall 2's heat.

Which row acts on which entry is a property of each drawing's graph, and it fixes where the reduced Jacobian of the rows can be nonzero, marked $*$:

$$\mathbf{J}_{\mathrm{A}} = \begin{bmatrix} * & 0 & * & 0 & * \\ 0 & * & 0 & * & * \\ 0 & 0 & 0 & 0 & * \\ 0 & 0 & * & * & 0 \\ 0 & 0 & * & 0 & 0 \\ 0 & 0 & * & * & * \\ * & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * \end{bmatrix}, \qquad \mathbf{J}_{\mathrm{B}} = \begin{bmatrix} * & 0 & * & 0 & * & 0 \\ 0 & * & 0 & * & 0 & * \\ 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & * & * & 0 & 0 \\ 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & * & 0 & * & 0 \\ * & 0 & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & * \end{bmatrix} \begin{matrix} \text{hall-1 reading} \\ \text{hall-2 reading} \\ \text{supply reading} \\ \text{heat meter} \\ \text{load meter} \\ \text{return reading} \\ \text{prior on } g_1 \\ \text{prior on } g_2 \\ \text{prior on } T_s \\ \text{prior on } T_{s2} \end{matrix} \tag{24.4}$$

with columns $(g_1, g_2, Q_1, Q_2, T_s)$ and, under drawing B, T_{s2} last; drawing A has the first nine of the rows named. Two rows differ between the drawings. Hall 2's reading row moves from the column of T_s to that of T_{s2}, and the column of T_{s2} holds only that row and its own prior: the new entry is seen by one reading, through a row that g_2 and Q_2 also enter. The return row loses the column of Q_2: under drawing A it sees the total load and the supply temperature, under drawing B hall 1's load and the supply temperature. The other seven rows are the same in both patterns.

The cooling's rows are the loss column given the units down. Under a drawing with threshold m, the days below the threshold form one count row on ξ_{o}, a binomial row of $k_<$ losses in $n_<$ days carried on the logit as in Section 4.2.7, equation (4.43), with the slope $n_<\nu_{\mathrm{o}} - k_<$, the curvature $n_<\nu_{\mathrm{o}}(1 - \nu_{\mathrm{o}})$ and

$$n_< = \sum_{j<m} n_j, \qquad k_< = \sum_{j<m} k_j,$$

and the prior Beta(1, 99) is the prior row of one pseudo-loss in a hundred pseudo-days, equation (5.15); the days at or above the threshold, $n_{\geq}$ of them with $k_{\geq}$ losses, form no row on ξ_{o}, and on the softer record they are the count row of $n_{\geq} - k_{\geq}$ misses in $n_{\geq}$ days on ξ_{m}. The drawing decides which days enter which row: the same days are other-cause days under one structure and structural losses under another. The pump's row is the count row of four failures in 2000 h on ω, the count row of a rate of equation (6.22), with the slope $\mu - k$ and the curvature μ for $\mu = T_{\mathrm{tot}}\lambda$, beside the prior row of the engineering prior, two pseudo-failures in a thousand pseudo-hours; the handbook drawing has no entry for them to act on. None of the cooling's and the pump's rows is Gaussian, and none is a physics row.

24.2 The posterior of the state (step 3)

In the book's notation $p(\mathbf{z} \mid \mathbf{y}, M)$ is the posterior density of a drawing's state, $p(\mathbf{y} \mid M)$ its evidence, $f(\mathbf{z})$ a quantity derived from the state and $\tilde{y}_a$ a reading not yet taken.

All five are read from the drawings' posteriors and their evidences, one posterior per drawing with the evidence its normalizing constant. This section derives them and computes them for the case; Section 24.3 reads every answer from them.

The objects of each drawing's fit recur in the specialized forms. The *MAP state* $\hat{\mathbf{z}}_d$ of drawing d is the minimum of the objective that its rows put on its state; the *information matrix* $\mathbf{\Lambda}_d = \mathbf{J}_d^{\mathsf{T}}\mathbf{J}_d$ is the curvature of that objective at the minimum, with $\mathbf{J}_d$ the Jacobian of its normalized rows; the *Laplace covariance* $\mathbf{\Sigma}_d = \mathbf{\Lambda}_d^{-1}$ makes the posterior of its state approximately normal; and the evidence $p(\mathbf{y} \mid M_d)$, the normalizing constant of that posterior, is read from the same three objects together with the constants the objective omits. Beside them stand the closed forms that two of the cases admit: the evidence of a count or of a sequence of trials under a conjugate prior, the normalizer of the conjugate update.

24.2.1 The evidence of a drawing

Every question of Table 24.2 rests on the evidence of a drawing and the odds it produces. Within one drawing M, Bayes' theorem for the state reads

$$p(\mathbf{z} \mid \mathbf{y}, M) = \frac{p(\mathbf{y} \mid \mathbf{z}, M)\, p(\mathbf{z} \mid M)}{p(\mathbf{y} \mid M)}, \tag{24.5}$$

where every term carries the drawing in its conditioning: the rows of the likelihood and the priors are the drawing's own. The posterior integrates to one over the state, so integrating both sides and multiplying by the denominator gives

$$p(\mathbf{y} \mid M) = \int p(\mathbf{y} \mid \mathbf{z}, M)\, p(\mathbf{z} \mid M)\, \mathrm{d}\mathbf{z}, \tag{24.6}$$

the marginal likelihood of Chapter 3, equation (3.13), now with the drawing written in, and the evidence (2.26) of Chapter 2. It is the probability, or the probability density, that the drawing assigned to the record before the record was seen, with its unknowns averaged over their prior. Within one drawing it is a constant that the estimation, threshold and design questions of the earlier chapters could ignore, and the comparisons of Chapters 2, 7, 15, 22 and 23 could not; between drawings it is the only term that differs in kind.

Bayes' theorem between drawings. Treat the drawing as a discrete unknown d, with prior probabilities $\Pr(M_d)$ summing to one. Bayes' theorem for a discrete unknown, with the law of total probability in the denominator, is equation (2.46):

$$\Pr(M_d \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid M_d)\, \Pr(M_d)}{\sum_e p(\mathbf{y} \mid M_e)\, \Pr(M_e)}. \tag{24.7}$$

For two drawings, dividing the equation for M_1 by that for M_0 cancels the denominator, and taking logarithms,

$$\ln \frac{\Pr(M_1 \mid \mathbf{y})}{\Pr(M_0 \mid \mathbf{y})} = \ln B_{10} + \ln \frac{\Pr(M_1)}{\Pr(M_0)}, \qquad \ln B_{10} = \ln p(\mathbf{y} \mid M_1) - \ln p(\mathbf{y} \mid M_0) : \tag{24.8}$$

the log posterior odds are the log Bayes factor plus the log prior odds. Solving for the probability of M_1,

$$\Pr(M_1 \mid \mathbf{y}) = \frac{1}{1 + \exp\big[-\ln B_{10} - \ln\{\Pr(M_1)/\Pr(M_0)\}\big]}, \tag{24.9}$$

a logistic function of the log posterior odds. The Bayes factor is a property of the record and the two drawings; the posterior probability needs the prior odds as well.

The scale of the logarithm. The log Bayes factor is measured in nats. At even prior odds, by equation (24.9), $\ln B_{10} = 1$ is a factor $\mathrm{e} = 2.72$ and a posterior probability 0.731; $\ln B_{10} = 3$ is a factor 20.1 and a probability 0.953; $\ln B_{10} = 5$ is a factor 148 and a probability 0.9933. A conventional verbal scale, standard in the literature, calls a log Bayes factor below 1 not worth more than a mention, 1 to 3 positive, 3 to 5 strong and above 5 very strong. The scale is a reading aid; the decision belongs to the probabilities and to the consequences of acting on the wrong drawing, which a later part of the book weighs.

The evidence as a sequence of predictions. By the multiplication rule, the probability of a record of readings $y_1, \ldots, y_K$ is the product of each reading's probability given those before it, $p(\mathbf{y} \mid M) = \prod_{k=1}^{K} p(y_k \mid y_1, \ldots, y_{k-1}, M)$, so that

$$\ln B_{10} = \sum_{k=1}^{K} \ln \frac{p(y_k \mid y_1, \ldots, y_{k-1}, M_1)}{p(y_k \mid y_1, \ldots, y_{k-1}, M_0)} : \tag{24.10}$$

the log Bayes factor is the sum over the readings of how much better one drawing predicted each reading from the readings before it. The order is arbitrary and the sum is the same for every order. A reading that the two drawings predict with the same distribution adds nothing, however well both predict it, and the expected gain (24.33) of a candidate reading is the expectation of the last term of the sum.

24.2.2 The Laplace evidence and its constants

For a fitted plant with Gaussian rows the integral of equation (24.6) has no closed form, and the Laplace approximation supplies it, as Chapter 4 derived for one hall (Section 4.2.2). The constants that the objective $\mathcal{L}$ omits are what matters here, so they are written out. Let the drawing's rows be the reading rows $k \in \mathcal{R}$, each normalized with standard deviation σ_k; the proper prior rows $k \in \mathcal{P}$, Gaussian with standard deviation σ_k; and flat priors of constant density c_{f} on the entries that have no prior, such as the two loads. As Chapter 18 derived for the likelihood and the prior, equation (18.8), the product of the row densities is

$$p(\mathbf{y} \mid \mathbf{z}, M)\, p(\mathbf{z} \mid M) = c_{\mathrm{f}} \prod_{k \in \mathcal{R} \cup \mathcal{P}} \frac{1}{\sqrt{2\pi}\,\sigma_k} \exp\big[-\mathcal{L}(\mathbf{z})\big], \qquad \mathcal{L}(\mathbf{z}) = \tfrac{1}{2}\, r(\mathbf{z})^{\mathsf{T}} r(\mathbf{z}). \tag{24.11}$$

The fit is the Gauss–Newton iteration of Chapter 18, which solves the normal equations $(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^{\mathsf{T}} r$ from the prior point until the step vanishes and returns $\hat{\mathbf{z}}$ and $\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$ at it; with every row normalized the information matrix is $\mathbf{J}^{\mathsf{T}}\mathbf{J}$. Expanding $\mathcal{L}$ to second order about $\hat{\mathbf{z}}$, where its gradient vanishes, $\mathcal{L}(\mathbf{z}) \approx \mathcal{L}(\hat{\mathbf{z}}) + \frac{1}{2}(\mathbf{z} - \hat{\mathbf{z}})^{\mathsf{T}}\mathbf{H}(\mathbf{z} - \hat{\mathbf{z}})$, with the Hessian $\mathbf{H} = \mathbf{\Lambda} + \sum_k r_k\, \partial^2 r_k / \partial \mathbf{z}^2$, the information plus the curvature term of the nonlinear rows; with

the Gauss–Newton $\mathbf{\Lambda}$ in its place and $\mathbf{\Sigma} = \mathbf{\Lambda}^{-1}$, the Gaussian integral of equation (19.9) gives $\int \exp[-\mathcal{L}(\mathbf{z})]\,\mathrm{d}\mathbf{z} \approx \exp[-\mathcal{L}(\hat{\mathbf{z}})]\,(2\pi)^{n/2}\det(\mathbf{\Sigma})^{1/2}$ for a state of n entries. The logarithm of the evidence is then

$$\ln p(\mathbf{y} \mid M) \approx -\mathcal{L}(\hat{\mathbf{z}}) + \tfrac{1}{2}\ln\det\mathbf{\Sigma} + \tfrac{n}{2}\ln 2\pi - \sum_{k\in\mathcal{R}} \ln\bigl(\sqrt{2\pi}\,\sigma_k\bigr) - \sum_{k\in\mathcal{P}} \ln\bigl(\sqrt{2\pi}\,\sigma_k\bigr) + \ln c_{\mathrm{f}}. \quad (24.12)$$

It is exact when every row is linear in the state, so that $\mathcal{L}$ is quadratic. The curvature term is small when the residuals at $\hat{\mathbf{z}}$ are, as Chapter 19 derived; Section 24.4.1 measures it for the loop. The first three terms are what the fit returns. The rest are the constants of the drawing, and the comparison of two drawings turns on which of them cancel.

What cancels between two drawings. Four rules follow from equation (24.12); Chapters 22 and 23 applied them to their bias and fault drawings, and they are collected here in general.

1. *Reading rows.* Two drawings compared on the same readings with the same error standard deviations have the same sum over $\mathcal{R}$, which cancels. A comparison of evidences computed on different readings is not a comparison of drawings, and a drawing may not drop an awkward reading.
2. *Shared prior rows.* A prior row present in both drawings with the same σ_k cancels. A prior widened in one drawing, as for the fouled path of Chapter 23, leaves the difference $-\ln(\sigma_{\mathrm{wide}}/\sigma_{\mathrm{old}})$ in its normalizer, which the change in $\det\mathbf{\Sigma}$ partly returns.
3. *An added entry.* An entry freed in one drawing adds $\tfrac{1}{2}\ln 2\pi$ through n and $-\ln(\sqrt{2\pi}\,\sigma_0)$ through its prior row; the $\ln 2\pi$ cancels and leaves $-\ln\sigma_0$, which together with the added dimension of $\det\mathbf{\Sigma}$ is the Occam factor of the entry.
4. *Flat priors.* The constant c_{f} is arbitrary (it depends on the units of the entry) and cancels only when both drawings carry the same flat priors on the same entries. A quantity freed by one drawing alone must have a proper prior (Section 24.2.3).

For the loop the six reading rows and the three prior rows on g_1, g_2 and T_s are common, the flat priors on Q_1 and Q_2 are common, and drawing B adds the entry T_{s2} with its prior row of $\sigma_0 = 1\,\mathrm{K}$.

The added entry. Let drawing B be drawing A's entries $\mathbf{z}_{\mathrm{A}}$ plus one entry z_{n} with prior standard deviation σ_0, and order $\mathbf{\Lambda}_{\mathrm{B}}$ with the old entries first. By the determinant of a partitioned matrix, $\det\mathbf{\Lambda}_{\mathrm{B}} = \det\mathbf{\Lambda}_{\mathrm{B}}^{\mathrm{old}} \times s$, where $\mathbf{\Lambda}_{\mathrm{B}}^{\mathrm{old}}$ is the block of the old entries and s the Schur complement of that block, which by the block inversion of Section 3.5.3 is the inverse of the marginal posterior variance Σ_{nn} of the new entry, the last diagonal entry of $\mathbf{\Sigma}_{\mathrm{B}}$. Hence $\tfrac{1}{2}\ln\det\mathbf{\Sigma}_{\mathrm{B}} = -\tfrac{1}{2}\ln\det\mathbf{\Lambda}_{\mathrm{B}}^{\mathrm{old}} + \tfrac{1}{2}\ln\Sigma_{\mathrm{nn}}$. Subtracting equation (24.12) for drawing B from the same equation for drawing A, with the common constants cancelled and the new prior row's $-\ln\sigma_0$ kept,

$$\ln B_{AB} \approx \bigl[\mathcal{L}_{\mathrm{B}}(\hat{\mathbf{z}}_{\mathrm{B}}) - \mathcal{L}_{\mathrm{A}}(\hat{\mathbf{z}}_{\mathrm{A}})\bigr] + \tfrac{1}{2}\ln\frac{\det\mathbf{\Lambda}_{\mathrm{B}}^{\mathrm{old}}}{\det\mathbf{\Lambda}_{\mathrm{A}}} + \ln\frac{\sigma_0}{\sqrt{\Sigma_{\mathrm{nn}}}},$$

which is equation (24.25) with $z_{\mathrm{n}} = T_{s2}$. The last term is positive, in favor of the drawing without the entry, by as much as the record narrows the entry's prior; an entry the record does not inform costs almost nothing, and buys almost nothing.

24.2.3 The Occam factor

The simplest comparison shows why the evidence charges a freed quantity, and it is the one-entry case of equation (24.12). Let one quantity z be read once, $y = z + \varepsilon$ with $\varepsilon \sim \mathcal{N}(0, \sigma^2)$. Drawing

M_0 pins z at a value z^{h}, and its evidence is the density of the reading, $p(\mathbf{y} \mid M_0) = \mathcal{N}(y;\ z^{\mathrm{h}},\ \sigma^2)$. Drawing M_1 frees it with the prior $\mathcal{N}(z^0, \sigma_0^2)$: its state is the one number z, and its two rows, the reading row and the prior row, are Gaussian, with residuals and slopes that are numbers,

$$r_1 = \frac{z - y}{\sigma}, \quad j_1 = \frac{1}{\sigma}, \qquad r_0 = \frac{z - z^0}{\sigma_0}, \quad j_0 = \frac{1}{\sigma_0}, \qquad \mathcal{L}(z) = \tfrac{1}{2}(r_1^2 + r_0^2).$$

Both rows are linear, so the curvature $\mathcal{L}'' = j_1^2 + j_0^2$ is the same at every z and the Newton step $\Delta z = -\mathcal{L}'/\mathcal{L}''$ reaches the minimum in one step from z^0, where the slope is $\mathcal{L}' = j_1 r_1 + j_0 r_0 = (z^0 - y)/\sigma^2$. The information and the variance are constants:

$$\Lambda = j_1^2 + j_0^2 = \frac{1}{\sigma^2} + \frac{1}{\sigma_0^2}, \qquad \sigma_z^2 = \frac{1}{\Lambda}, \qquad \hat{z} = z^0 + \sigma_z^2 \, \frac{y - z^0}{\sigma^2} = \sigma_z^2 \left(\frac{y}{\sigma^2} + \frac{z^0}{\sigma_0^2} \right), \tag{24.13}$$

the posterior's precision and mean, as in Chapter 4. The objective is quadratic, so equation (24.12) with $n = 1$, the determinant reduced to the variance σ_z^2, one reading row and one prior row is exact:

$$\ln p(\mathbf{y} \mid M_1) = -\frac{(\hat{z} - y)^2}{2\sigma^2} - \frac{(\hat{z} - z^0)^2}{2\sigma_0^2} + \tfrac{1}{2} \ln \sigma_z^2 + \tfrac{1}{2} \ln 2\pi - \ln(\sqrt{2\pi}\,\sigma) - \ln(\sqrt{2\pi}\,\sigma_0).$$

Since $\sigma_z^2 = \sigma^2 \sigma_0^2/(\sigma^2 + \sigma_0^2)$, the minimized objective is $(y - z^0)^2/(2(\sigma^2 + \sigma_0^2))$ and $\sigma\sigma_0/\sigma_z = \sqrt{\sigma^2 + \sigma_0^2}$, so that

$$p(\mathbf{y} \mid M_1) = \mathcal{N}(y;\ z^0,\ \sigma^2 + \sigma_0^2)\ : \tag{24.14}$$

under the free drawing the reading is predicted with the spread of the error and the prior together. The same evidence, its terms regrouped around the MAP state, is

$$p(\mathbf{y} \mid M_1) = \underbrace{p(y \mid \hat{z})}_{\text{best fit}} \times \underbrace{\frac{\sigma_z}{\sigma_0}}_{\text{Occam factor}} \times \underbrace{\exp\left[-\frac{(\hat{z} - z^0)^2}{2\sigma_0^2} \right]}_{\text{prior misfit}}. \tag{24.15}$$

The best-fit likelihood is what a comparison of fits would look at. The Occam factor, never above one, is the fraction of the prior's width that the record left standing, the posterior spread of the entry over its prior spread: a prior wider than the record needs is charged in proportion. The misfit factor charges a prior placed away from where the record put the quantity. For an entry added to a larger drawing the same factor is $\sqrt{\Sigma_{\mathrm{nn}}}/\sigma_0$, the marginal posterior spread of the entry over its prior spread, the last term of equation (24.25).

The limit of a flat prior. As $\sigma_0 \to \infty$ the Occam factor $\sigma_z/\sigma_0 \to 0$ while the best fit and the misfit factor tend to finite limits, so $p(\mathbf{y} \mid M_1) \to 0$ and $\ln B_{10} \to -\infty$ whatever the reading: a freed quantity with an improper flat prior loses to any pinned value. This is Lindley's paradox in its simplest form. Its practical rule is that every quantity one drawing frees and the other does not must carry a proper prior, stated before the record is looked at, and that the Bayes factor must be reported with that prior.

24.2.4 Conjugate rows in closed form

When the prior and the rows are conjugate, the integral of equation (24.6) is a closed form, the normalizer of the conjugate update, and the Laplace evidence of the corresponding fit in the logit or the log-rate is its approximation. Two families cover the records of the book. For a Poisson count k in an exposure T_{tot} with the prior $\lambda \sim \mathrm{Gamma}(a, b)$, Chapter 6 evaluated the evidence, equation (6.68), the negative-binomial probability of k:

$$p(\mathbf{y} \mid M_1) = \frac{\Gamma(a + k)}{\Gamma(a)\, k!} \left(\frac{b}{b + T_{\mathrm{tot}}} \right)^a \left(\frac{T_{\mathrm{tot}}}{b + T_{\mathrm{tot}}} \right)^k. \tag{24.16}$$

For a probability p with the prior $\mathrm{Beta}(a, b)$ and a sequence of n trials of which k are failures, the likelihood of the sequence is $p^k(1-p)^{n-k}$, and the Beta integral of Chapter 5, equation (5.13), gives

$$\int_0^1 p^k(1-p)^{n-k}\,\frac{p^{a-1}(1-p)^{b-1}}{\mathrm{B}(a,b)}\,\mathrm{d}p = \frac{\mathrm{B}(a+k,\ b+n-k)}{\mathrm{B}(a,b)}. \tag{24.17}$$

For the count of failures rather than the sequence the binomial coefficient $\binom{n}{k}$ multiplies the result; it cancels from a ratio of two drawings that see the same n and k, and not from one whose drawings split the record into different groups of trials, which is why the cooling's record is taken as its sequence of days.

The evidence of the cooling's drawings. By the multiplication rule the probability of the record under a drawing M factorizes,

$$p(\mathbf{y}_{\mathrm{down}}, \mathbf{y}_{\mathrm{loss}} \mid M) = p(\mathbf{y}_{\mathrm{down}})\; p(\mathbf{y}_{\mathrm{loss}} \mid \mathbf{y}_{\mathrm{down}}, M), \tag{24.18}$$

and the first factor, the same under every drawing, cancels. The second is the loss column given the units down. Given ν_{o}, the days below the threshold contribute $\nu_{\mathrm{o}}^{k_<}(1-\nu_{\mathrm{o}})^{n_<-k_<}$, the likelihood of the count row on ξ_{o}, and the days at or above it contribute one if every one of them was lost and zero otherwise. The prior row and the count row are those of Chapter 5, whose exact posterior is the Beta update (5.24), $\nu_{\mathrm{o}} \mid \mathbf{y}, M \sim \mathrm{Beta}(a_{\mathrm{o}} + k_<,\ b_{\mathrm{o}} + n_< - k_<)$, and the evidence is the normalizer of that update. Integrating ν_{o} over its prior $\mathrm{Beta}(a_{\mathrm{o}}, b_{\mathrm{o}})$ with the Beta function of equation (5.13),

$$p(\mathbf{y}_{\mathrm{loss}} \mid \mathbf{y}_{\mathrm{down}}, M) = \frac{\mathrm{B}(a_{\mathrm{o}} + k_<,\ b_{\mathrm{o}} + n_< - k_<)}{\mathrm{B}(a_{\mathrm{o}}, b_{\mathrm{o}})}\;\mathbf{1}\{k_{\geq} = n_{\geq}\}. \tag{24.19}$$

With the miss probability of the softer record, a day at or above the threshold is lost with probability $1-\nu_{\mathrm{m}}$, the indicator becomes $\nu_{\mathrm{m}}^{n_{\geq}-k_{\geq}}(1-\nu_{\mathrm{m}})^{k_{\geq}}$, the likelihood of the count row on ξ_{m}, and its integral over $\mathrm{Beta}(a_{\mathrm{m}}, b_{\mathrm{m}})$ gives

$$p(\mathbf{y}_{\mathrm{loss}} \mid \mathbf{y}_{\mathrm{down}}, M) = \frac{\mathrm{B}(a_{\mathrm{o}} + k_<,\ b_{\mathrm{o}} + n_< - k_<)}{\mathrm{B}(a_{\mathrm{o}}, b_{\mathrm{o}})}\;\frac{\mathrm{B}(a_{\mathrm{m}} + n_{\geq} - k_{\geq},\ b_{\mathrm{m}} + k_{\geq})}{\mathrm{B}(a_{\mathrm{m}}, b_{\mathrm{m}})}. \tag{24.20}$$

The two nuisance probabilities are independent a priori and enter separate groups of days, so their integrals separate, and the two entries ξ_{o} and ξ_{m} have disjoint columns. Both equations are exact. Each loss that a drawing must attribute to other causes multiplies its evidence by about the other-cause probability, a cost of about $\ln(1/\hat{\nu}_{\mathrm{o}})$ nats, and one day at or above the threshold that kept its cooling sets the evidence of a deterministic drawing to zero. On the state ξ_{o}, one number, the same two terms are the count row of $k_<$ losses in $n_<$ days and the prior row of a_{o} pseudo-losses in $a_{\mathrm{o}} + b_{\mathrm{o}}$ pseudo-days, both binomial rows in the logit (Section 4.2.7, equation (4.43)), whose curvatures add to the information

$$\mathcal{L}_c'' = n_<\,\nu_{\mathrm{o}}(1-\nu_{\mathrm{o}}), \qquad \mathcal{L}_0'' = (a_{\mathrm{o}}+b_{\mathrm{o}})\,\nu_{\mathrm{o}}(1-\nu_{\mathrm{o}}), \qquad \Lambda = \mathcal{L}_c'' + \mathcal{L}_0'' = (a_{\mathrm{o}}+b_{\mathrm{o}}+n_<)\,\nu_{\mathrm{o}}(1-\nu_{\mathrm{o}}),$$

at the minimum $\hat{\nu}_{\mathrm{o}} = (a_{\mathrm{o}}+k_<)/(a_{\mathrm{o}}+b_{\mathrm{o}}+n_<)$; the Laplace evidence in the logit, $(a_{\mathrm{o}}+k_<)\ln\hat{\nu}_{\mathrm{o}} + (b_{\mathrm{o}}+n_<-k_<)\ln(1-\hat{\nu}_{\mathrm{o}}) - \ln\mathrm{B}(a_{\mathrm{o}}, b_{\mathrm{o}}) + \frac{1}{2}\ln(2\pi/\Lambda)$, is the other route, set beside equation (24.19) in Section 24.4.1. Two letters carry more than one meaning here and are kept apart by their form: an italic B with subscripts is a Bayes factor, the upright $\mathrm{B}(\cdot,\cdot)$ is the Beta function, and the letters A and B in roman type name drawings; n without a subscript is the number of entries of a state, and with one (n_j, $n_<$) a number of days, of which k with the same subscript counts the losses. The loss-record evidences (24.19) and (24.20) are this ratio applied to each group of days, and the pump's factor (24.28) is equation (24.16) divided by the Poisson probability at the pinned rate. Both are exact; no fit and no approximation is involved.

24.2.5 A nested point by the Savage–Dickey ratio

Let drawing M_1 carry an entry z_j together with the other entries $\mathbf{z}_{-j}$, and let M_0 be M_1 with z_j pinned at a value z_j^{h}. Suppose the prior of the other entries under M_0 is their conditional prior under M_1 at that value,

$$p(\mathbf{z}_{-j} \mid M_0) = p(\mathbf{z}_{-j} \mid z_j = z_j^{\mathrm{h}}, M_1), \tag{24.21}$$

which holds in particular when the priors of the entries are independent and M_0 keeps the others' priors unchanged. The evidence of M_0 is then

$$p(\mathbf{y} \mid M_0) = \int p(\mathbf{y} \mid z_j^{\mathrm{h}}, \mathbf{z}_{-j})\, p(\mathbf{z}_{-j} \mid z_j^{\mathrm{h}}, M_1)\, \mathrm{d}\mathbf{z}_{-j} = p(\mathbf{y} \mid z_j = z_j^{\mathrm{h}}, M_1),$$

the likelihood of z_j under M_1 at the pinned value, with the other entries integrated out. Bayes' theorem for the entry z_j under M_1, evaluated at z_j^{h}, reads

$$p(z_j^{\mathrm{h}} \mid \mathbf{y}, M_1) = \frac{p(\mathbf{y} \mid z_j = z_j^{\mathrm{h}}, M_1)\, p(z_j^{\mathrm{h}} \mid M_1)}{p(\mathbf{y} \mid M_1)}.$$

Replacing the likelihood by $p(\mathbf{y} \mid M_0)$ and dividing,

$$B_{01} = \frac{p(\mathbf{y} \mid M_0)}{p(\mathbf{y} \mid M_1)} = \frac{p(z_j^{\mathrm{h}} \mid \mathbf{y}, M_1)}{p(z_j^{\mathrm{h}} \mid M_1)}, \qquad \ln B_{10} = \ln p(z_j^{\mathrm{h}} \mid M_1) - \ln p(z_j^{\mathrm{h}} \mid \mathbf{y}, M_1)\,: \tag{24.22}$$

the Bayes factor in favor of the pinned value is the posterior density of the entry at that value divided by its prior density there, both under the free drawing. Chapter 7 used the one-entry case, equation (7.68), with the flat prior on the log-scale common to both models. The ratio needs one fit, of the free drawing, and the marginal posterior density of one entry at one point; written in another coordinate of the entry, both densities take the same change-of-variable factor at the pinned value, and the ratio is unchanged. It needs a proper prior on z_j, whose density at z_j^{h} must be a finite number, and it needs the posterior density at the pinned value, often in a tail, where a Laplace normal can be badly wrong, as Chapter 7 showed where the posterior is skewed.

Table 24.3 collects the three computations of the evidence: for each, the drawings it applies to, its formula and the case of the chapter that uses it.

Table 24.3: The three computations of the evidence.

Computation	Applies to	Formula	Used for
Closed form	conjugate rows: counts with a gamma prior, outcomes with a Beta prior	negative binomial (24.16); Beta ratio (24.17), (24.19), (24.20)	the exact values beside the cooling and the pump
Savage–Dickey ratio	a drawing nested in another at a point, with condition (24.21)	$B_{01} = p(z_j^{\mathrm{h}} \mid \mathbf{y}, M_1)/p(z_j^{\mathrm{h}} \mid M_1)$, equation (24.22)	the exact value beside the pump
Laplace evidence	a fitted plant with near-Gaussian posterior	$-\mathcal{L}(\hat{\mathbf{z}}) + \frac{1}{2}\ln\det\boldsymbol{\Sigma} + \frac{n}{2}\ln 2\pi$ plus the row constants, equation (24.12); for an added entry, equation (24.25)	every reported comparison

24.2.6 The posterior of the drawings

The inputs are the records of Section 24.1.3: the six readings of each loop record with the priors of Table 17.2 and the second set point; the loss record and its softer variant with the

nuisance priors Beta(1, 99) and Beta(1, 19); the pump's four failures in 2000 h with the prior Gamma(2, 1000 h); and Chapter 13's fleet posterior Beta(3, 97) of the unit unavailability.

Each loop drawing is fitted to each record by the Gauss–Newton iteration of Chapter 18, and its evidence is evaluated by equation (24.12). Table 24.4 lists the four fits: each column is one drawing on one record, and the rows give the fitted conductances, loads and supply temperatures, the return temperature the fit predicts against the reading, the largest normalized residual and the row it belongs to, the objective with its tail probability, the log evidence and the log Bayes factor of drawing A against drawing B. The log evidences include every reading and prior constant, with the flat densities of the two loads taken as one per kW; only their differences carry meaning.

Table 24.4: The fits of the two drawings of the loop to the two records. Tail: $\Pr(\chi^2_4 > 2\mathcal{L}(\hat{\mathbf{z}}))$, four degrees of freedom for both drawings (nine rows over five entries, ten over six). Log evidence by equation (24.12) with the Gauss–Newton $\boldsymbol{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$. The probability of drawing A is at even prior odds.

	Record A ($y_{T_r} = 17.50$°C)		Record B ($y_{T_r} = 15.10$°C)	
Quantity	drawing A	drawing B	drawing A	drawing B
$\hat{G}_1$ [kW/K]	5.53	6.94	4.83	5.48
$\hat{G}_2$ [kW/K]	4.11	3.93	4.87	5.99
$\hat{Q}_1$, $\hat{Q}_2$ [kW]	30.70, 22.39	33.25, 19.79	30.99, 17.32	31.04, 21.08
$\hat{T}_s$ [°C]	12.15	13.06	11.23	12.04
$\hat{T}_{s2} \pm$ sd [°C]	–	12.53 ± 0.86	–	11.43 ± 0.78
Predicted T_r [°C]	17.46	16.38	16.06	15.14
$r_k(\hat{\mathbf{z}})$ largest in size [–]	-0.66, g_2 prior	-3.73, return	$+3.19$, return	$+0.60$, g_2 prior
$2\mathcal{L}(\hat{\mathbf{z}})$ [–]	0.79	29.85	23.67	0.98
Tail [–]	0.94	5.3×10^{-6}	9.3×10^{-5}	0.91
$\ln p(\mathbf{y} \mid M)$ [–]	-4.50	-18.94	-15.72	-4.44
$\ln B_{AB}$ [nat]	$+14.44$		-11.28	
$\Pr(M_{\mathrm{A}} \mid \mathbf{y})$ [–]	$1 - 5.3 \times 10^{-7}$		1.3×10^{-5}	

On record A drawing A fits with $2\mathcal{L} = 0.79$ and every normalized residual below one, while drawing B moves the supply temperature to 13.06°C, 3.2 standard deviations above its reading, and the hall-1 load to 33.25 kW, 1.5 standard deviations (2.25 kW) above its meter, and still predicts a return 3.7 standard deviations below the reading. On record B drawing B fits with $2\mathcal{L} = 0.98$, while drawing A lowers the supply temperature to 11.23°C, 2.9 standard deviations below its reading, and the total load to 48.3 kW, 2.1 standard deviations (4.2 kW) below the heat meter, and still predicts a return 3.2 standard deviations above the reading.

The unit unavailability. The unit unavailability q shared by the four units is read from the whole record: Chapter 13's fleet posterior Beta(3, 97), three units down in a hundred checks, is updated by the loss record's units-down column, 57 units down in 1600 unit-checks, through the Beta update (5.24), to

$$q \mid \mathbf{y} \sim \mathrm{Beta}(3 + 57,\ 97 + 1600 - 57) = \mathrm{Beta}(60,\ 1640). \tag{24.23}$$

This uses the column that the factorization (24.18) cancels from the structure comparison: it is the same under every drawing, so it leaves the weights unchanged, and it is the larger part of the record on q, so a report that left it out would rest on less than the record that sets the weights. It treats the 1600 unit-checks as independent trials with one shared unavailability, which fails if units go down together; the same column tests that (Section 24.4.5). On the state, one number, the logit $\xi_q = \ln(q/(1-q))$, the same units-down counts and the fleet

posterior's pseudo-counts are the count row and the prior row of Chapter 5, and the fit stops, by equations (5.20) and (5.21), at

$$\hat{\xi}_q = \ln \frac{60}{1640} = -3.3081, \qquad \Lambda = \frac{60 \times 1640}{1700} = 57.88, \qquad \sigma_q = \frac{1}{\sqrt{\Lambda}} = 0.13144, \tag{24.24}$$

the Laplace normal $\xi_q \mid \mathbf{y} \approx \mathcal{N}(-3.3081,\ 0.13144^2)$; the fleet posterior Beta$(3, 97)$ alone has $\hat{\xi}_q = -3.4761$ and $\sigma_q = 0.5862$.

The cooling and the pump. The cooling needs no iteration: under each drawing the posterior of the other-cause probability is the Beta update of its count row, Beta$(4, 488)$ under 3-out-of-4, of mean 0.0081, and Beta$(11, 488)$ under 2-out-of-4, of mean 0.022; on the logit, $\hat{\xi}_{\mathrm{o}} = -4.804$ with $\Lambda = 3.967$ and $\sigma_{\xi_{\mathrm{o}}} = 0.502$, and $\hat{\xi}_{\mathrm{o}} = -3.792$ with $\Lambda = 10.758$ and $\sigma_{\xi_{\mathrm{o}}} = 0.305$. The unit unavailability's posterior (24.23), Beta$(60, 1640)$, has mean 0.0353. The pump's fit is that of Chapter 6: $\hat{\omega} = \ln 2 = 0.6931$, $\Lambda = 6$, $\sigma_\omega^2 = 1/6$, $\sigma_\omega = 0.4082$, with the exact posterior Gamma$(6,\ 3000\,\mathrm{h})$.

These fits, their $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$ or their one entry and its variance, are what step 4 uses: every answer of the chapter is computed from them, a Laplace evidence for each comparison and the moments of the state for each estimate. Where a closed form or an importance estimate exists, it gives the value beside the answer, with the gap.

24.3 The answers (step 4)

The five questions of Table 24.2 fall into three categories of Chapter 2: hypothesis comparison (the first three), estimation over a structure that stays open (the fourth) and design (the fifth). Each is answered from the evidences and the posteriors of Section 24.2: its engineering question, the general form of its category, the form it takes for the case, the quantities it uses, and the case's numbers. The questions name the handbook rate $\lambda_{\mathrm{h}} = 10^{-3}\,\mathrm{h}^{-1}$ and the design file's prior odds $0.7 : 0.3$. Every answer is computed from the fits of Section 24.2.6, by the Laplace evidence of each drawing or the moments of the state, and that is the value the chapter reports; the closed form or the importance estimate gives the value set beside it, with the gap.

24.3.1 The comparisons of drawings

Engineering questions. Is hall 2 really on the shared loop? Is it really parallel: does N+1 hold, does the hall's cooling need three of its four units, or two? Does the pump's record contradict the handbook rate? Each asks which of two descriptions of the plant the record supports, and how strongly.

General form. Hypothesis comparison, equation (2.14), the ratio of the evidences of two drawings fitted to the same record, which multiplies their prior odds into their posterior odds, equation (2.47):

$$B_{AB} = \frac{p(\mathbf{y} \mid M_A)}{p(\mathbf{y} \mid M_B)}, \qquad \frac{\Pr(M_A \mid \mathbf{y})}{\Pr(M_B \mid \mathbf{y})} = B_{AB}\,\frac{\Pr(M_A)}{\Pr(M_B)},$$

with each evidence the integral (2.26) of the drawing's likelihood over its prior.

Specialized form. The three questions are equation (2.14) with M_A and M_B the two drawings of the case and each evidence the Laplace evidence (24.12) of the drawing's fit, the height of the integrand at $\hat{\mathbf{z}}$ times the posterior volume. Each case enters with the drawings and the choices of Section 24.1.5.

The loop. Each drawing is fitted by the Gauss–Newton iteration of Chapter 18 on its own rows, nine over five entries under drawing A and ten over six under drawing B, which returns $\hat{\mathbf{z}}_d$, $\mathbf{\Lambda}_d = \mathbf{J}_d^{\mathsf{T}}\mathbf{J}_d$ and $\mathbf{\Sigma}_d$. Its evidence is the height of the integrand at the MAP state, $\mathrm{e}^{-\mathcal{L}_d(\hat{\mathbf{z}}_d)}$ times the normalizing constants of its rows, multiplied by the posterior volume $(2\pi)^{n_d/2}\det(\mathbf{\Sigma}_d)^{1/2}$, with n_d the number of entries (Section 24.2.2). In the ratio the constants of the six reading rows and of the three shared prior rows cancel, and what is left is

$$\ln B_{AB} \approx \left[\mathcal{L}_{\mathrm{B}}(\hat{\mathbf{z}}_{\mathrm{B}}) - \mathcal{L}_{\mathrm{A}}(\hat{\mathbf{z}}_{\mathrm{A}})\right] + \tfrac{1}{2}\ln\frac{\det \mathbf{\Lambda}_{\mathrm{B}}^{\mathrm{old}}}{\det \mathbf{\Lambda}_{\mathrm{A}}} + \ln\frac{\sigma_0}{\sqrt{\Sigma_{T_{s2}T_{s2}}}}, \tag{24.25}$$

with $\mathbf{\Lambda}_{\mathrm{B}}^{\mathrm{old}}$ the 5×5 block of $\mathbf{\Lambda}_{\mathrm{B}}$ on the entries the two drawings share, $\sigma_0 = 1\,\mathrm{K}$ the prior standard deviation of T_{s2} and $\Sigma_{T_{s2}T_{s2}}$ the last diagonal entry of $\mathbf{\Sigma}_{\mathrm{B}}$, the marginal posterior variance of the new entry. The three terms are the difference of the misfits, the change of the curvature of the shared entries, and the inverse of the Occam factor of the freed set point, positive and in favor of the drawing without it by as much as the record narrows its prior. The posterior probability of drawing A follows from the odds, equation (2.47). The misfit term is read from the return row's column pattern: a drawing whose return row counts the wrong heat can meet the return reading only through the entries that row shares with others, the supply temperature and the loads, so it moves them away from their own readings and still leaves part of the disagreement in the return row; its objective then carries the disagreement in several rows at once, and the difference of the misfits is the term that decides the ratio.

The cooling. The state is the one number ξ_{o}, with ξ_{m} added on the softer record, and the factor B_{32} is the ratio of the two drawings' Laplace evidences in the logit,

$$\ln B_{32} \approx \Big[(a_{\mathrm{o}} + k_<)\ln\hat{\nu}_{\mathrm{o}} + (b_{\mathrm{o}} + n_< - k_<)\ln(1-\hat{\nu}_{\mathrm{o}}) + \tfrac{1}{2}\ln\frac{2\pi}{\Lambda}\Big]_{M_3} - \Big[\cdots\Big]_{M_2}, \tag{24.26}$$

the term $-\ln \mathrm{B}(a_{\mathrm{o}}, b_{\mathrm{o}})$ cancelling; on the softer record each drawing has the same bracket again for ξ_{m}, the two entries having disjoint columns. The value beside it comes from the exact evidences of the loss column, equation (24.19), or equation (24.20) on the softer record. The units-down column cancels by the factorization (24.18).

The pump. The free drawing's state is the one number ω, with the count row and the prior row of step 2. Its fit is that of Chapter 6: the minimum $\hat{\omega} = \ln\big((a+k)/((b+T_{\mathrm{tot}})\lambda_{\mathrm{ref}})\big)$, the information $\Lambda = \mathcal{L}''(\hat{\omega}) = a + k$, the sum of the two rows' curvatures, their expected counts at the minimum, and the variance $\sigma_\omega^2 = 1/(a+k)$. Its Laplace evidence is

$$\ln p(\mathbf{y} \mid M_1) \approx \ln p(k \mid \hat{\omega}) + \ln p(\hat{\omega}) + \tfrac{1}{2}\ln(2\pi\sigma_\omega^2), \tag{24.27}$$

and that of M_0, with nothing to integrate, the Poisson probability at the pinned rate. The value beside it is exact: the evidence of M_1 is the negative-binomial probability of k, equation (6.68), and that of M_0, with nothing to integrate, the Poisson probability $(\lambda_{\mathrm{h}}T_{\mathrm{tot}})^k \mathrm{e}^{-\lambda_{\mathrm{h}}T_{\mathrm{tot}}}/k!$; in the ratio the factorial and the power T_{tot}^k cancel,

$$B_{10} = \frac{\Gamma(a+k)}{\Gamma(a)}\,\frac{b^a}{(b+T_{\mathrm{tot}})^{a+k}}\,\frac{\mathrm{e}^{\lambda_{\mathrm{h}}T_{\mathrm{tot}}}}{\lambda_{\mathrm{h}}^k}. \tag{24.28}$$

Because M_0 is M_1 with the entry pinned, the same factor is a ratio of two densities of the free drawing at the pinned value, the Savage–Dickey ratio (Section 24.2.5): with the gamma densities $\mathrm{Gamma}(\lambda;\ \alpha,\beta) = \beta^\alpha\lambda^{\alpha-1}\mathrm{e}^{-\beta\lambda}/\Gamma(\alpha)$,

$$B_{01} = \frac{p(\lambda_{\mathrm{h}} \mid \mathbf{y}, M_1)}{p(\lambda_{\mathrm{h}} \mid M_1)} = \frac{\mathrm{Gamma}(\lambda_{\mathrm{h}};\ a+k,\ b+T_{\mathrm{tot}})}{\mathrm{Gamma}(\lambda_{\mathrm{h}};\ a,\ b)}. \tag{24.29}$$

Written in ω the two densities are each multiplied by the change-of-variable factor $\mathrm{d}\lambda/\mathrm{d}\omega = \lambda_{\mathrm{h}}$ at $\omega = 0$, which cancels from the ratio. The two directions of the comparison differ in kind.

Evidence against the handbook grows without bound as the count rises, because the handbook's Poisson probability of many failures falls faster than any power. Evidence for it is capped: by equation (24.29), B_{01} is the posterior density at λ_{h} over the prior density there, and a record of a given exposure can raise the posterior density at λ_{h} only as far as it narrows the posterior around it. Support for a point value is bought with exposure.

Building blocks. The fits of Section 24.2.6, each drawing's $\hat{\mathbf{z}}_d$ and $\mathbf{\Sigma}_d$ on the loop and its $\hat{\xi}$ and Λ on the cooling and the pump; the Laplace evidence with its constants and the split (24.25) (Section 24.2.2); the Occam factor of a freed entry (Section 24.2.3); the counts $n_<$, $k_<$, $n_\geq$, $k_\geq$ of each cooling drawing. For the exact values beside them, the closed forms of conjugate rows (Section 24.2.4), the gamma densities at λ_{h} of the Savage–Dickey ratio (Section 24.2.5), and the importance estimates of the loop.

Hall 2 and the shared loop. On record A the misfit difference is $\frac{1}{2}(29.85 - 0.79) = 14.53$ nats. By equation (24.25) the curvature term adds -0.234 for the shared entries and the Occam factor of T_{s2}, whose marginal posterior standard deviation $\sqrt{\Sigma_{T_{s2}T_{s2}}}$ is 0.863 K against its prior's 1 K, adds $\ln(1/0.863) = +0.147$:

$$\ln B_{AB} = 14.530 - 0.234 + 0.147 = 14.443.$$

At even prior odds the probability of drawing B is $\mathrm{e}^{-14.44} = 5.3 \times 10^{-7}$; even a prior of 0.99 on drawing B would leave it at 5.3×10^{-5}. The value beside it comes from an importance estimate of the same integrals, 14.424, 0.019 below.

On record B the misfit difference is $-\frac{1}{2}(23.67 - 0.98) = -11.348$; the shared-entry curvature adds -0.189 and the Occam factor of T_{s2}, with $\sqrt{\Sigma_{T_{s2}T_{s2}}} = 0.778$ K, adds $+0.252$:

$$\ln B_{AB} = -11.348 - 0.189 + 0.252 = -11.285,$$

and the probability of drawing A at even odds is 1.3×10^{-5}; the importance estimate is -11.234, 0.051 above. On both records the misfit term is the largest of the three.

Three of four or two of four. Equation (24.26) needs four counts per drawing. Table 24.5 lists them for the four candidate structures, with the log evidence of the loss column on the record and on the softer record, the Laplace value first and the exact one, by equations (24.19) and (24.20), in parentheses. Each row is a drawing; the columns give its threshold, the days below it and their losses, the days at or above it and their losses, and the two log evidences. The common factor $p(\mathbf{y}_{\mathrm{down}})$ of equation (24.18) is omitted.

The record. Under 3-out-of-4 every one of the eight days with two or more units down was lost, as the drawing requires, and the three remaining losses in 392 days are other causes. At the fit of step 3, $\hat{\nu}_{\mathrm{o}} = 4/492$ and $\Lambda = 3.967$, the Laplace evidence is

$$4 \ln \frac{4}{492} + 488 \ln \frac{488}{492} + \ln 99 + \tfrac{1}{2} \ln \frac{2\pi}{3.967} = -19.249 - 3.984 + 4.595 + 0.230 = -18.407;$$

under 2-out-of-4 the seven losses on the seven two-down days must also be other causes, ten in 399 days, and at $\hat{\nu}_{\mathrm{o}} = 11/499$ and $\Lambda = 10.758$

$$11 \ln \frac{11}{499} + 488 \ln \frac{488}{499} + \ln 99 + \tfrac{1}{2} \ln \frac{2\pi}{10.758} = -41.962 - 10.878 + 4.595 - 0.269 = -48.513,$$

so that $\ln B_{32} = 30.106$. The value beside it comes from the exact evidences:

$$\ln p(\mathbf{y}_{\mathrm{loss}} \mid \mathbf{y}_{\mathrm{down}}, M_3) = \ln \frac{\mathrm{B}(1+3,\ 99+389)}{\mathrm{B}(1, 99)} = \ln \frac{\mathrm{B}(4, 488)}{\mathrm{B}(1, 99)} = -18.39.$$

Table 24.5: Evidences of the cooling drawings on the loss record, with $\nu_{\mathrm{o}} \sim \mathrm{Beta}(1, 99)$, and on the softer record ($k_2 = 5$), with $\nu_{\mathrm{m}} \sim \mathrm{Beta}(1, 19)$: the Laplace evidence in the logit, with the exact value of equation (24.19) or (24.20) in parentheses. Counts are those of the loss record; in the softer record $k_{\geq}$ is smaller by two for the first two drawings and $k_{<}$ smaller by two for the last two.

Drawing	Threshold m [–]	$n_{<}$, $k_{<}$ [days]	$n_{\geq}$, $k_{\geq}$ [days]	$\ln p(\mathbf{y}_{\mathrm{loss}} \mid \mathbf{y}_{\mathrm{down}}, M)$, record [–]	$\ln p(\mathbf{y}_{\mathrm{loss}} \mid \mathbf{y}_{\mathrm{down}}, M)$, softer record with ν_{m} [–]
M_4, 4-out-of-4	1	352, 2	48, 9	$-\infty$	-55.83 (-55.79)
M_3, 3-out-of-4	2	392, 3	8, 8	-18.41 (-18.39)	-24.57 (-24.52)
M_2, 2-out-of-4	3	399, 10	1, 1	-48.51 (-48.51)	-40.76 (-40.67)
1-out-of-4	4	400, 11	0, 0	-52.33 (-52.32)	-44.65 (-44.64)
$\ln B_{32}$ [nat]				30.11 (30.12)	16.19 (16.15)

and under 2-out-of-4, where only the one three-down day is explained,

$$\ln p(\mathbf{y}_{\mathrm{loss}} \mid \mathbf{y}_{\mathrm{down}}, M_2) = \ln \frac{\mathrm{B}(11, 488)}{\mathrm{B}(1, 99)} = -48.51.$$

The prior normalizers cancel and, by $\mathrm{B}(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha + \beta)$ and $\Gamma(x + 1) = x\,\Gamma(x)$, the ratio is a product of seven factors, one per loss that 2-out-of-4 must attribute to other causes:

$$\ln B_{32} = \ln \frac{\Gamma(4)\,\Gamma(499)}{\Gamma(11)\,\Gamma(492)} = \sum_{i=0}^{6} \ln \frac{492 + i}{4 + i} = 4.81 + 4.59 + 4.41 + 4.26 + 4.13 + 4.01 + 3.91 = 30.12,$$

0.013 nats above the Laplace value. The more redundant drawing's other-cause rate, 0.022, is 2.2 times the prior's one day in a hundred and 2.7 times that of 3-out-of-4. The 4-out-of-4 drawing has zero evidence: thirty-nine of the forty one-down days passed without a loss.

The softer record. With two two-down days that kept their cooling, 3-out-of-4 without a miss probability is impossible, and the comparison is made with $\nu_{\mathrm{m}} \sim \mathrm{Beta}(1, 19)$ in both drawings. Under 3-out-of-4 the eight days at or above its threshold now hold two misses, $\ln[\mathrm{B}(3, 25)/\mathrm{B}(1, 19)] = -6.14$, and the other-cause part is unchanged at -18.39, a total of -24.52. Under 2-out-of-4 the five two-down losses are other causes, eight in 399 days, $\ln[\mathrm{B}(9, 490)/\mathrm{B}(1, 99)] = -40.62$, and its single three-down day has no miss, $\ln(19/20) = -0.05$, a total of -40.67. The factor falls to $\ln B_{32} = 16.15$, and the Laplace evidences in the two logits, -24.57 and -40.76, give 16.19, the value reported, 0.04 nats above the exact one: two misses at a five-percent rate cost 6.1 nats, and five losses at a one-percent rate cost 22.

The handbook rate. For the pump, at the fit of step 3, $\hat{\omega} = \ln 2$ and $\sigma_\omega^2 = 1/6$, equation (24.27) gives $\ln p(\mathbf{y} \mid M_1) = -1.633 - 0.614 + 0.023 = -2.224$, and the handbook drawing the Poisson probability $\ln(2^4 \mathrm{e}^{-2}/4!) = -2.405$, so that

$$\ln B_{10} = -2.224 + 2.405 = 0.182.$$

The value beside it comes from the exact evidences: equation (24.16) with $a = 2$, $b = 1000\,\mathrm{h}$, $k = 4$ and $T_{\mathrm{tot}} = 2000\,\mathrm{h}$ is the value 0.1097 of Chapter 6, and the handbook drawing gives the Poisson probability of four failures at a mean of $\lambda_{\mathrm{h}} T_{\mathrm{tot}} = 2$, $2^4 \mathrm{e}^{-2}/4! = 0.0902$. By equation (24.28),

$$B_{10} = \frac{\Gamma(6)}{\Gamma(2)} \frac{1000^2}{3000^6} \frac{\mathrm{e}^2}{10^{-12}} = \frac{120 \times 7.389 \times 10^{18}}{7.29 \times 10^{20}} = 1.216, \qquad \ln B_{10} = 0.196.$$

The Savage–Dickey ratio (24.29) gives the same number from the two gamma densities at λ_h: the prior density is $10^6 \times 10^{-3}\,e^{-1} = 367.9\,h$ and the posterior density $6.075 \times 10^{18} \times 10^{-15}\,e^{-3} = 302.5\,h$, so $\ln B_{10} = \ln(367.9/302.5) = 0.196$, 0.014 nats above the Laplace value. Four failures in 2000 h are predicted 1.20 times better by a free rate than by the handbook's (1.22 exactly): a factor not worth more than a mention.

Table 24.6 repeats the comparison for other failure counts in the same exposure. Each row is a count k; the columns give the two log evidences, the log Bayes factor from the fit with the exact value beside it, the exact posterior density at the handbook rate (the prior density is 367.9 h for every row), the central 95 % credible interval of the rate, $\hat{\omega} \pm 1.96\sigma_\omega$ mapped to the rate with the exact interval beside it, and whether both contain λ_h.

Table 24.6: The handbook comparison of the pump for failure counts k in $T_{tot} = 2000\,h$, prior Gamma(2, 1000 h), handbook rate $\lambda_h = 10^{-3}\,h^{-1}$. Laplace evidence and interval from the fit in ω; in parentheses the exact values, from the posterior Gamma($2 + k$, 3000 h).

k [–]	$\ln p(\mathbf{y} \mid M_1)$ [–]	$\ln p(\mathbf{y} \mid M_0)$ [–]	$\ln B_{10}$ [nat]	$p(\lambda_h \mid \mathbf{y}, M_1)$ [h]	95 % interval $[10^{-3}\,h^{-1}]$	Contains λ_h
0	−2.239 (−2.197)	−2.000	−0.239 (−0.197)	448.1	[0.17, 2.67] ([0.08, 1.86])	yes
2	−1.930 (−1.910)	−1.307	−0.623 (−0.603)	672.1	[0.50, 3.55] ([0.36, 2.92])	yes
4	−2.224 (−2.210)	−2.405	+0.182 (+0.196)	302.5	[0.90, 4.45] ([0.73, 3.89])	yes
6	−2.695 (−2.684)	−4.420	+1.726 (+1.736)	64.8	[1.33, 5.33] ([1.15, 4.81])	no
8	−3.252 (−3.244)	−7.059	+3.807 (+3.816)	8.1	[1.79, 6.20] ([1.60, 5.69])	no
12	−4.504 (−4.498)	−13.669	+9.166 (+9.172)	0.038	[2.76, 7.88] ([2.55, 7.41])	no

Figure 24.3 draws the comparison. In panel (a) the horizontal axis is the rate in $10^{-3}\,h^{-1}$ and the vertical axis the density per $10^{-3}\,h^{-1}$; the blue curve is the prior Gamma(2, 1000 h), the green curve the posterior Gamma(6, 3000 h), and the dashed line the handbook rate, where the two dots mark the densities 0.368 and 0.302 whose ratio is the Bayes factor. Panel (b) plots $\ln B_{10}$ against the failure count from 0 to 14, the Laplace evidence dotted in blue and the exact evidence solid in purple, the dots marking the counts of Table 24.6, with the gray lines at 0 and 3. In panel (a) the posterior has moved its mode from 1 to $1.67 \times 10^{-3}\,h^{-1}$ and still holds at λ_h a density close to the prior's: the record neither piled probability onto the handbook value nor drained it away. In panel (b) the curve rises steeply to the right, 3.8 nats at eight failures and 9.2 at twelve, and is shallow to the left, with its lowest value −0.63 at one failure by the Laplace evidence and −0.60 at one and at two failures exactly.

The largest posterior density at λ_h that 2000 h can produce, at one or two failures, is 672.1 h, $672.1/367.9 = 1.83$ times the prior's: the most the record can say for the handbook is 0.63 nats by the Laplace evidence, 0.60 exactly.

24.3.2 The availability under an open structure

Engineering question. What availability is reported when the drawing is uncertain? The availability report needs one number for the hall's cooling, with its band, whether or not the structure is settled.

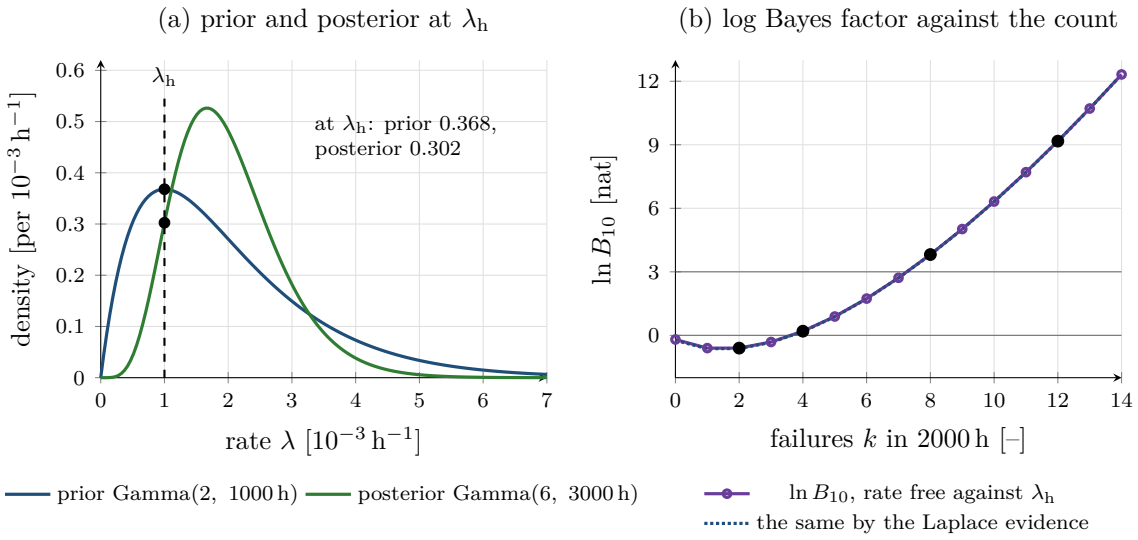


Figure 24.3: The handbook comparison of the pump. (a) The prior density $x\,e^{-x}$ and the posterior density $6.075\,x^5e^{-3x}$ of the rate x in $10^{-3}\,h^{-1}$, with their values at the handbook rate; their ratio $0.368/0.302 = 1.216$ is the Bayes factor B_{10} by equation (24.29). (b) $\ln B_{10}$ of equation (24.28) and of the Laplace evidence (24.27) against the number of failures in 2000 h; gray lines at 0 and 3 nats.

General form. Estimation, the posterior moments of a derived quantity, equation (2.10), with the drawing as one more unknown; by the law of total expectation over the drawing,

$$\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}] = \sum_d \Pr(M_d \mid \mathbf{y})\, \mathbb{E}[f(\mathbf{z}) \mid M_d, \mathbf{y}]. \tag{24.30}$$

Specialized form. The question is equation (24.30) with $f(\mathbf{z}) = U_c$, the state $\mathbf{z}$ the one number ξ_q, each $\Pr(M_d \mid \mathbf{y})$ the weight of the Laplace evidences of Section 24.3.1, and $p(\mathbf{z} \mid M_d, \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\xi}_q, \sigma_q^2)$ of equation (24.24). Under each drawing U_c is a polynomial in $q = 1/(1+e^{-\xi_q})$, $1-(1-q)^3(1+3q)$ under 3-out-of-4 and $4q^3(1-q)+q^4$ under 2-out-of-4; its moments are the Gauss–Hermite sum of equation (5.33) over ξ_q, and because it increases with q, its target probability is the normal tail

$$\Pr(U_c < 10^{-3} \mid M_d, \mathbf{y}) \approx \Phi\left(\frac{\xi_d^* - \hat{\xi}_q}{\sigma_q}\right), \tag{24.31}$$

with ξ_d^* the logit of the q at which the drawing's polynomial equals 10^{-3}. The values beside them come from Chapter 13's exact Beta enumeration on the posterior (24.23): its moments are sums of the Beta moments (13.29) and its target probability an incomplete Beta function, equation (13.37). The prior odds of the structures are the design file's $0.7 : 0.3$, the only prior information on the structure: they enter the weights through equation (2.46), and Section 24.3.2 evaluates even odds beside them.

Here $f(\mathbf{z}) = U_c$, the unavailability of the hall's cooling. Write $\bar{f}_d = \mathbb{E}[f \mid M_d, \mathbf{y}]$ for its mean under drawing d, $\mathrm{Var}(f \mid M_d, \mathbf{y})$ for its variance there and $\bar{f}$ for the mixture mean. The laws of total expectation and total variance over the drawing give

$$\begin{aligned} \mathbb{E}[f \mid \mathbf{y}] &= \sum_d \Pr(M_d \mid \mathbf{y})\, \bar{f}_d, \\ \mathrm{Var}(f \mid \mathbf{y}) &= \sum_d \Pr(M_d \mid \mathbf{y})\, \mathrm{Var}(f \mid M_d, \mathbf{y}) + \sum_d \Pr(M_d \mid \mathbf{y})\, (\bar{f}_d - \bar{f})^2, \end{aligned} \tag{24.32}$$

the mixture of equations (13.50) and (13.51) with the prior weights of Chapter 13 replaced by the posterior ones: a variance within the drawings and a variance between them, the second the doubt about the drawing. The probability of meeting a target is the weighted sum of the drawings' probabilities, equation (13.54), $\Pr(U_{\mathrm{c}} < 10^{-3} \mid \mathbf{y}) = \sum_d \Pr(M_d \mid \mathbf{y})\ \Pr(U_{\mathrm{c}} < 10^{-3} \mid M_d, \mathbf{y})$.

Building blocks. The posterior weights from the Laplace $\ln B_{32}$ of the cooling and the prior odds; the state $\hat{\xi}_q$ and σ_q of equation (24.24) with the Gauss–Hermite nodes and the thresholds ξ_d^*; and, for the exact values beside them, the within-drawing moments and target probabilities of Chapter 13's enumeration under the posterior (24.23).

The mixture over drawings. Table 24.7 evaluates the Gauss–Hermite sum with five nodes. Each row is a node: the first two columns give the node t_i and its weight, the third the state $\xi_i = \hat{\xi}_q + \sqrt{2}\,\sigma_q t_i$, the fourth the unit unavailability there, and the last two the unavailability of the cooling under the two drawings.

Table 24.7: The Gauss–Hermite rule with five nodes over the Laplace posterior $\xi_q \mid \mathbf{y} \approx \mathcal{N}(-3.3081,\ 0.13144^2)$ of equation (24.24).

Node t_i	Weight	State ξ_i	q_i	U_{c} under 3-out-of-4	U_{c} under 2-out-of-4
−2.0202	0.01126	−3.6836	0.02452	3.489×10^{-3}	5.785×10^{-5}
−0.9586	0.22208	−3.4863	0.02970	5.087×10^{-3}	1.025×10^{-4}
0	0.53333	−3.3081	0.03529	7.127×10^{-3}	1.712×10^{-4}
0.9586	0.22208	−3.1299	0.04189	9.950×10^{-3}	2.848×10^{-4}
2.0202	0.01126	−2.9326	0.05057	1.433×10^{-2}	4.976×10^{-4}
Weighted sum				7.341×10^{-3}	1.836×10^{-4}

The weighted sums are the means of the state route, under 3-out-of-4 7.341×10^{-3} and under 2-out-of-4 1.836×10^{-4}; forty nodes change no digit shown, and with the sum of the squares they give the standard deviations 1.839×10^{-3} and 7.14×10^{-5}. The polynomials reach 10^{-3} at $q = 0.01302$ and $q = 0.06404$, $\xi^* = -4.3279$ and -2.6821, so that equation (24.31) gives $\Phi((-4.3279 + 3.3081)/0.13144) = \Phi(-7.76) = 4.3 \times 10^{-15}$ under 3-out-of-4 and $\Phi(4.76) = 0.999999$ under 2-out-of-4.

The values beside them come from the exact posterior. Under the unit posterior (24.23), Beta(60, 1640), Chapter 13's enumeration gives, under 3-out-of-4, a mean unavailability of 7.230×10^{-3} with standard deviation 1.793×10^{-3} and a probability 1.5×10^{-11} of meeting $U_{\mathrm{c}} < 10^{-3}$, and under 2-out-of-4 1.793×10^{-4} with 6.85×10^{-5} and 1.000. Under that chapter's fleet posterior Beta(3, 97), which leaves the units-down column out, they are 6.673×10^{-3}, 7.573×10^{-3} and 0.139, and 2.228×10^{-4}, 4.411×10^{-4} and 0.956; the state route on $\hat{\xi}_q = -3.4761$, $\sigma_q = 0.5862$ gives 9.145×10^{-3}, 1.240×10^{-2} and 0.070, and 3.953×10^{-4}, 1.170×10^{-3} and 0.908, the probabilities by the delta method on $\ln U$ (0.073 and 0.912 with the threshold mapped to the normal), far from them because three units down leave the logit normal poorly matched to the Beta. Table 24.8 evaluates the mixture (24.32) for six cases. Each row is a case: the first two use the fleet posterior alone, with the design file's weights (Chapter 13's report) and with the loss record's posterior weights; the next two use the whole record's unit posterior with the same two sets of weights; the last two are an illustrative, less decisive record with $\ln B_{32} = 1$ at even prior odds and at the design file's odds, on the whole record's unit posterior. The columns give the unit posterior, the weights, the mixture mean, the two parts of the variance, the standard deviation, the between-drawing share and the probability of meeting the target $U_{\mathrm{c}} < 10^{-3}$, each from the state with the exact value in parentheses.

Table 24.8: The mixture over the cooling drawings, equation (24.32), under independent units: the within-drawing moments of the state by Gauss–Hermite over ξ_q, and in parentheses those of Chapter 13's exact enumeration. Unit posterior: the fleet posterior Beta$(3, 97)$, which leaves the loss record's units-down column out, or the whole record's Beta$(60, 1640)$ of equation (24.23). Weights are $\Pr(M_3 \mid \mathbf{y})$ and $\Pr(M_2 \mid \mathbf{y})$; share is the between-drawing part of the variance; the target probability $\Pr(U_{\mathrm{c}} < 10^{-3} \mid \mathbf{y})$ is the weighted sum of the drawings' probabilities. The reported case is the fourth row.

Case	Unit posterior	Weights [–]	Mean $[10^{-3}]$	Within $[10^{-6}]$	Between $[10^{-6}]$	Std. dev. $[10^{-3}]$	Share [–]	Target prob. [–]
Design file, no structure record	Beta$(3, 97)$	0.700, 0.300	6.52 (4.74)	108 (40.2)	16.1 (8.74)	11.1 (7.00)	0.130 (0.179)	0.321 (0.384)
Loss record, units-down column left out	Beta$(3, 97)$	1, 3.6×10^{-14}	9.15 (6.67)	154 (57.4)	2.8×10^{-12} (1.5×10^{-12})	12.4 (7.57)	0.000 (0.000)	0.070 (0.139)
Design file, whole record on q	Beta$(60, 1640)$	0.700, 0.300	5.19 (5.12)	2.37 (2.25)	10.8 (10.4)	3.62 (3.56)	0.820 (0.823)	0.300 (0.300)
Loss record, whole record	Beta$(60, 1640)$	1, 3.6×10^{-14}	7.34 (7.23)	3.38 (3.21)	1.8×10^{-12} (1.8×10^{-12})	1.84 (1.79)	0.000 (0.000)	4.0×10^{-14} (1.5×10^{-11})
$\ln B_{32} = 1$, even prior odds	Beta$(60, 1640)$	0.731, 0.269	5.42 (5.33)	2.47 (2.35)	10.1 (9.78)	3.54 (3.48)	0.803 (0.806)	0.269 (0.269)
$\ln B_{32} = 1$, prior odds 0.7/0.3	Beta$(60, 1640)$	0.864, 0.136	6.37 (6.27)	2.92 (2.78)	6.03 (5.85)	2.99 (2.94)	0.674 (0.678)	0.136 (0.136)

The loss record. From the design file's prior odds 0.7/0.3, equation (24.8) gives log posterior odds $30.106 + \ln(7/3) = 30.95$ and a weight of 3.6×10^{-14} on 2-out-of-4; from even odds the weight is $\mathrm{e}^{-30.106} = 8.4 \times 10^{-14}$ (exactly $30.12 + \ln(7/3) = 30.97$, the same 3.6×10^{-14}, and $\mathrm{e}^{-30.12} = 8.3 \times 10^{-14}$). On the whole record the reported unavailability is $7.34 \times 10^{-3} \pm 1.84 \times 10^{-3}$, and the probability of meeting the target is 4.0×10^{-14}, the weight of 2-out-of-4 plus the 3-out-of-4 tail, against 0.321 for the design-file mixture and 0.070 for the loss record read without its units-down column, the two rows on Beta$(3, 97)$ carrying the delta method on $\ln U$ of the structures chapter. The values beside them come from the exact posterior: $7.23 \times 10^{-3} \pm 1.79 \times 10^{-3}$, 53 % above the design-file mixture's 4.74×10^{-3}, and a probability of 1.5×10^{-11}, against 0.384 and 0.139. The gap on the mean is 1.1×10^{-4}, 1.5 %; the target probability differs by a factor of 370 in a tail 7.8 standard deviations out, where the logit normal is thinner than the Beta, and both are far below any assurance.

A less decisive record. With $\ln B_{32} = 1$ at even prior odds the weights are $\mathrm{e}/(1 + \mathrm{e}) = 0.731$ and 0.269, which give 5.42×10^{-3} on the state route. One row of the arithmetic, with the exact moments:

$$\bar{f} = 0.731 \times 7.230 \times 10^{-3} + 0.269 \times 1.793 \times 10^{-4} = 5.286 \times 10^{-3} + 4.82 \times 10^{-5} = 5.33 \times 10^{-3},$$

the within-drawing variance is $0.731 \times (1.793 \times 10^{-3})^2 + 0.269 \times (6.85 \times 10^{-5})^2 = 2.349 \times 10^{-6} + 1.3 \times 10^{-9} = 2.35 \times 10^{-6}$, and the between-drawing variance is $0.731 \times 0.269 \times (7.051 \times 10^{-3})^2 = 9.78 \times 10^{-6}$, so that the standard deviation is $\sqrt{1.213 \times 10^{-5}} = 3.48 \times 10^{-3}$ and the open structure carries 81 % of the variance.

24.3.3 The reading that settles a structure question

Engineering question. Which reading settles the loop question, before it is installed? With the five readings of Chapter 17 alone, an engineer choosing one more sensor wants the one that would tell the two drawings apart.

General form. Design, the preposterior expectation (2.16),

$$U(a) = \int u \, p(\tilde{y}_a \mid \mathbf{y}) \, \mathrm{d}\tilde{y}_a,$$

with $p(\tilde{y}_a \mid \mathbf{y})$ the predictive of the outcome and u the information the outcome adds about the drawing: the log Bayes factor it contributes in favor of the drawing that is true.

Specialized form. The question is equation (2.16) with a a candidate reading, u the log Bayes factor its outcome contributes, and each drawing's $p(\mathbf{z} \mid \mathbf{y}, M)$ the posterior of step 3, the Laplace normal of its fit, whose predictive is normal below. The candidates are ranked from the record the plant had before the return sensor, the five readings of Chapter 17: this is the ranking an engineer could have made before buying the sensor, and Section 24.3.3 sets it against the gain the installed sensor realized. The drawing that is true is not known when the sensor is bought, so the predictive is the mixture of the two drawings' predictives with their posterior probabilities on that record as weights: this is the gain a buyer can expect, and the gain conditional on each drawing, which the mixture averages, is reported beside it. Each drawing's predictive of the reading is the Laplace normal of its fit: this gives the gain in closed form, and it would misstate a gain where the predictive is skewed, which the realized gains check.

By the multiplication rule, $p(\mathbf{y}, \tilde{y}_a \mid M) = p(\mathbf{y} \mid M)\, p(\tilde{y}_a \mid \mathbf{y}, M)$, so adding a reading $\tilde{y}_a$ to a record $\mathbf{y}$ adds to the log Bayes factor the log ratio of the two drawings' predictive densities of $\tilde{y}_a$ given $\mathbf{y}$, one more term of the sum (24.10). If drawing A is true, u is that log ratio and its expected value is

$$\mathbb{E}\left[\ln \frac{p(\tilde{y}_a \mid \mathbf{y}, M_\mathrm{A})}{p(\tilde{y}_a \mid \mathbf{y}, M_\mathrm{B})} \,\middle|\, M_\mathrm{A} \right] = \int p(\tilde{y}_a \mid \mathbf{y}, M_\mathrm{A}) \, \ln \frac{p(\tilde{y}_a \mid \mathbf{y}, M_\mathrm{A})}{p(\tilde{y}_a \mid \mathbf{y}, M_\mathrm{B})} \, \mathrm{d}\tilde{y}_a, \tag{24.33}$$

the Kullback–Leibler divergence of drawing A's predictive from drawing B's, $\mathrm{KL}(\mathrm{A} \,\|\, \mathrm{B})$, which is never negative: a reading never moves the evidence the wrong way on average. If drawing B is true, the expected gain is $\mathrm{KL}(\mathrm{B} \,\|\, \mathrm{A})$, the same expression with the drawings exchanged, and averaged over which drawing is true,

$$U(a) = \Pr(M_\mathrm{A} \mid \mathbf{y}) \, \mathrm{KL}(\mathrm{A} \,\|\, \mathrm{B}) + \Pr(M_\mathrm{B} \mid \mathbf{y}) \, \mathrm{KL}(\mathrm{B} \,\|\, \mathrm{A}), \tag{24.34}$$

with the posterior probabilities of equation (2.46) on the record the ranking is made from. Under the Laplace posterior each drawing predicts the reading with a normal distribution, mean $\mu_d = f_d(\hat{\mathbf{z}}_d)$ and variance $\varsigma_d^2 = \mathbf{g}_d^\mathsf{T} \boldsymbol{\Sigma}_d \mathbf{g}_d + \sigma_a^2$, the delta method of Chapter 20 with the gradient $\mathbf{g}_d$ of the reading's quantity at $\hat{\mathbf{z}}_d$, plus the reading's own error variance σ_a^2; the predictive standard deviation is written ς_d to keep it apart from the reading's σ_a. For two normal distributions the divergence is

$$\mathrm{KL}(\mathrm{A} \,\|\, \mathrm{B}) = \ln \frac{\varsigma_\mathrm{B}}{\varsigma_\mathrm{A}} + \frac{\varsigma_\mathrm{A}^2 + (\mu_\mathrm{A} - \mu_\mathrm{B})^2}{2\varsigma_\mathrm{B}^2} - \frac{1}{2} \approx \frac{1}{2} \left(\frac{\mu_\mathrm{A} - \mu_\mathrm{B}}{\varsigma} \right)^2 \quad (\varsigma_\mathrm{A} \approx \varsigma_\mathrm{B} = \varsigma), \tag{24.35}$$

half the squared separation of the two predictions in units of their spread.

Four candidates are ranked, each with its quantity and its gradient in the order of the state (24.1), G_ref absorbed. The return sensor, $\sigma_a = 0.3\,\mathrm{K}$, reads $T_s + (Q_1 + Q_2)/\dot{m}c_p$ under drawing A and $T_s + Q_1/\dot{m}c_p$ under drawing B,

$$\mathbf{g}_\mathrm{A} = \begin{pmatrix} 0 & 0 & 1/\dot{m}c_p & 1/\dot{m}c_p & 1 \end{pmatrix}^\mathsf{T}, \qquad \mathbf{g}_\mathrm{B} = \begin{pmatrix} 0 & 0 & 1/\dot{m}c_p & 0 & 1 & 0 \end{pmatrix}^\mathsf{T};$$

a sensor on hall 2's own supply, $\sigma_a = 0.3\,\mathrm{K}$, reads T_s under drawing A and T_{s2} under drawing B, with $\mathbf{g}$ the unit vector of that entry; a second temperature sensor in hall 2, $\sigma_a = 0.5\,\mathrm{K}$, reads

hall 2's balance, with $\mathbf{g}_\mathrm{A} = (0,\ -Q_2\mathrm{e}^{-g_2},\ 0,\ \mathrm{e}^{-g_2},\ 1)^\mathsf{T}$ and the last entry moved to T_{s2} under drawing B; and a load meter on hall 2, $\sigma_a = 1.5\,\mathrm{kW}$, reads Q_2 under both. The reading that settles the question is the one whose predictions differ most across the drawings in units of their spread. The five readings already in hand are the opposite case: drawing B's extra entry T_{s2} is seen only by hall 2's row, the single row in its column of pattern (24.4) besides its prior, and that row has g_2 free to absorb any departure of the reading, so both drawings predict the five readings alike and their factor is near zero, a statement that the record is silent rather than a verdict. By equation (24.35) the expected gain falls below the 3 nats of the strong range once the predictive spread exceeds $|\mu_\mathrm{A} - \mu_\mathrm{B}|/\sqrt{6}$, and with $\varsigma_d^2 = \mathbf{g}_d^\mathsf{T}\boldsymbol{\Sigma}_d\mathbf{g}_d + \sigma_a^2$ that fixes the largest sensor error at which a candidate still reaches it, before the sensor is chosen.

Building blocks. The fits of both drawings to the five readings, their covariances and the gradients above.

The return sensor as the discriminating row. With the five readings of Chapter 17 alone, both drawings fit both records: $2\mathcal{L} = 0.74$ against 0.66 on record A and 1.60 against 0.94 on record B, on three degrees of freedom. The log Bayes factors are

$$\ln B_{AB} = +0.05 \text{ (record A)}, \qquad \ln B_{AB} = -0.02 \text{ (record B)},$$

and the posterior probabilities of drawing A at even odds are 0.513 and 0.496. The posterior standard deviation of T_{s2} is 0.87 K on record A and 0.78 K on record B, and its Occam term costs $\ln(1/0.869) = 0.14$ nats on record A.

Four candidate readings. For the return sensor, the fits of record A predict $\mu_\mathrm{A} = 17.40^\circ\mathrm{C}$ with $\sqrt{\mathbf{g}_\mathrm{A}^\mathsf{T}\boldsymbol{\Sigma}_\mathrm{A}\mathbf{g}_\mathrm{A}} = 0.333\,\mathrm{K}$, the virtual reading of Chapter 21, and $\mu_\mathrm{B} = 15.15^\circ\mathrm{C}$ with 0.316 K; the predictive spreads are $\varsigma_\mathrm{A} = 0.448$ and $\varsigma_\mathrm{B} = 0.436\,\mathrm{K}$ and the predictions are 2.25 K apart. On record B the predictions are 17.18 and 15.19°C, 1.99 K apart. For the hall-2 supply sensor drawing B predicts the fitted T_{s2}, 12.36 or 11.43°C, with posterior spread 0.87 or 0.78 K and predictive spread 0.92 or 0.83 K. Table 24.9 lists the gains of the four candidates: each row is a candidate, and for each record the columns give the gain conditional on drawing A, $\mathrm{KL}(\mathrm{A}\,\|\,\mathrm{B})$, the gain conditional on drawing B, $\mathrm{KL}(\mathrm{B}\,\|\,\mathrm{A})$, and their average (24.34) with the posterior probabilities above.

Table 24.9: Expected evidence gains of the candidate readings from the five-reading fits, by equations (24.35) and (24.34), in nats. Weights $\Pr(M_\mathrm{A} \mid \mathbf{y}) = 0.513$ on record A and 0.496 on record B. Realized gain of the installed return sensor, $\sigma_{T_r} = 0.3\,\mathrm{K}$: 14.39 on record A, 11.27 on record B.

	Record A			Record B		
Candidate	given A	given B	$U(a)$	given A	given B	$U(a)$
return sensor, 0.3 K	13.37	12.64	13.02	10.41	10.12	10.26
hall-2 supply sensor, 0.3 K	0.443	1.388	0.903	0.560	1.821	1.196
second hall-2 sensor, 0.5 K	0.0003	0.0003	0.0003	0.016	0.017	0.017
hall-2 load meter, 1.5 kW	0.0010	0.0010	0.0010	0.010	0.011	0.010

The return sensor is expected to add 13.0 nats on record A and 10.3 on record B, and it realized 14.39, $\ln B_{AB}$ moving from +0.05 to +14.44, and 11.27, from −0.02 to −11.28. The ranking, return sensor far ahead of the hall-2 supply sensor and both far ahead of the other two, holds on both records and under either drawing.

Figure 24.4 shows how the return sensor's precision sets the answer. The horizontal axis, logarithmic, is the standard deviation σ_{T_r} of the return sensor's error from 0.1 to 3 K, and the

vertical axis the log Bayes factor of drawing A against drawing B. The solid curves are the realized values on the two records, refitted at each σ_{T_r}, and the dashed curves the expected values of equation (24.35) conditional on the drawing that generated each record, signed as the factor it would move; the dots mark the installed sensor, $\sigma_{T_r} = 0.3\,\mathrm{K}$, and the gray lines are $\ln B_{AB} = \pm 3$. At 0.1 K the records give +25.1 and −19.1 nats; at 1 K, +2.5 and −2.0; at 3 K, +0.35 and −0.26. The strong range of 3 nats is reached for a sensor error below 0.91 K on record A and 0.80 K on record B (realized). The expected gain conditional on the generating drawing falls below 3 nats above 0.86 K on record A and 0.74 K on record B, and the posterior-weighted gain (24.34) above 0.86 and 0.75 K. On record A the approximate form of equation (24.35) gives the same threshold: $\varsigma > 2.25/\sqrt{6} = 0.92\,\mathrm{K}$, a sensor error of $\sqrt{0.92^2 - 0.33^2} = 0.86\,\mathrm{K}$.

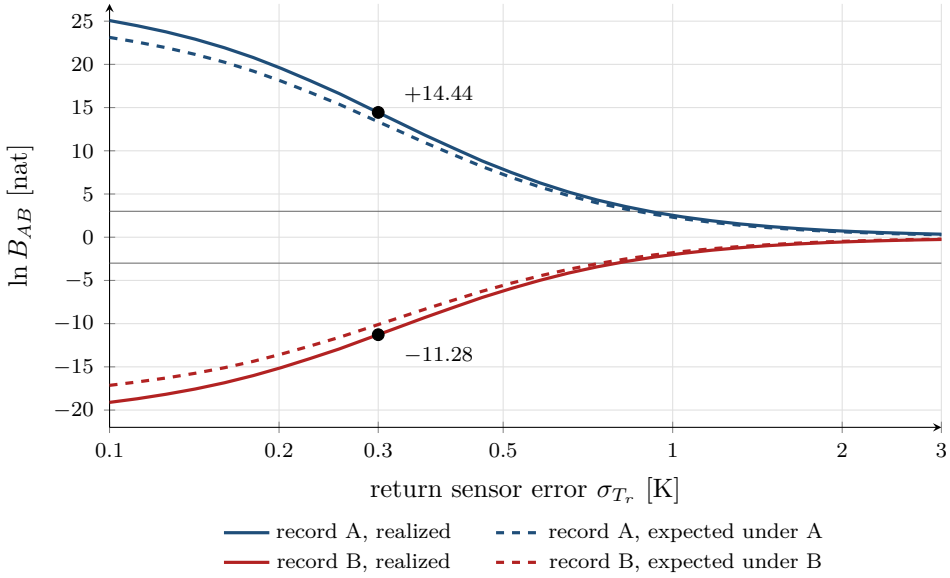


Figure 24.4: The log Bayes factor of drawing A (shared supply) against drawing B (hall 2 on its own supply) as a function of the return sensor's error. Solid: realized on records A and B, each drawing refitted at each σ_{T_r} and its Laplace evidence equation (24.12). Dashed: the expected value of equation (24.35) conditional on the drawing that generated the record, from the fits without the return reading. Dots: the installed sensor, $\sigma_{T_r} = 0.3\,\mathrm{K}$. Gray: ± 3 nats.

24.4 The checks (step 5)

Each evidence of Sections 24.2 and 24.3 rests on a route with its own assumption, and each is set beside a calculation that does not make it. The loop's drawings carry the physics rows of Chapter 17 in the limit $\sigma_p \to 0$, and an evidence keeps constants that the limit could leave behind; the cooling and the pump have no physics row. An evidence also depends on the prior of the quantity a drawing frees, as no posterior does, and the mixture rests on an independence of the units that the loss record itself can test.

24.4.1 The evidences against fuller calculations

Table 24.10 lists the checks: each row is one evidence or log Bayes factor, and the columns give the value reported in Section 24.3, the value by the check, and the check itself.

The Laplace evidence of the loop. The importance estimate is an unbiased estimate of the integral of equation (24.6) itself, with the same constants: the draws $\mathbf{z}^{(i)}$ come from a

Table 24.10: The evidences checked. For the loop the Laplace values use the Gauss–Newton curvature; the importance estimates use 400 000 draws from a Student-t distribution with five degrees of freedom centered at $\hat{\mathbf{z}}$ with scale $1.5^2\boldsymbol{\Sigma}$, effective sample size 150 000 to 175 000 and standard error 0.002; the grid integrates T_{s2} over $[4,\ 20]^{\circ}$C in steps of 0.025 K with the other five entries integrated by the Laplace formula at each point. For the pump the Laplace route is the fit in ω.

Quantity	Used [–]	Check [–]	Check
Loop, record A, $\ln p(\mathbf{y} \mid M_{\mathrm{A}})$	-4.497	-4.486	importance sampling
Loop, record A, $\ln p(\mathbf{y} \mid M_{\mathrm{B}})$	-18.940	-18.910	importance sampling; grid on T_{s2}: -18.901
Loop, record A, $\ln B_{AB}$	14.443	14.424	importance sampling; full Hessian: 14.428
Loop, record B, $\ln p(\mathbf{y} \mid M_{\mathrm{A}})$	-15.724	-15.704	importance sampling
Loop, record B, $\ln p(\mathbf{y} \mid M_{\mathrm{B}})$	-4.439	-4.470	importance sampling; grid on T_{s2}: -4.472
Loop, record B, $\ln B_{AB}$	-11.285	-11.234	importance sampling; full Hessian: -11.243
Loop, no return row, $\ln B_{AB}$	$+0.05$, -0.02	$+0.02$, 0.00	importance sampling, records A and B
Cooling, $\ln p(\mathbf{y}_{\mathrm{loss}} \mid \mathbf{y}_{\mathrm{down}}, M_3)$, M_2	-18.387, -48.506	-18.387, -48.506	numerical quadrature over ν_{o}
Cooling, $\ln p(\mathbf{y}_{\mathrm{loss}} \mid \mathbf{y}_{\mathrm{down}}, M_3)$, M_2	-18.407, -48.513	-18.387, -48.506	the exact Beta integral
Cooling, $\ln B_{32}$	30.106	30.119	the exact Beta integral
Cooling, softer record, M_3, M_2	-24.522, -40.674	-24.522, -40.674	quadrature over ν_{o} and ν_{m}
Cooling, softer record, M_3, M_2	-24.571, -40.765	-24.522, -40.674	the exact Beta integrals
Pump, $\ln B_{10}$	0.196	0.196	Savage–Dickey ratio against the closed form
Pump, $\ln p(\mathbf{y} \mid M_1)$	-2.224	-2.210	the exact negative binomial
Pump, $\ln B_{10}$	0.182	0.196; 0.464	the closed form; the Savage–Dickey ratio with the Laplace normal in the numerator

heavy-tailed proposal q around the MAP state, and the evidence is the average of $\exp[-\mathcal{L}(\mathbf{z}^{(i)})]$ times the constants divided by $q(\mathbf{z}^{(i)})$, the reweighting of Chapter 20 with the normalizer kept. The Laplace log evidences differ from it by 0.01 to 0.03 nats, and the log Bayes factors by 0.02 and 0.05, against 14 and 11. The grid over T_{s2} isolates the entry that drawing B adds: at each value of T_{s2} the other five entries are integrated by the Laplace formula, and the outer integral is done numerically against the set-point prior. It agrees with the six-entry Laplace value within 0.04 nats on both records, so the freed entry is as Gaussian as the others. The full Hessian $\mathbf{H}$, with the curvature term $\sum_k r_k\, \partial^2 r_k/\partial \mathbf{z}^2$ that the Gauss–Newton $\mathbf{\Lambda}$ drops, moves the log Bayes factor by 0.015 and 0.04 nats: the curvature term is small because the residuals are, except in the poor fits, where it does not matter. For this plant and these records the Laplace evidence is adequate to two decimals in the log Bayes factor.

The closed forms. The Beta integrals of the cooling agree with a numerical quadrature over the nuisance probabilities to every printed digit, as they must, since equations (24.19) and (24.20) are exact. The fit on the logit state of Section 24.3.1 is the route reported: under 3-out-of-4 it stops at $\hat{\xi}_{\mathrm{o}} = -4.804$, $\hat{\nu}_{\mathrm{o}} = 4/492$, with the curvatures 3.161 of the count row and 0.806 of the prior row, $\Lambda = 3.967$ and $\sigma_{\xi_{\mathrm{o}}} = 0.502$; under 2-out-of-4 at $\hat{\xi}_{\mathrm{o}} = -3.792$ with the curvatures 8.602 and 2.156 and $\Lambda = 10.758$. Its Laplace evidences, -18.407 and -48.513, are 0.02 and 0.01 nats below the exact ones, and the factor 30.106 is 0.013 nats short of the exact 30.119. For the pump the Savage–Dickey ratio and the ratio of the negative-binomial to the Poisson probability are the same number, an identity under the conjugate prior; the check is on the arithmetic, not on an approximation.

The pump in the log-rate. The route reported for the pump is the fit of Section 24.2.6 in ω. Its evidence is the height of the integrand at the minimum times the posterior volume, as in Section 2.3.6, $\ln p(k \mid \hat{\omega}) + \ln p(\hat{\omega}) + \frac{1}{2}\ln(2\pi\sigma_\omega^2) = -1.633 - 0.614 + 0.023 = -2.224$, against the exact -2.210, and a log Bayes factor of 0.182 against 0.196. The Laplace Savage–Dickey ratio reads instead the normal $\mathcal{N}(\ln 2,\ 1/6)$ at the pinned value $\omega = 0$, a density of 0.231 against the exact posterior's 0.302 and the prior's 0.368, and gives 0.464, more than twice the exact value: the pinned value lies 1.7 posterior standard deviations from the minimum, in the tail where the posterior of ω, a log-gamma density, departs from the normal. With eight failures the gap widens, 6.02 by the Laplace ratio against the exact 3.82, while the Laplace evidence gives 3.81. The evidence integrates the whole posterior, and its Laplace error stays small; the ratio reads one point of it, in a tail, which is why the chapter's state route for the pump is the evidence and not the ratio.

24.4.2 The exact balances as a limit

An evidence keeps the normalizing constants that a posterior drops, and the reduced state removes the physics rows together with theirs. In the full state of Chapter 17 each drawing of the loop has nine or ten entries, the hall temperatures, the chiller's heat and the return temperature among them, and four physics rows with the tolerance σ_p: the two hall balances, the header balance and the return header's balance, $T_r = T_s + (Q_1 + Q_2)/\dot{m}c_p$ under drawing A and $T_s + Q_1/\dot{m}c_p$ under drawing B. Each physics row's constant $-\ln(\sqrt{2\pi}\,\sigma_p)$ enters equation (24.12), and each carried entry's posterior width, proportional to σ_p, enters $\det \mathbf{\Sigma}$; the two cancel as $\sigma_p \to 0$, as equation (4.19) found for one hall, and both drawings carry four physics rows, so their constants cancel from the ratio at every tolerance. Table 24.11 checks the limit. Each column is one tolerance, σ_p in kelvin for the three temperature balances and the same number in kilowatts for the header, and the last column is the reduced state; the rows are the log evidences and log Bayes factors on the two records and the spread of T_{s2} under drawing B.

Table 24.11: The loop's evidences on the full state, with four physics rows of tolerance σ_p and every row's constant in equation (24.12), against the reduced state of Chapter 17. Spread of T_{s2}: $\sqrt{\Sigma_{T_{s2}T_{s2}}}$ under drawing B.

Quantity	$\sigma_p = 0.1$	0.01	0.001	reduced
Record A, $\ln p(\mathbf{y} \mid M_\mathrm{A})$ [–]	−4.523	−4.497	−4.497	−4.497
Record A, $\ln p(\mathbf{y} \mid M_\mathrm{B})$ [–]	−18.240	−18.933	−18.940	−18.940
Record A, $\ln B_{AB}$ [nat]	13.717	14.436	14.443	14.443
Record B, $\ln p(\mathbf{y} \mid M_\mathrm{A})$ [–]	−15.198	−15.718	−15.724	−15.724
Record B, $\ln p(\mathbf{y} \mid M_\mathrm{B})$ [–]	−4.466	−4.439	−4.439	−4.439
Record B, $\ln B_{AB}$ [nat]	−10.732	−11.279	−11.285	−11.285
Record B, spread of T_{s2} [K]	0.7788	0.7776	0.7776	0.7776

The log Bayes factor approaches the reduced value as the square of the tolerance: it is short by 0.73 and 0.55 nats at 0.1 K, by 0.008 and 0.006 at 0.01 K, and by 10^{-4} at 0.001 K. A tolerance of a tenth of a kelvin is a balance that may fail by a third of a sensor's spread, and it lets the drawing that misfits absorb part of the return reading's disagreement in its balances; that is a different, looser drawing, not the one compared. The reduced state's evidences are the limit, and they are the ones reported.

24.4.3 Each drawing tested by its own residuals

The χ^2 check of Chapter 18 tests one drawing against its own residuals; the evidence compares two. On the loop with the return reading the two agree: drawing B on record A has $2\mathcal{L} = 29.85$ on four degrees of freedom, a tail of 5×10^{-6} in Table 24.4, and fails its test, while drawing A passes with a tail of 0.94; on record B the roles reverse, with tails of 9.3×10^{-5} and 0.91. They part in two ways. Without the return reading both drawings pass their tests on both records, with tails of 0.86 and 0.88 on record A and 0.66 and 0.82 on record B (drawings A and B, three degrees of freedom), and a test would report both as adequate; the evidence reports that the record cannot tell them apart, which is the more useful statement because it names what is missing. And a drawing that fails a test is not thereby rejected in favor of anything: on the softer cooling record the 3-out-of-4 drawing with a deterministic loss rule is impossible, a test with tail zero, yet with the miss probability it is preferred to 2-out-of-4 by 16 nats. A test says whether one drawing could have produced the record; only a comparison says which of the candidates did, and only among the candidates written down.

24.4.4 Sensitivity to the prior of the freed quantity

An evidence integrates over the prior of every quantity, and the priors of the quantities two drawings share largely cancel from their ratio; the prior of the quantity one drawing frees and the other does not does not cancel. The one-reading case of Section 24.2.3 shows the mechanism. Figure 24.5 sets it at the illustrative values $y = 2$, $\sigma = 1$ and $z^\mathrm{h} = z^0 = 0$. In panel (a) the horizontal axis is the quantity z and the vertical axis a density; the black dashed curve is the likelihood of the reading as a function of z, and the three colored curves are the priors of the free drawing with standard deviations $\sigma_0 = 0.5$, 2 and 8. Panel (b) plots, on the same horizontal axis, the product of the likelihood and each prior; by equation (24.6) the area under each curve is that drawing's evidence.

By equation (24.13) the MAP states are $\hat{z} = 0.40$, 1.60 and 1.97 and the posterior spreads

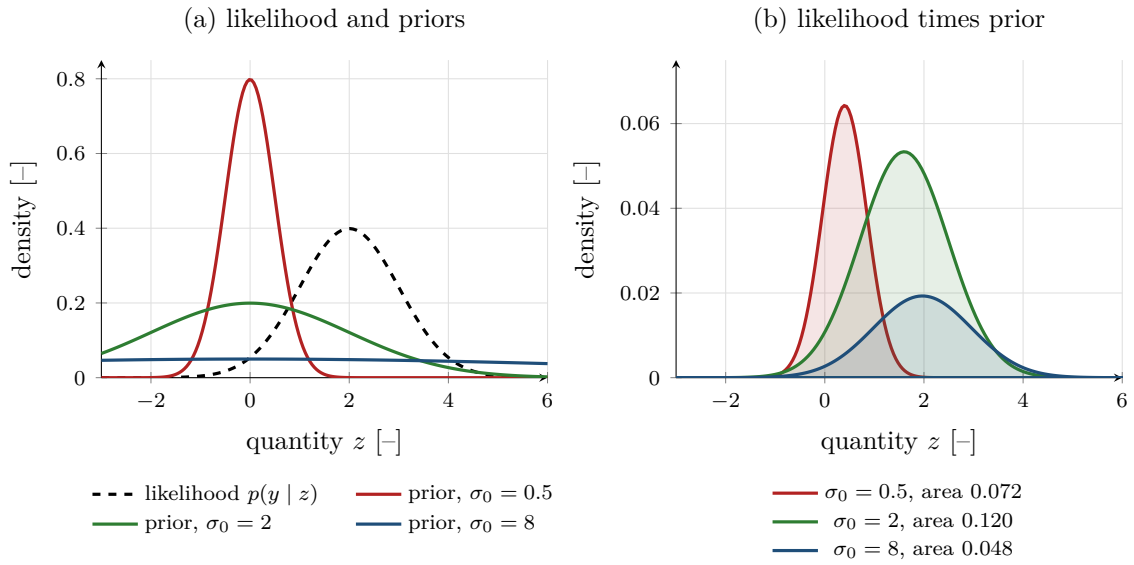


Figure 24.5: The Occam factor for one reading $y = 2$ with error $\sigma = 1$ of a quantity z freed with the prior $\mathcal{N}(0, \sigma_0^2)$; illustrative values. (a) The likelihood $\mathcal{N}(2;\ z,\ 1)$ as a function of z and the priors for $\sigma_0 = 0.5$, 2 and 8. (b) The products of the likelihood and each prior; the area under each is the evidence $\mathcal{N}(2;\ 0,\ 1 + \sigma_0^2)$ of equation (24.14). The pinned drawing $z = 0$ has evidence 0.054.

$\sigma_z = 0.447$, 0.894 and 0.992, and equation (24.15) splits each area into its three factors:

$$\begin{aligned} \sigma_0 = 0.5: &\quad 0.111 \times 0.894 \times 0.726 = 0.072, \\ \sigma_0 = 2: &\quad 0.368 \times 0.447 \times 0.726 = 0.120, \\ \sigma_0 = 8: &\quad 0.399 \times 0.124 \times 0.970 = 0.048, \end{aligned}$$

best fit, Occam factor and misfit in that order. The narrow prior loses on its best fit: it holds $\hat{z}$ at 0.40, where the likelihood of the reading is 0.111, and its product in panel (b) is tall but placed at the wrong z. The intermediate prior lets $\hat{z}$ reach 1.60 and pays the same misfit factor, 0.726, as the narrow one, with an Occam factor of half. The wide prior fits best, 0.399, and pays almost no misfit, but it spends nearly all its mass where the record does not fall, and its Occam factor 0.124 cuts its area to the smallest of the three. Against the pinned drawing, whose evidence is $\mathcal{N}(2;\ 0,\ 1) = 0.054$, the log Bayes factors are 0.29, 0.80 and -0.12: the same reading supports freeing the quantity under the middle prior and opposes it under the widest one, which fits it best.

Table 24.12 recomputes the three comparisons under other priors for the quantity each frees. Its upper part varies the prior standard deviation σ_0 of the hall-2 supply temperature T_{s2} of drawing B, centered at 12°C, and gives the log Bayes factor of the loop on both records with the return reading and on record A without it, and the estimate $\hat{T}_{s2}$ on record B. Its middle part varies the shape a of the pump's prior with the mean held at $2 \times 10^{-3}\,\mathrm{h}^{-1}$, so that $b = a/(2 \times 10^{-3})$, and gives $\ln B_{10}$ for four, eight and two failures; the limit $a \to \infty$ is the point $\lambda = 2 \times 10^{-3}\,\mathrm{h}^{-1}$. Its lower part varies the nuisance priors of the cooling.

The loop. With the return reading the verdict does not move: 14.4 to 16.0 nats on record A and -11.3 to -9.4 on record B over a factor of thirty-two in the prior's width. As the prior of T_{s2} widens from 0.5 to 2 K, record A's factor falls slightly, from 14.52 to 14.43, because the wider prior lets drawing B fit a little better, and beyond 2 K it rises, to 15.99 at 16 K, as drawing B's Occam charge grows; on record B the wider prior lets T_{s2} move toward the true 9°C, which it

Table 24.12: The log Bayes factors under other priors for the quantity each comparison frees. Upper: the loop, prior $T_{s2} \sim \mathcal{N}(12\,^\circ\mathrm{C},\ \sigma_0^2)$. Middle: the pump, prior Gamma$(a,\ 500a\,\mathrm{h})$, mean $2 \times 10^{-3}\,\mathrm{h}^{-1}$. Lower: the cooling, the prior Beta$(1, b_\mathrm{o})$ of ν_o on the record and Beta$(1, b_\mathrm{m})$ of ν_m on the softer record. Pump and cooling rows: the Laplace evidence, with the exact value in the row below. The chapter's priors are in bold.

Loop, σ_0 of T_{s2} [K]	0.5	**1**	2	4	8	16
$\ln B_{AB}$, record A [nat]	14.52	**14.44**	14.43	14.75	15.33	15.99
$\ln B_{AB}$, record B [nat]	−11.33	**−11.28**	−11.09	−10.65	−10.05	−9.38
$\ln B_{AB}$, record A, no return row [nat]	0.05	**0.05**	0.18	0.59	1.19	1.86
$\hat{T}_{s2}$ [°C], record B	11.78	**11.43**	10.96	10.65	10.52	10.48
Pump, shape a [–]	0.5	1	**2**	4	16	∞
$\ln B_{10}$, $k = 4$ [nat]	−0.48	−0.11	**0.18**	0.41	0.66	0.77
exact	−0.46	−0.10	**0.20**	0.42	0.66	0.77
$\ln B_{10}$, $k = 8$ [nat]	3.38	3.66	**3.81**	3.84	3.71	3.55
exact	3.39	3.67	**3.82**	3.85	3.72	3.55
$\ln B_{10}$, $k = 2$ [nat]	−1.04	−0.78	**−0.62**	−0.56	−0.57	−0.61
exact	−1.01	−0.75	**−0.60**	−0.55	−0.57	−0.61
Cooling, b_o [–]	9	49	**99**	199	999	
$\ln B_{32}$, record, b_o varied [nat]	28.70	29.36	**30.11**	31.39	37.36	
exact	28.71	29.37	**30.12**	31.41	37.37	
Cooling, b_m [–]	4	9	**19**	49	99	
$\ln B_{32}$, softer record, b_m varied [nat]	17.40	16.96	**16.19**	14.81	13.59	
exact	17.35	16.92	**16.15**	14.77	13.54	

does only to 10.5°C because the conductance prior still takes its share of the deficit; the better fit offsets part of the growing Occam charge, and the factor moves toward drawing A by 1.9 nats without approaching zero. Without the return reading the factor is pure Occam charge: both drawings fit equally and $\ln B_{AB}$ rises with $\ln \sigma_0$, by 0.67 nats from 8 to 16 K, close to $\ln 2 = 0.69$. This is Lindley's paradox at work: on a record that cannot discriminate, a wide enough prior on the freed quantity manufactures support for the drawing without it.

The pump. With four failures the sign of $\ln B_{10}$ depends on the prior: -0.48 (-0.46 exactly) under a prior twice as wide in its coefficient of variation, $+0.77$ in the limit of a prior concentrated at its mean. No prior in the family makes the record a strong argument either way, and the conclusion, that four failures in 2000 h neither confirm nor contradict the handbook, holds across it. Eight failures stay between 3.4 and 3.9 nats, against the handbook under every prior. Two failures stay between -1.0 and -0.56 (-0.55 exactly): weak support for the handbook under every prior, and strongest under the widest, Lindley's direction.

The cooling. The factor ranges from 28.7 to 37.4 nats as the prior mean of the other-cause probability goes from 0.1 to 0.001: a rarer other cause makes each unexplained loss dearer. On the softer record the miss prior moves it from 17.4 to 13.6 (13.5 exactly). No reasonable prior brings either record near the boundary of the strong range.

24.4.5 The units-down column against independence

The unit posterior (24.23) reads the units-down column as 1600 independent unit-checks; the same column tests that reading. Read without it, the column measures the 3-out-of-4 unavailability directly: 8 of the 400 morning checks found two or more units down, 2.0 % of the time, and with a uniform prior Beta$(1, 1)$ on that fraction the posterior is Beta$(9, 393)$, of mean 0.022 and 95 % interval [0.010, 0.039], three times the 7.23×10^{-3} of independent units. Under independence and the posterior (24.23) the expected number of such days is $400\,\mathbb{E}[U_{\mathrm{c}} \mid M_3, \mathbf{y}] = 2.9$, and the posterior predictive probability of eight or more, the binomial tail averaged over ξ_q by Gauss–Hermite, is 0.018 (0.017 under the exact Beta); under the fleet posterior alone, which leaves the column out, it is 0.132 (0.082). A predictive tail of 0.018 is a hint that units go down together more often than independence allows, not a verdict: the verdict is a common-cause drawing of the four units, the subject of Chapter 15, compared with the independent one by its evidence on the same column, which the next measurement takes up (Section 24.6.1). Both readings put the 3-out-of-4 unavailability far above the target of 10^{-3}; they differ on its level.

24.5 The report (step 6)

The availability report that the drawings as drawn would have produced is wrong about the target. With the design file's weights the hall's cooling met its target of 10^{-3} with probability 0.321 (0.384 exactly); the loss record says that the hall needs three of its four units, the N+1 of the design, by 30.1 nats over two, and its units-down column, 57 units down in 1600 unit-checks, fixes the unit unavailability near 0.035, so that the expected unavailability is $7.34 \times 10^{-3} \pm 1.84 \times 10^{-3}$ from the posterior of the state, a rise of 13 % over the design-file mixture ($7.23 \times 10^{-3} \pm 1.79 \times 10^{-3}$ exactly, a rise of 53 %), and the probability of meeting the target falls to 4.0×10^{-14} (1.5×10^{-11} exactly): the cooling misses its target by both. Read with the loss record's weights but without its units-down column, the same mixture would have said 0.070 (0.139 exactly). The level rests on independent units, and the same column hints, with a predictive tail of 0.018, that units go down together; its direct count reads three times higher, which would move the level and not the verdict. The loop is settled, by the

return-header reading alone: with it the record says which drawing the plant is by 14.4 or 11.3 nats, and without it the two drawings, both of which pass their tests, differ by 0.05. And the pump's interval that contains the handbook rate is no support for the handbook: the factor is 0.18 nats (0.20 exactly).

24.5.1 The answers and their bands

Table 24.13 gives the answers. Each row is a question of Table 24.2, in its order; the second column gives the answer under the chapter's priors, from the fit of the state, the third the exact value beside it, the fourth its band, which for a comparison is the range over the priors of the freed quantity in Table 24.12, and the last the evidence behind it.

Table 24.13: The answers to the questions of Table 24.2. Bands of the comparisons: the range of the log Bayes factor over the priors of Table 24.12.

Question	Answer (state route)	Exact	Band	Evidence
Is hall 2 really on the shared loop?	on record A yes, $\ln B_{AB} = +14.44$, $\Pr(M_{\mathrm{A}} \mid \mathbf{y}) = 1 - 5 \times 10^{-7}$; on record B no, -11.28, $\Pr = 1.3 \times 10^{-5}$	importance estimates 14.424, -11.234	14.4 to 16.0 and -11.3 to -9.4 for σ_0 of T_{s2} from 0.5 to 16 K	the return reading; without it $\ln B_{AB} = +0.05$ and -0.02
Is it really parallel? Does N+1 hold? Does the cooling need three units or two?	three, N+1: 3-out-of-4 over 2-out-of-4 (N+2) by $\ln B_{32} = 30.11$; 4-out-of-4 impossible; softer record 16.19	30.12; softer record 16.15	28.7 to 37.4 over the other-cause prior; softer record 13.6 to 17.4 over the miss prior (13.5 to 17.4 exact)	400 days; seven two-down losses that 2-out-of-4 must call other causes
Does the pump's record contradict the handbook rate?	no, and it does not confirm it: $\ln B_{10} = 0.18$; the interval $[0.90,\ 4.45] \times 10^{-3}\,\mathrm{h}^{-1}$ contains λ_{h}	0.20, Savage–Dickey 367.9/302.5; interval $[0.73,\ 3.89]$	-0.48 to $+0.77$ over the prior's shape at its mean (-0.46 to $+0.77$ exact)	at most 0.63 from 2000 h (0.60 exact); eight failures would give 3.8, twelve 9.2
What availability is reported when the drawing is uncertain?	3-out-of-4's, from the whole record (unit posterior Beta(60, 1640)) under independent units: $U_{\mathrm{c}} = 7.34 \times 10^{-3} \pm 1.84 \times 10^{-3}$, $\Pr(U_{\mathrm{c}} < 10^{-3} \mid \mathbf{y}) = 4.0 \times 10^{-14}$, the target missed; withdrawn: the design-file mixture's 6.52×10^{-3} and 0.321, and the 0.070 of the loss record without its units-down column; the level not final until a common-cause drawing is compared	$7.23 \times 10^{-3} \pm 1.79 \times 10^{-3}$, 1.5×10^{-11}; withdrawn 4.74×10^{-3}, 0.384 and 0.139	posterior weights 1 and 3.6×10^{-14} from the design file's odds, 8.4×10^{-14} from even odds (8.3×10^{-14} exact); the direct count of the units-down column, 0.022 $[0.010,\ 0.039]$, three times the independence answer	8 of 400 checks at or past the threshold against 2.9 expected under independence, predictive probability 0.018 of eight or more (0.017 exact)
Which reading settles the loop question, before it is installed?	the return-header temperature: predicted 2.25 and 1.99 K apart, expected gain 13.0 and 10.3 nats at $\sigma_{T_r} = 0.3$ K, realized 14.39 and 11.27	– (the realized gains check it)	strong (3 nats) expected for a sensor error below 0.86 K on record A and 0.75 K on record B; realized below 0.91 and 0.80 K	from the five-reading fits: hall-2 supply sensor 0.90 and 1.20; second hall-2 sensor 0.0003 and 0.017; hall-2 load meter 0.001 and 0.010

24.5.2 The Bayes factor beside the credible interval

For the pump with four failures the 95 % credible interval of the state, $[0.90,\ 4.45] \times 10^{-3}\,\mathrm{h}^{-1}$ (Chapter 6's exact $[0.73,\ 3.89]$), contains the handbook rate, and the log Bayes factor is 0.18 (0.20 exactly). The two say the same thing differently. The interval answers an estimation question inside the free drawing: given the engineering prior and the record, the handbook value is a plausible value of the rate, with $\Phi(-0.6931/0.4082) = 4.5\,\%$ of the posterior below it (8.4 % exactly). The Bayes factor answers a comparison between drawings: a rate free over the engineering prior predicted four failures 1.20 times better (1.22 exactly) than the handbook rate did. The interval does not depend on whether the handbook drawing exists; the Bayes factor depends on the width of the prior the free drawing spreads, from -0.48 (-0.46 exactly) to $+0.77$ over the family of priors with the same mean. The Bayes factor says it with the prior's width on the table, which is its strength when that width is engineering knowledge and its weakness when it is arbitrary.

Up to six failures the two agree in direction: the interval contains λ_h and the factor is below 1 nat, or the interval excludes it and the factor is positive. They part when the record is long and the rate close to the handbook's, the case of Exercise 24.6(b): the interval excludes the handbook value while the factor favors the handbook point. The interval then says the rate is not exactly $10^{-3}\,\mathrm{h}^{-1}$; the Bayes factor says that the point $10^{-3}\,\mathrm{h}^{-1}$ predicted the record better than a prior spread over 0.24 to $5.6 \times 10^{-3}\,\mathrm{h}^{-1}$, most of which the record excludes. Both statements are true. The engineering question decides which is wanted: whether the handbook rate may be used as it is, an estimation question answered by the interval and by how much the difference matters, or whether the pump is the machine the handbook describes, a comparison answered by the factor with the prior of the alternative stated.

24.6 The next measurement (step 7)

The report settles the loop's structure and the cooling's within the family compared, and leaves the pump's question open. Each opens a different next step: for the loop, the value of the node the new drawing introduced; for the cooling, a drawing the family lacks; for the pump, exposure.

24.6.1 The candidate measurements

The hall-2 supply temperature. Under drawing B the fit of record B puts hall 2's supply at $\hat{T}_{s2} = 11.43$°C with $\sqrt{\Sigma_{T_{s2}T_{s2}}} = 0.78\,\mathrm{K}$, and the record was generated at 9°C. The estimate is pulled because the record determines only the sum $T_{s2} + Q_2/G_2$ that hall 2's sensor reads, with Q_2 fixed by the two meters at about 21 kW, and the column of T_{s2} in pattern (24.4) holds no other reading; how the sum divides between a low supply temperature and a high conductance is decided by the two priors, the set point $12 \pm 1\,\mathrm{K}$ and the handbook conductance 5 kW/K with 30 % spread. At the prior point the hall would read $12 + 21.1/5 = 16.2$°C, 1.4 K above its reading; the fit takes 0.57 K of that from the set point, 0.70 K from the conductance, raised to 5.99 kW/K against a true 4.0, and leaves 0.15 K as residual. The two priors share the deficit roughly in proportion to their variances in temperature units, $1\,\mathrm{K}^2$ for the set point and $(4.2 \times 0.3)^2 = 1.6\,\mathrm{K}^2$ for the conductance term. The true 9°C lies 3.1 standard deviations below the estimate. The candidates are the readings that would add a row in the column of T_{s2} or g_2: a sensor on hall 2's own supply at $\sigma_a = 0.1$, 0.3 or 0.5 K, whose Jacobian row is the unit vector of T_{s2} over σ_a; a second temperature sensor in hall 2 at 0.5 K, the gradient of Section 24.3.3 over σ_a; and a load meter on hall 2 at 1.5 kW.

The cooling and the pump. For the cooling the record is decisive within the family, and further days would serve the units' unavailabilities, not the structure; what the units-down column asks for is a drawing, a common-cause model of the units, set beside the independent one and compared with it by its evidence on the same 400 days, before the level of the unavailability is reported as final. For the pump the handbook question can be settled in favor of the handbook only by exposure, and against it quickly if the rate is in fact higher.

24.6.2 The preposterior of the hall-2 supply temperature

A candidate reading adds its Jacobian row $\mathbf{J}_a$ to the fit of drawing B on record B, and the covariance it would leave is the preposterior of equation (2.17), $\mathbf{\Sigma}_a = (\mathbf{\Sigma}_\mathrm{B}^{-1} + \mathbf{J}_a^\mathsf{T}\mathbf{J}_a)^{-1}$, which needs no outcome. For a supply sensor the row is $\mathbf{J}_a = \mathbf{e}_6^\mathsf{T}/\sigma_a$, with $\mathbf{e}_6$ the unit vector of T_{s2}, and the update of that entry is the scalar one, $\Sigma_{a,T_{s2}T_{s2}} = (\Sigma_{T_{s2}T_{s2}}^{-1} + \sigma_a^{-2})^{-1}$. Table 24.14 ranks the candidates. Each row is a candidate; the columns give the information it adds on T_{s2}, the spread of T_{s2} after it, the width it removes from the present 0.778 K, the spread of the log-conductance g_2 after it, from the present 0.224, and the expected information gain of equation (2.18) on the whole state.

Table 24.14: The preposterior of drawing B on record B for candidate readings, $\mathbf{\Sigma}_a = (\mathbf{\Sigma}_\mathrm{B}^{-1} + \mathbf{J}_a^\mathsf{T}\mathbf{J}_a)^{-1}$. Present spreads: $\sqrt{\Sigma_{T_{s2}T_{s2}}} = 0.778$ K, $\sqrt{\Sigma_{g_2g_2}} = 0.224$. Width removed: $0.778 - \sqrt{\Sigma_{a,T_{s2}T_{s2}}}$. Information gain by equation (2.18).

Candidate	Added on T_{s2} [K^{-2}]	Spread of T_{s2} [K]	Width removed [K]	Spread of g_2 [–]	Gain [nat]
present	–	0.778	–	0.224	–
hall-2 supply sensor, 0.1 K	100	0.099	0.678	0.158	2.06
hall-2 supply sensor, 0.3 K	11.1	0.280	0.498	0.167	1.02
hall-2 supply sensor, 0.5 K	4	0.421	0.357	0.179	0.62
second hall-2 sensor, 0.5 K	4	0.764	0.013	0.220	0.32
hall-2 load meter, 1.5 kW	0	0.765	0.013	0.220	0.62

A sensor on hall 2's own supply is the only candidate that reaches T_{s2} by itself: at 0.3 K it removes 0.50 of the 0.78 K, and because hall 2's reading then fixes the rise Q_2/G_2, it narrows the conductance as well, from 0.224 to 0.167 in the log. A second sensor in hall 2 adds the same information as hall 2's first sensor, on the sum $T_{s2} + Q_2/G_2$, and leaves the split where the priors put it; the load meter informs Q_2 and gains as much information on the state as the 0.5 K supply sensor, but almost none of it on T_{s2}. For the pump the expected support grows with exposure, as half the logarithm of the exposure: a record exactly at the handbook rate reaches $\ln B_{10} = -1$ at about 5500 h, -1.60 at 20 000 h, -2.73 at 200 000 h and -3 at about 340 000 h, three hundred and forty failures, while eight failures in the same 2000 h would give 3.8 nats against it.

24.6.3 The recommendation

The next measurement on the loop is a temperature sensor on hall 2's own supply branch, of 0.3 K or better. It was a weak discriminator of the drawings, 0.90 and 1.20 nats against the return sensor's 13.0 and 10.3, and it is the strong one for the quantity the settled drawing introduced, reducing the spread of T_{s2} from 0.78 to 0.28 K and with it the error that the set-point prior forces on hall 2's conductance. The hall's cooling is reported as missing its target of 10^{-3}: under independent units the probability of meeting it is 4.0×10^{-14} (1.5×10^{-11} exactly), and the direct count of the units-down column, which does not assume independence, reads 0.022. Its level is not reported as final: 7.34×10^{-3} (7.23×10^{-3} exactly) under independence, possibly

up to about three times that if units go down together, until a common-cause drawing of the four units is compared with the independent one on the same 400 days; no further days are needed for the structure. The pump's handbook rate may be used for planning only as a point that the record neither confirms nor refutes; to rely on it as the pump's own rate needs either years of exposure or a stated alternative narrower than the engineering prior.

24.6.4 What the analysis showed

The first reading took every drawing that fits as the drawing of the plant: the shared-loop fit passed its test, the pump's interval contained the handbook rate, and the design file's weights went into the availability report as they were. The analysis found that fitting settled none of the three: the loss log, read as the evidence of each structure, settles the cooling against the design file's hopes; one well-chosen reading settles the loop; and the pump's record cannot settle its question either way. On the cooling the loss log, not the design file, decides the structure, and it decides against the forgiving one, so the hall's cooling is N+1, and read with the whole log, units down as well as losses, it misses its target outright where the design file's weights left it a fair chance; the units-down column then hints past both structures to units that fail together, which the plant has to examine before it trusts the level of any unavailability built on independent units. On the loop the five readings that every fit of the earlier chapters used could not tell a shared supply from a separate one, and a single sensor on the return header, chosen because the two drawings predict it differently, decided the question; the sensor that now matters is a different one, on hall 2's own supply, because a settled structure still leaves its new node to the priors. On the pump, an interval that contains the handbook value is no evidence that the pump is the handbook's machine: such evidence costs years of operation, and a high count would refute the handbook within the same two thousand hours.

24.6.5 Limits of the model

The prior of the freed quantity. Every Bayes factor of the chapter is conditional on the priors of the quantities its drawings disagree on, and Lindley's paradox is the reason it must be: a factor for a freed quantity with an improper prior is zero, and with a very wide prior it favors the drawing without the quantity whatever the record. The remedy is the one the chapter applies: priors that state engineering knowledge, written before the record is seen, with the factor reported beside its sensitivity. A factor that changes sign over the plausible priors, as the pump's does at four failures, says that the record does not discriminate, not that either drawing is right.

Drawings outside the family. The posterior probabilities of equation (24.7) sum to one over the drawings written down. If the plant is a drawing outside the family, the comparison picks the candidate that predicts the record best, with a probability near one, and is wrong. Exercise 24.5 adds a third drawing to the loop that mimics drawing B on the return reading. The check against this is the test of each drawing by its own residuals, which reveals a record that no candidate explains, and the habit of generating candidates from the graph, as Chapter 23 did.

The assumptions under the closed forms. The cooling's evidence treats days as independent given the drawing and the nuisance probabilities, and the units-down column as independent of the drawing. A common cause that takes two units down and the cooling with them, the subject of Chapter 15, would couple the two columns and is a further drawing, not a nuisance; the units-down column already hints at it (Section 24.4.5). The unit posterior treats the 1600 unit-checks as independent trials with one shared unavailability, and the reported level

of the unavailability rests on that. A cause that lingers over several days, a failed controller left unrepaired, would make the losses of consecutive days dependent and overstate the evidence of each day; the pump's evidence assumes a constant rate, which an aging pump violates.

Common mistake

Four mistakes recur. Reading a Bayes factor as a posterior probability: $\ln B = 1$ is a probability of 0.73 only at even prior odds, and 0.86 from the design file's 0.7. Comparing evidences that do not share their constants: a drawing scored on fewer readings, or with a prior normalizer dropped in one drawing and kept in the other, produces a factor that measures the bookkeeping; the rules of Section 24.2.2 say which constants cancel. Freeing a quantity with a flat prior and reporting the factor: it is zero, or arbitrary through the units of the flat density. And taking a credible interval that contains the handbook value as confirmation of the handbook: the interval is an estimate inside the free drawing, and support for the point is a comparison of drawings (Section 24.5.2).

The next chapter turns from which drawing the plant is to what its failures cost: the expected consequence over the enumerated states of a plant, with its band, and the acceptance of a risk against a limit.

24.7 What to retain

Definition 24.1 (Evidence of a drawing and the Bayes factor between drawings)**.** The evidence of a drawing M is $p(\mathbf{y} \mid M) = \int p(\mathbf{y} \mid \mathbf{z}, M)\, p(\mathbf{z} \mid M)\, \mathrm{d}\mathbf{z}$, the probability of the record under the drawing with its unknowns integrated over their prior. Two drawings compared on the same record have the Bayes factor $B_{10} = p(\mathbf{y} \mid M_1)/p(\mathbf{y} \mid M_0)$, and the log posterior odds are $\ln B_{10}$ plus the log prior odds. The evidence reported is computed from the drawing's fit, its MAP state $\hat{\mathbf{z}}$, its information $\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$ and its covariance $\mathbf{\Sigma} = \mathbf{\Lambda}^{-1}$, by the Laplace formula

$$\ln p(\mathbf{y} \mid M) \approx -\mathcal{L}(\hat{\mathbf{z}}) + \tfrac{1}{2} \ln \det \mathbf{\Sigma} + \tfrac{n}{2} \ln 2\pi$$

plus the constants of the reading and prior rows, of which those common to both drawings cancel; beside it stands the closed form for conjugate rows, or the Savage–Dickey ratio $B_{01} = p(z_j^{\mathrm{h}} \mid \mathbf{y}, M_1)/p(z_j^{\mathrm{h}} \mid M_1)$ for a drawing nested at a point. A quantity freed by one drawing is charged by its Occam factor, the ratio $\sqrt{\Sigma_{\mathrm{nn}}}/\sigma_0$ of its posterior to its prior spread, and must carry a proper prior. When the drawing stays uncertain, a reported quantity is the mixture of its values under the drawings weighted by their posterior probabilities, with a between-drawing variance. The expected evidence a candidate reading adds is the divergence of the drawings' predictions of it, averaged over the drawings with their posterior probabilities.

Three conclusions carry beyond the three cases. A structure question is decided by a row the drawings predict differently, and the expected evidence of a candidate row, about half the squared separation of its predictions in units of their spread, ranks the candidates before any is installed: the return-header sensor decides whether hall 2 is on the shared loop by 14 and 11 nats. A posterior inside a drawing and the evidence between drawings answer different questions: the record can settle which drawing the plant is while leaving the quantity the drawing introduces, hall 2's own supply temperature, at the mercy of its prior, and a credible interval that contains the handbook rate is not evidence for the handbook. And the evidence depends on the prior of the freed quantity in a way no posterior does; the factor is reported with its sensitivity, and where it changes sign over plausible priors the record does not discriminate.

Table 24.15 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 24.15: The terms of model selection and Bayes factors and where each is defined.

Term	Expression	Units	Meaning	Defined in
Evidence	$p(\mathbf{y} \mid M) = \int p(\mathbf{y} \mid \mathbf{z}, M)\, p(\mathbf{z} \mid M)\, \mathrm{d}\mathbf{z}$	varies	probability of the record	24.2.1
Posterior model probability	$p(\mathbf{y} \mid M_d) \Pr(M_d) / \sum_e p(\mathbf{y} \mid M_e) \Pr(M_e)$	–	weight of a drawing	24.2.1
Log posterior odds	$\ln B_{10} + \ln[\Pr(M_1)/\Pr(M_0)]$	nat	Bayes factor and prior	24.2.1
Sequential form	$\ln B_{10} = \sum_k \ln[p(y_k \mid y_{<k}, M_1)/p(y_k \mid y_{<k}, M_0)]$	nat	a sum of predictions	24.2.1
Laplace evidence	$-\mathcal{L}(\hat{\mathbf{z}}) + \frac{1}{2}\ln\det\boldsymbol{\Sigma} + \frac{n}{2}\ln 2\pi - \sum \ln(\sqrt{2\pi}\sigma_k) + \ln c_\mathrm{f}$	–	fitted plant	24.2.2
Added-entry split	misfit $+ \frac{1}{2}\ln(\det \boldsymbol{\Lambda}_\mathrm{B}^\mathrm{old} / \det \boldsymbol{\Lambda}_\mathrm{A}) + \ln(\sigma_0/\sqrt{\Sigma_\mathrm{nn}})$	nat	where $\ln B$ comes from	24.3.1
Occam factor	σ_z/σ_0, one entry; $\sqrt{\Sigma_\mathrm{nn}}/\sigma_0$, an added entry	–	width left standing	24.2.3
Gaussian evidence	$\mathcal{N}(y;\ z^0,\ \sigma^2 + \sigma_0^2)$	varies	one reading, one freed entry	24.2.3
Negative-binomial evidence	$\frac{\Gamma(a+k)}{\Gamma(a)k!}(\frac{b}{b+T_\mathrm{tot}})^a(\frac{T_\mathrm{tot}}{b+T_\mathrm{tot}})^k$	–	counts, gamma prior	24.2.4
Handbook Bayes factor	$\frac{\Gamma(a+k)}{\Gamma(a)}\frac{b^a}{(b+T_\mathrm{tot})^{a+k}}\frac{\mathrm{e}^{\lambda_\mathrm{h} T_\mathrm{tot}}}{\lambda_\mathrm{h}^k}$	–	free rate against λ_h	24.3.1
Beta ratio	$\mathrm{B}(a+k,\ b+n-k)/\mathrm{B}(a,b)$	–	a sequence of trials	24.2.4
Loss-record evidence	Beta ratio over days below the threshold, times $\mathbf{1}\{k_\geq = n_\geq\}$	–	a structure with other causes	24.2.4
Savage–Dickey ratio	$B_{01} = p(z_j^\mathrm{h} \mid \mathbf{y}, M_1)/p(z_j^\mathrm{h} \mid M_1)$	–	nested point	24.2.5
Mixture over drawings	$\sum_d \Pr(M_d \mid \mathbf{y})\bar{f}_d$; within plus between variance	varies	report under an open drawing	24.3.2
Between-drawing variance	$\sum_d \Pr(M_d \mid \mathbf{y})(\bar{f}_d - \bar{f})^2$	varies	doubt about the drawing	24.3.2
Expected evidence gain	$\sum_d \Pr(M_d \mid \mathbf{y})\,\mathrm{KL}(d \,\|\, d')$, $\mathrm{KL} \approx \frac{1}{2}(\mu_\mathrm{A} - \mu_\mathrm{B})^2/\varsigma^2$, $\varsigma_d^2 = \mathbf{g}_d^\uparrow \boldsymbol{\Sigma}_d \mathbf{g}_d + \sigma_a^2$	nat	value of a reading for a comparison	24.3.3

24.8 Exercises

Exercise 24.1. *Formulation.* (a) The design file gives the 3-out-of-4 drawing a prior probability of 0.7, and a short record gives $\ln B_{32} = -1.5$. Write the question as posterior odds and compute the posterior probability of 3-out-of-4. (b) What log Bayes factor would raise a drawing with prior probability 0.3 to a posterior probability of 0.95? (c) One reading $y = 2$ with error $\sigma = 1$ is taken of a quantity that one drawing pins at 0 and another frees with the prior $\mathcal{N}(0, \sigma_0^2)$. For $\sigma_0 = 1$ and $\sigma_0 = 4$ write the state, its two rows and their slopes, compute the information Λ, the variance σ_z^2, $\hat{z}$, the two evidences, $\ln B_{10}$, and the three factors of equation (24.15). ((a) $\Pr(M_3 \mid \mathbf{y})/\Pr(M_2 \mid \mathbf{y}) = \mathrm{e}^{-1.5} \times 0.7/0.3 = 0.521$, so $\Pr(M_3 \mid \mathbf{y}) = 0.521/1.521 = 0.342$. (b) Posterior odds 19: $\ln B = \ln 19 + \ln(0.7/0.3) = 2.944 + 0.847 = 3.79$ nats. (c) The state is the one number z; the rows are $r_1 = (z-2)/1$ and $r_0 = z/\sigma_0$, with the slopes $j_1 = 1$ and $j_0 = 1/\sigma_0$; $p(\mathbf{y} \mid M_0) = \mathcal{N}(2;\ 0,\ 1) = 0.0540$. For $\sigma_0 = 1$: $\Lambda = 2$, $\sigma_z^2 = 0.5$, $\sigma_z = 0.707$, $\hat{z} = 1$; $p(\mathbf{y} \mid M_1) = \mathcal{N}(2;\ 0,\ 2) = 0.1038$, $\ln B_{10} = 0.653$; best fit 0.2420, Occam factor 0.707, misfit $\mathrm{e}^{-1/2} = 0.607$, product 0.1038. For $\sigma_0 = 4$: $\Lambda = 1.0625$, $\sigma_z = 0.970$, $\hat{z} = 1.882$; $\mathcal{N}(2;\ 0,\ 17) = 0.0860$, $\ln B_{10} = 0.466$; best fit 0.3962, Occam factor 0.243, misfit 0.895: the

wider prior fits better and is charged more.)

Exercise 24.2. *Formulation.* A record gives the two cooling drawings the weights 0.6 for 3-out-of-4 and 0.4 for 2-out-of-4. With the whole record's within-drawing moments of Section 24.3.2, compute the reported unavailability, the two parts of its variance, its standard deviation, the between-drawing share and the probability of meeting $U_{\mathrm{c}} < 10^{-3}$, given the drawings' probabilities 1.5×10^{-11} and 1.000. (Mean $0.6 \times 7.230 \times 10^{-3} + 0.4 \times 1.793 \times 10^{-4} = 4.41 \times 10^{-3}$; within $0.6 \times 3.214 \times 10^{-6} + 0.4 \times 4.69 \times 10^{-9} = 1.930 \times 10^{-6}$; between $0.6 \times 0.4 \times (7.051 \times 10^{-3})^2 = 1.193 \times 10^{-5}$; standard deviation $\sqrt{1.386 \times 10^{-5}} = 3.72 \times 10^{-3}$; share 0.861; probability $0.6 \times 1.5 \times 10^{-11} + 0.4 \times 1.000 = 0.400$.)

Exercise 24.3. *Derivation.* (a) Derive the Laplace evidence (24.12) with all its constants from the product of the row densities, and state when it is exact. (b) Derive the split (24.25) for a drawing that adds one entry with a prior row. (c) Apply (a) to one entry with one reading and recover equation (24.15). ((a) $p(\mathbf{y}, \mathbf{z} \mid M) = c_{\mathrm{f}} \prod_k (\sqrt{2\pi}\sigma_k)^{-1} \mathrm{e}^{-\mathcal{L}(\mathbf{z})}$; $\mathcal{L} \approx \mathcal{L}(\hat{\mathbf{z}}) + \frac{1}{2}\boldsymbol{\delta}^{\mathsf{T}}\boldsymbol{\Lambda}\boldsymbol{\delta}$ with $\boldsymbol{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$ at $\hat{\mathbf{z}}$; $\int \mathrm{e}^{-\mathcal{L}} \approx \mathrm{e}^{-\mathcal{L}(\hat{\mathbf{z}})} (2\pi)^{n/2} \det(\boldsymbol{\Sigma})^{1/2}$ with $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}^{-1}$; exact when every row is linear in the state, so that $\mathcal{L}$ is quadratic. (b) $\det \boldsymbol{\Lambda}_{\mathrm{B}} = \det \boldsymbol{\Lambda}_{\mathrm{B}}^{\mathrm{old}} / \Sigma_{\mathrm{nn}}$ by the partitioned determinant; the added $\frac{1}{2}\ln 2\pi$ cancels the $\ln\sqrt{2\pi}$ of the new prior row, leaving $-\ln\sigma_0 + \frac{1}{2}\ln\Sigma_{\mathrm{nn}}$ in $\ln p(\mathbf{y} \mid M_{\mathrm{B}})$. (c) $n = 1$, the state the number z, the slopes $j_1 = 1/\sigma$ and $j_0 = 1/\sigma_0$, $\Lambda = j_1^2 + j_0^2 = 1/\sigma^2 + 1/\sigma_0^2$, $\sigma_z^2 = 1/\Lambda$, constants $-\ln(2\pi\sigma\sigma_0)$: $\ln p(\mathbf{y} \mid M_1) = -\frac{(y-\hat{z})^2}{2\sigma^2} - \frac{(\hat{z}-z^0)^2}{2\sigma_0^2} + \frac{1}{2}\ln\sigma_z^2 + \frac{1}{2}\ln 2\pi - \ln(2\pi\sigma\sigma_0) = \ln \mathcal{N}(y;\ \hat{z},\ \sigma^2) + \ln(\sigma_z/\sigma_0) - \frac{(\hat{z}-z^0)^2}{2\sigma_0^2}$.)

Exercise 24.4. *Derivation.* (a) Derive the Savage–Dickey ratio (24.22) in general, name the step that uses condition (24.21), and verify it for the pump in the log-rate $\omega = \ln(\lambda/\lambda_{ref})$, where the handbook drawing is $\omega = 0$. (b) Derive the evidence (24.19) of the loss record and write $\ln B_{32}$ as a sum of seven logarithms. ((a) $p(\mathbf{y} \mid M_0) = \int p(\mathbf{y} \mid z_j^{\mathrm{h}}, \mathbf{z}_{-j})\, p(\mathbf{z}_{-j} \mid M_0)\, \mathrm{d}\mathbf{z}_{-j}$, and the condition replaces $p(\mathbf{z}_{-j} \mid M_0)$ by $p(\mathbf{z}_{-j} \mid z_j^{\mathrm{h}}, M_1)$, giving $p(\mathbf{y} \mid z_j^{\mathrm{h}}, M_1)$; Bayes' theorem for z_j at z_j^{h} then gives the ratio. In ω the densities are λ times those in λ, $10^{-3} \times 367.9 = 0.368$ and $10^{-3} \times 302.5 = 0.302$, and the change-of-variable factor cancels from the ratio, $\ln B_{10} = 0.196$. (b) Factor $p(\mathbf{y}_{\mathrm{down}}, \mathbf{y}_{\mathrm{loss}} \mid M) = p(\mathbf{y}_{\mathrm{down}})\, p(\mathbf{y}_{\mathrm{loss}} \mid \mathbf{y}_{\mathrm{down}}, M)$; given ν_{o} the days below the threshold give $\nu_{\mathrm{o}}^{k<}(1-\nu_{\mathrm{o}})^{n<-k<}$ and those above an indicator; the Beta integral gives $\mathrm{B}(4, 488)$ and $\mathrm{B}(11, 488)$, and $\ln B_{32} = \sum_{i=0}^{6} \ln[(492+i)/(4+i)] = 30.12$.)

Exercise 24.5. *Application.* A third drawing of the loop puts hall 2 on the shared loop, as drawing A does, with an unmetered bypass from the supply to the return header downstream of the flow meter. The bypass can only add flow, so the capacity rate at the return sensor is $C_w = \dot{m}c_p + C_{\mathrm{by}}$ with $\dot{m}c_p = 10\,\mathrm{kW/K}$ and a freed bypass capacity rate $C_{\mathrm{by}} \geq 0$, given the one-sided (half-normal) prior $p(C_{\mathrm{by}}) = 2\,\mathcal{N}(C_{\mathrm{by}};\ 0,\ 5^2)$ in kW/K. (a) Write its return row, say what in the evidence differs from drawing A, and why the Laplace formula alone will not do. (b) Its evidences give log Bayes factors against drawing A of -1.80 on record A and $+10.24$ on record B, and against drawing B of $+12.65$ and -1.05; the posterior of C_{by} has its mode at 0 on record A and at 5.8 kW/K on record B, where $2\mathcal{L} = 1.77$. Interpret them and name a reading that separates drawings B and C. (c) If the plant were drawing C and only A and B were compared, what would be reported? ((a) $r_{\mathrm{ret}}^{\mathrm{C}} = [T_s + (Q_1 + Q_2)/(\dot{m}c_p + C_{\mathrm{by}}) - y_{T_r}]/\sigma_{T_r}$; the reading rows and the three shared priors cancel, and what remains is the integral over C_{by} against its prior, the Occam charge of the freed bypass. The prior's mode sits on the boundary $C_{\mathrm{by}} = 0$, where record A also puts the posterior's, so the posterior is not a Gaussian around an interior mode; the evidence is computed as a one-dimensional integral over C_{by} of the Laplace evidence of the other five entries, the grid check of Section 24.4.1. (b) On record A drawing C reproduces A with no bypass and pays 1.8 nats for the freedom; on record B it explains the low return by a bypass of about 6 kW/K, beating A by 10.2 nats and losing to B by only 1.05: both predict a return near 15.1°C, so the return sensor cannot separate them. A flow meter on the

return header downstream of the bypass separates them (about 16 against 10 kW/K), as does a sensor on hall 2's supply (T_{s2} against T_s). (c) Drawing B with probability $1 - 1.3 \times 10^{-5}$, and drawing B's own residuals would pass ($2\mathcal{L} = 0.98$): the comparison is confidently wrong about a drawing it was not given, and nothing in the record warns of it.)

Exercise 24.6. *Application.* (a) Replace the pump's handbook point by a handbook interval, the drawing M_0' with $\lambda \sim \text{Gamma}(25,\ 25\,000\,\text{h})$, of mean λ_h and coefficient of variation 0.2, and compute its log Bayes factor against the free drawing for four and for eight failures in 2000 h. (b) A long record has 231 failures in 200 000 h. Compute its 95 % interval, the posterior probability below λ_h, and the log Bayes factors of the free drawing against the point and against the interval drawing, and explain why the interval and the factors disagree. ((a) By equation (24.16) for both drawings, $\ln B_{10'} = 0.198$ for four failures, against 0.196 for the point, and 3.33 for eight, against 3.82: a count near the handbook rate cannot tell a point from an interval, while a high count is absorbed in part by the interval's width. (b) Posterior Gamma(233, 201 000 h), interval $[1.015,\ 1.313] \times 10^{-3}\,\text{h}^{-1}$, which excludes λ_h; $\Pr(\lambda < \lambda_\text{h} \mid \mathbf{y}) = 0.015$; $\ln B_{10} = -0.38$ against the point and $\ln B_{10'} = -1.26$ against the interval. The interval measures the distance of λ_h from the estimate in posterior standard deviations, which a long record shrinks; the factor compares how well the handbook and the broad engineering prior predicted the record, and the engineering prior spent its mass on rates the record excludes. The rate is not exactly the handbook's, and the handbook is still the better prediction: Lindley's paradox.)

Part VII

Risk, Importance and Acceptance

Chapter 25

Risk and Expected Consequences

25.1 The expected loss of the hall's cooling

The owner of the data hall of Chapter 14 is asked by the operator of the computing load to state that the hall's cooling will cost it less than R$ 1.5 million a year, with an assurance of 0.9, and the facilities engineer is asked for the figure. The records are the three series of availability checks of that chapter: of a hundred checks of the four CRAC units, three found a unit down; of four hundred checks of the two chilled-water pumps, four; of a thousand checks of the control system, one. The operator's accounts price each condition of the hall by the hour: an hour without cooling sheds the computing load at R$ 20 000, an hour with one CRAC unit down costs R$ 150 in lost margin and fan power, and an hour with one pump down R$ 50. The first reading takes the fractions found down as the units' unavailabilities, finds the hall without cooling 55 hours a year and with a unit down about an eighth of the time, and prices the hours: R$ 1.25 million a year, below the limit, so the statement can be signed.

The engineer who must sign it asks what that figure rests on. The fractions come from three, four and one units found down, which fix no unavailability closely; the loss of cooling needs two CRAC units down together, so the uncertain CRAC unavailability enters the cost squared; and the cost of one particular year depends on how many losses of cooling the year happens to contain, which the rate alone does not say. The figure at the fractions settles none of this. The question of this chapter is the owner's, stated in full:

what is the expected consequence: the expected loss per year of the hall's cooling, summed over every configuration of its units with each weighted by its probability and by its cost, with the band that the evidence leaves on it; whether it lies below a stated limit with a stated assurance; what a bad year costs; and how much each design change removes.

What changes from Chapters 13 and 14 is the weight given to each configuration: the enumeration and the posterior of three logits are the same, and every configuration is now weighted by its cost per hour as well as by its probability, so that the derived quantity is an expected consequence in reais rather than a probability.

Jargon: probabilistic risk assessment, risk, expected loss, consequence, consequence model, loss schedule, state class, IT load shed, risk-informed acceptance, assurance, plug-in risk, risk curve, exceedance probability, value at risk, conditional value at risk, tail mean, compound Poisson loss

A *probabilistic risk assessment* (PRA) answers three questions of a system: what can go wrong (the scenarios), how likely each is (their probabilities) and what each costs (their *consequences*); the *risk* is the probability-weighted sum of the consequences over mutually exclusive scenarios, an *expected loss*. A *consequence model* is the rule that assigns a cost to each scenario; when the operator supplies it as a table it is a *loss schedule*. A *state class* is a group of configurations with the same kind of consequence: loss of the service, degraded service, nominal service. *IT load shed* is the switching off of computing load when a hall can no longer be cooled. *Risk-informed acceptance* accepts a design or a plant when its risk lies below a limit with a stated posterior probability, the *assurance*. The *plug-in risk* is the risk evaluated at the posterior means of the unknowns. The *risk curve* is the complementary cumulative distribution of a loss, the *exceedance probability* $\Pr(X > x)$ drawn against the loss x. The *value at risk* at level p is the p-quantile of a loss distribution; the *conditional value at risk*, or *tail mean*, is the mean of the loss in the tail beyond it. A *compound Poisson loss* is a sum of a Poisson number of random amounts, here one amount per loss of cooling.

Key idea

The risk is a derived quantity of the state: a sum over the enumerated configurations of their probabilities times their consequences, a polynomial in the units' unavailabilities that is linear in the consequences. Its expectation under a shared posterior exceeds its plug-in value when coincidences of units dominate the loss, by half the trace of the posterior covariance of the unavailabilities times the risk's curvature; its band and its acceptance probability come from the same posterior as any other derived quantity, and acceptance is a tail probability of the risk, not a comparison of its mean with a limit. The posterior of the risk says how well the loss rate is known; the loss of one year is a second, aleatory variable around that rate, and the two must be reported apart.

25.1.1 Risk as an expected loss

Probabilistic risk assessment writes the expected loss of a plant as a sum over mutually exclusive scenarios of probability times consequence, the expected consequence of Chapter 2, equation (2.23), with the scenarios as its events. The companion volume *Reliability Engineering* develops the method in its chapters on risk and on event trees, with scenarios built from an initiating event and the branches that follow it, and its vocabulary is used here. For a plant of independent units already enumerated, the scenarios need no construction: the 2^n configurations of Chapter 13 are mutually exclusive and exhaust the possibilities, and each one has a consequence. The loss of cooling is the union of 113 of the 128 configurations of the hall, the top event of Chapter 14; the risk weighs those 113 by the cost of a lost hall and the other fifteen by their own, smaller costs. Each configuration's probability depends on unavailabilities known only through their records, so the risk has a posterior, and the cost incurred in one particular year is a second, aleatory quantity around it.

25.1.2 The physical system and its graph (step 0)

The plant is the one of Chapter 14, drawn in Figure 14.1: four CRAC units C1 to C4, of which three must run; two chilled-water pumps P1 and P2, of which one must run; and one control system K, which must run. Units of one design checked under one record share one entry, so

the state has three entries, the logits of the three shared unavailabilities,

$$\mathbf{z} = \begin{bmatrix} \xi_C \\ \xi_P \\ \xi_K \end{bmatrix}, \qquad q_j = \frac{1}{1 + \mathrm{e}^{-\xi_j}}, \quad j \in \{C, P, K\}, \tag{25.1}$$

the state of equation (14.2), while the enumeration runs over the configurations of the seven units. In this chapter $\mathbf{q} = (q_C, q_P, q_K)^{\mathsf{T}}$ denotes the three shared unavailabilities, of which every unit's unavailability is one. The records are those of Section 14.1.5: $k_C = 3$ units found down in $n_C = 100$ checks of the CRAC units, $k_P = 4$ in $n_P = 400$ checks of the pumps and $k_K = 1$ in $n_K = 1000$ checks of the controller, each record standing alone, with no prior drawn from other halls. The evidence $\mathbf{y}$ is the three records.

The structure is that of the fault tree of Figure 14.2: the top event, the loss of hall cooling, occurs when two or more CRAC units are down, when both pumps are down, or when the controller is down. Its probability given the unavailabilities is equation (14.26),

$$U_{\mathrm{sys}}(\mathbf{q}) = 1 - R_{\mathrm{C}}(q_C)\,(1 - q_P^2)\,(1 - q_K), \qquad R_{\mathrm{C}}(q) = (1-q)^3(1+3q), \tag{25.2}$$

with R_{C} the probability that at most one CRAC unit is down. The plant graph is not extended for the risk. Step 0 was paid once, and the risk is a new question asked of the same graph and the same posterior.

25.1.3 The consequence schedule

The operator's loss schedule gives a cost per hour for each class of configuration. Table 25.1 lists it: each row is a class, and the columns give the condition that defines it, the number of the 128 configurations it contains, its consequence rate and what the rate stands for. When the cooling is lost the hall sheds its computing load, at R\$ 20 000 h^{-1}. When the cooling is kept with one CRAC unit down the hall runs without its margin and the other three units run harder, at R\$ 150 h^{-1} for each unit down; with one pump down it runs without the pumps' redundancy and with a small change in pumping energy, at R\$ 50 h^{-1}. With every unit up the cost is zero. In a configuration in which the cooling is kept, at most one CRAC unit and at most one pump can be down, so the degraded configurations have the three costs R\$ 150, R\$ 50 and R\$ 200 h^{-1}. The owner's limit on the annual cost is $\mathcal{R}^* =$ R\$ 1.5 million, to be met with the assurance $\kappa = 0.9$, and it is compared in the questions below with the limits of R\$ 1.0 and R\$ 2.0 million as well.

Table 25.1: The consequence schedule of the hall's cooling: the classes of configurations, their number among the $2^7 = 128$ and the consequence rate the operator assigns to each.

Class	Configurations [–]	Consequence rate [R\$/h]	Stands for
loss of cooling (top event)	113	20 000	IT load shed
one CRAC unit down, pumps up	4	150	lost margin, fan power
one pump down, CRAC units up	2	50	lost redundancy, pumping
one CRAC unit and one pump down	8	200	both of the above
nominal, every unit up	1	0	–
total	128		

The consequence function. Write $m_{\mathrm{C}}(\mathbf{s})$ and $m_{\mathrm{P}}(\mathbf{s})$ for the numbers of CRAC units and of pumps down in configuration $\mathbf{s}$, and $\phi(\mathbf{s})$ for the structure function, one when the hall is cooled. The schedule is the function

$$C(\mathbf{s}) = C_{\mathrm{L}}\bigl(1 - \phi(\mathbf{s})\bigr) + \phi(\mathbf{s})\Bigl[C_{\mathrm{C}}\, m_{\mathrm{C}}(\mathbf{s}) + C_{\mathrm{P}}\, \mathbf{1}\bigl(m_{\mathrm{P}}(\mathbf{s}) \geq 1\bigr)\Bigr], \tag{25.3}$$

with $C_{\mathrm{L}} = \mathrm{R\$}\,20\,000\,\mathrm{h}^{-1}$, $C_{\mathrm{C}} = \mathrm{R\$}\,150\,\mathrm{h}^{-1}$ and $C_{\mathrm{P}} = \mathrm{R\$}\,50\,\mathrm{h}^{-1}$. It is a consequence *rate*, a cost per hour spent in the configuration, so that the cost of a stay is proportional to its length.

What a consequence model is and is not. The schedule is a set of pinned inputs: numbers taken from the operator's accounts, with no uncertainty attached in this chapter. It maps each configuration to a cost and adds no row to the graph and no entry to the state. It is not a model of how long a loss lasts: the duration is carried by the unavailability, which is a long-run fraction of time, and the consequence is per hour of it. It is not a function of the load, the season or the time of day, so a loss at full load costs the same as one at night. And it is not a model of recovery: shedding load, restarting it and the damage a hot hall may do to equipment are folded into one rate. A later chapter on repair and replacement decisions frees the costs as uncertain quantities; Section 25.6.6 states what the pinned schedule leaves out.

Figure 25.1 attaches the schedule to the tree. On the left is the fault tree of Figure 14.2: the top event at the top, its OR gate, the voting gate 2/4 on the four CRAC units, the AND gate on the two pumps and the controller entering directly. On the right are the three classes of configuration, each with its number of configurations and its consequence rate: the loss of cooling, reached when the top event occurs; and, when it does not, the degraded and the nominal configurations.

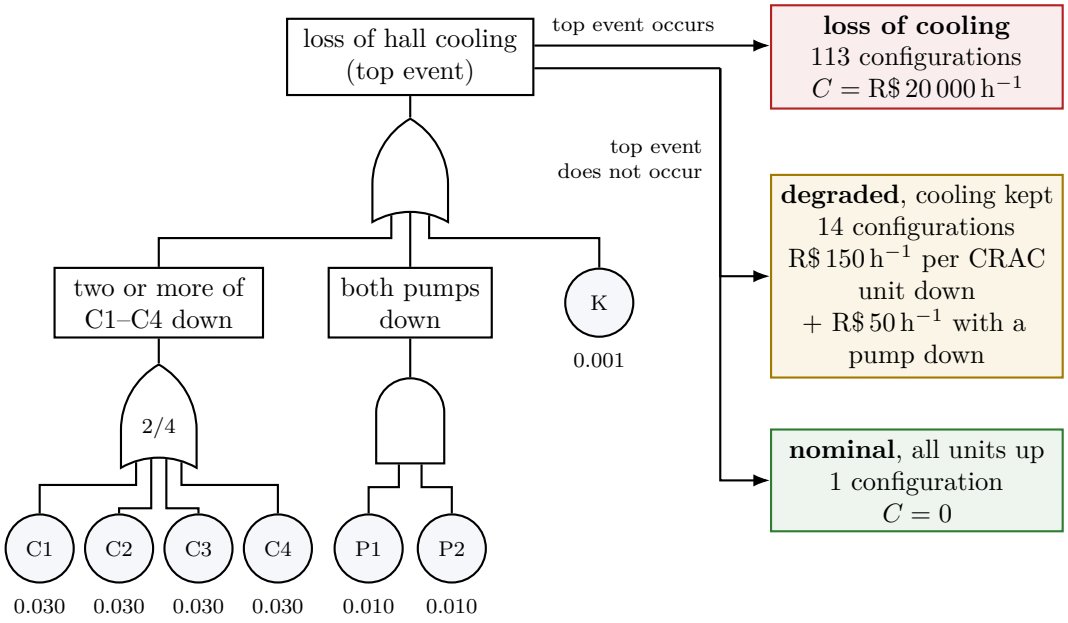


Figure 25.1: The fault tree of the loss of hall cooling (left, as in Chapter 14, with the fraction of each group's checks that found a unit down under its basic events) and the three classes of configuration it separates, with the consequence rate of each (right): the 113 configurations of the top event cost $\mathrm{R\$}\,20\,000\,\mathrm{h}^{-1}$; of the fifteen in which the hall is cooled, fourteen have a unit down and cost R$ 50 to $\mathrm{R\$}\,200\,\mathrm{h}^{-1}$, and one has every unit up.

25.1.4 The probabilistic questions (step 1)

The owner asks six questions of the schedule and the three records, and step 1 writes each as a quantity over the state, with the unknown on the left of the bar and the records on the right. Table 25.2 lists them: the first column is the question in engineering language; the second the general form of its category in the notation of Chapter 2, for a quantity $f(\mathbf{z})$ derived from the state, a level f^* of it, the loss $L(a, \mathbf{z})$ of an action a and its expected value $\mathcal{R}(a) = \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$; the third its mathematical form for the hall; the fourth its category; and the last the subsection where it is answered. The first is the owner's question as stated in Section 25.1, an estimation of a derived quantity whose computation is the expected consequence. The second compares two summaries of that quantity, its value at the posterior means and its expectation, and asks why they differ. The third attributes the quantity to classes of configuration and to cut sets, a decomposition of the same sum. The fourth is risk-informed acceptance, a prediction under a threshold on the derived quantity. The fifth asks for a tail of the derived quantity and, with it, of the loss of one year, the predictive of a future outcome; it separates the uncertainty of the loss rate from the variability of one year's loss. The sixth is design by alternative drawings, as in Chapter 14: a change to the plant brings no new evidence and is valued on the present posterior, whereas the next record of step 7 values evidence not yet taken. The atlas of Appendix A lists all but the second.

25.1.5 The evidence as rows (step 2)

Step 2 attaches the same rows as in Chapter 14: each check record is one count row on the logit of its group, the hundred CRAC checks on ξ_C, the four hundred pump checks on ξ_P and the thousand controller checks on ξ_K. Which row acts on which entry is a property of the graph, and it fixes where the Jacobian of the rows can be nonzero, marked $*$:

$$\mathbf{J} = \begin{bmatrix} * & 0 & 0 \\ 0 & * & 0 \\ 0 & 0 & * \end{bmatrix} \begin{matrix} \text{CRAC checks} \\ \text{pump checks} \\ \text{controller checks} \end{matrix}, \qquad \text{columns } \xi_C, \xi_P, \xi_K, \tag{25.4}$$

the pattern of equation (14.3). No row acts on two entries, at any state, and no row is a physics row, so no model tolerance σ_p arises. The consequence schedule adds no row: it is a pinned function of the configuration, equation (25.3), and the risk is a derived quantity of the state built from it and from the enumeration. Every question of Table 25.2 is a question about that derived quantity, its parts or its tails.

The modeling choices. Four choices make the model of the risk, and each buys something and can fail. The schedule of Table 25.1 is taken as a set of pinned costs per hour: this makes the risk linear in the costs and adds no entry to the state, and it misleads if the costs are uncertain or depend on the load, which Section 25.4.4 varies and Section 25.6.6 discusses. The plant is taken to be in steady state through the year, each unavailability the long-run fraction of time its units are down: this turns an expected cost per hour into a loss per year by one factor t_{yr}, and it fails if the unavailabilities change with the season or with age, the case of Exercise 25.5. The units are taken as down independently given their unavailabilities, as in Chapter 14: this makes every configuration's probability a product and the risk a polynomial in three entries, and it fails under a cause that takes two units down at once (Section 25.6.6). Each group's posterior is its record alone: this lets the records set the answer, and with three, four and one units found down it makes the level of the answer depend on any prior added, which Section 25.4.3 measures.

The rows of the check records. A record alone is a flat prior on the logit, so each entry has one row, the count of units found down. With k_j units found down in n_j checks of group j,

Table 25.2: The questions asked of the hall's cooling with its consequence schedule, the general form of their category, their mathematical form for the hall, their category and where each is answered. $\mathbf{y}$ is the three check records; $\mathcal{R}_{\mathrm{yr}}$ the annual risk, the expected loss of a year given the unavailabilities; $\mathcal{R}^{*}$ a limit on it and κ the assurance; L_{yr} the loss realized in one year; $L(a, \mathbf{z})$ the loss of action a given the state, here the annual risk under design a, and $\mathcal{R}(a)$ its posterior expectation.

Engineering question	General form	Mathematical form	Category	Answered in
What is the expected consequence? What is the expected loss per year of the hall's cooling, and how well is it known?	$\mathcal{R}(a) = \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$, quantiles	$\mathbb{E}[\mathcal{R}_{\mathrm{yr}}(\mathbf{q}) \mid \mathbf{y}]$, $\mathcal{R}_{\mathrm{yr}} = 8760\,\mathrm{h} \times \sum_{\mathbf{s}} \Pr(\mathbf{s} \mid \mathbf{q})\, C(\mathbf{s})$, with its central 95 % band	estimation of a derived quantity (an expected consequence)	25.3.1
What is the plug-in risk, and why is it below the expected risk?	$L(a, \hat{\mathbf{z}})$ against $\mathcal{R}(a)$	$\mathcal{R}_{\mathrm{yr}}(\bar{\mathbf{q}})$ against $\mathbb{E}[\mathcal{R}_{\mathrm{yr}} \mid \mathbf{y}]$; the gap	estimation (two summaries compared)	25.3.1
Where does the risk come from: the loss states or the degraded states, and which cut sets?	$\mathbb{E}[f_k(\mathbf{z}) \mid \mathbf{y}]$ for the parts f_k of f, quantiles	the sum split by state class and by minimal cut set	estimation (attribution)	25.3.2
Is the annual loss below R$ 1.5 million with an assurance of 0.9?	$\Pr(f(\mathbf{z}) \geq f^{*} \mid \mathbf{y}) \geq \kappa$	$\Pr(\mathcal{R}_{\mathrm{yr}}(\mathbf{q}) < \mathcal{R}^{*} \mid \mathbf{y}) \geq \kappa$, $\mathcal{R}^{*} = $ R$ 1.5 million, $\kappa = 0.9$	prediction under a threshold (risk-informed acceptance)	25.3.3
What is the loss in a bad year?	$\int p(y_{\mathrm{new}} \mid \mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}$, a quantile	the 95th percentile and the tail mean of $\mathcal{R}_{\mathrm{yr}}$ and of the predictive L_{yr}; the risk curve	prediction under a threshold (a quantile of a predictive)	25.3.4
How much risk does each design change remove?	$\mathcal{R}(a)$ for each action a, the drawings ranked; $\Pr(L(a, \mathbf{z}) < L(a', \mathbf{z}) \mid \mathbf{y})$	$\mathbb{E}[\Delta\mathcal{R}_{\mathrm{yr}}(\mathbf{q}; a) \mid \mathbf{y}]$ for a fifth CRAC unit, a redundant controller and both, with bands	design, by alternative drawings	25.3.5

each row puts its negative log-likelihood into the objective, and each enters Gauss–Newton in the working form of a count row, the binomial-in-the-logit pair of Chapter 4, equation (4.43), written for this state in Chapter 14, equations (14.4) and (14.5):

$$\begin{aligned} \mathcal{L}(\mathbf{z}) &= \sum_{j\in\{C,P,K\}} \mathcal{L}_j(\mathbf{z}), \qquad \mathcal{L}_j(\mathbf{z}) = -k_j \ln q_j - (n_j - k_j)\ln(1-q_j), \\ r_j(\mathbf{z}) &= \frac{n_j q_j - k_j}{\sqrt{n_j q_j (1-q_j)}}, \qquad \mathbf{J}(\mathbf{z}) = \operatorname{diag}\Big(\sqrt{n_j q_j (1-q_j)}\Big), \end{aligned} \tag{25.5}$$

three rows on three entries and a 3×3 Jacobian with the pattern (25.4). The gradient of the objective is $\mathbf{J}^\mathsf{T} r$, and its curvature is $\mathbf{J}^\mathsf{T}\mathbf{J} + \mathbf{S}$ with $\mathbf{S}$ the curvature term of the count rows; the logit is the canonical parameter of the binomial row, so $\mathbf{S} = 0$ at every state and the curvature equals $\mathbf{J}^\mathsf{T}\mathbf{J}$ exactly. With the rows normalized, as here, the information matrix is $\mathbf{\Lambda} = \mathbf{J}^\mathsf{T}\mathbf{J}$, with diagonal entries $n_j q_j (1-q_j)$.

25.2 The posterior of the state (step 3)

The six questions of Table 25.2 rest on one derived quantity, the risk, and one posterior, the MAP state $\hat{\mathbf{z}}$ and the covariance $\mathbf{\Sigma}$ of the fit, with the gradient $\mathbf{g}$ of a derived quantity. The problem also admits an exact posterior, three independent Beta posteriors, one per record: their moments give every expectation of the risk in closed form, because the risk is a polynomial in the unavailabilities, and their draws give its quantiles and tail probabilities. This section derives the posterior and computes it for the hall; Section 25.3 computes every answer from it.

25.2.1 The posterior from the rows

The fit solves the normal equations $(\mathbf{J}^\mathsf{T}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^\mathsf{T} r$, equation (14.6), three separate equations because $\mathbf{J}^\mathsf{T}\mathbf{J}$ is diagonal. The gradient vanishes where every working residual is zero, at $q_j = k_j/n_j$, so the MAP state and the Laplace covariance are those of equation (14.7),

$$\hat{z}_j = \ln\frac{k_j}{n_j - k_j}, \qquad \mathbf{\Sigma} = \operatorname{diag}\Big(\frac{1}{k_j} + \frac{1}{n_j - k_j}\Big), \tag{25.6}$$

with zero off-diagonal entries: the three posteriors are independent, a result of the pattern (25.4) and not an assumption. The problem also admits the exact posterior, equation (14.8),

$$q_j \mid \mathbf{y} \sim \operatorname{Beta}(k_j,\ n_j - k_j), \qquad \mathbb{E}[q_j \mid \mathbf{y}] = \frac{k_j}{n_j} = q(\hat{z}_j), \tag{25.7}$$

independently across the groups, so that the plug-in values are the posterior means and the unavailabilities at the MAP state. Every expectation of a product of powers of the three unavailabilities is a product of the moments of equation (14.20), and every expectation of the risk with it, since the risk is such a polynomial; a quantity of several entries has its band read from draws of the three posteriors; and the exact variance of a logit is the trigamma sum of equation (14.9), the counterpart of Σ_{jj}.

25.2.2 The posterior of the hall

The hall's inputs are the three check records, three CRAC units found down in 100 checks, four pumps in 400 and one controller in 1000.

The fit is that of Chapter 14. Its Table 14.2 runs the normal equations from the starting state $\mathbf{z}_0 = (-4,\ -4,\ -4)^\mathsf{T}$, and the CRAC and pump entries settle in four steps and the

controller's in six. At the MAP state every working residual of equation (25.5) is zero, and the Jacobian, the information matrix and the Laplace covariance there are, by equation (14.10),

$$\hat{\mathbf{z}} = \begin{bmatrix} -3.4761 \\ -4.5951 \\ -6.9068 \end{bmatrix}, \qquad \mathbf{J}(\hat{\mathbf{z}}) = \begin{bmatrix} 1.7059 & 0 & 0 \\ 0 & 1.9900 & 0 \\ 0 & 0 & 0.9995 \end{bmatrix},$$
$$\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J} = \begin{bmatrix} 2.9100 & 0 & 0 \\ 0 & 3.9600 & 0 \\ 0 & 0 & 0.9990 \end{bmatrix}, \qquad \mathbf{\Sigma} = \mathbf{\Lambda}^{-1} = \begin{bmatrix} 0.3436 & 0 & 0 \\ 0 & 0.2525 & 0 \\ 0 & 0 & 1.0010 \end{bmatrix}. \tag{25.8}$$

The MAP state maps to $\bar{q}_C = 0.030$, $\bar{q}_P = 0.010$ and $\bar{q}_K = 0.001$, the diagonal of $\mathbf{\Sigma}$ is $1/k_j + 1/(n_j - k_j)$ by equation (25.6), and the logits are read off with $\sqrt{\Sigma_{jj}} = 0.586$, 0.503 and 1.000. The exact posteriors are $q_C \sim \text{Beta}(3,\ 97)$, $q_P \sim \text{Beta}(4,\ 396)$ and $q_K \sim \text{Beta}(1,\ 999)$, Table 14.3, with the posterior means at the same values; the central 95 % interval of q_C, [0.0063, 0.0711], spans a factor of eleven.

The pair $(\hat{\mathbf{z}}, \mathbf{\Sigma})$ is what step 4 uses: every answer of the chapter is computed from it. Because this problem also admits the exact posterior, each answer has an exact value beside it, from the moments or the draws of the three Beta posteriors, and the gap between the two is reported with it.

25.3 The answers (step 4)

Given the state, the expected loss of a year is the annual risk, the loss $L(a, \mathbf{z})$ of equation (2.23) for the present design, and its posterior expectation is the expected consequence $\mathcal{R}(a)$. Each question is answered from the posterior of the state, Section 25.2: its engineering question, the general form of its category, the form it takes for the hall, the quantities of the posterior it uses, and the hall's numbers. The questions add the tree of Figure 14.2 and the trees of the three alternative designs; the schedule of Table 25.1; the year of $t_{\text{yr}} = 8760\,\text{h}$; the mean down time of a loss of cooling, MDT = 12 h; and, for the quantities without a closed form, 2×10^5 triples (q_C, q_P, q_K) drawn from the three Beta posteriors, seed 24, each passed through equations (25.2) and (25.14), for the exact values beside the answers. The computation needs only standard functions, Beta moments, quantiles and draws and Poisson probabilities, and its operations are these.

1. *Checks of the inputs.* Every configuration has a consequence, the consequences are non-negative, the loss configurations are exactly those of the top event, and the configuration probabilities at the posterior means sum to one.
2. *Fit.* The normal equations of the three count rows, for $\hat{\mathbf{z}}$ and $\mathbf{\Sigma}$; the exact Beta posteriors beside them.
3. *Risk.* The sum (25.10) over the 2^n configurations, times t_{yr}; its split by class, equation (25.13), and by cut set, equations (25.25) and (25.26).
4. *Expectation.* Each configuration's expected probability as a product of group moments, equation (25.15); the plug-in value at the means; the gap.
5. *Band and acceptance.* The risk on each posterior draw; its quantiles and the probabilities (25.27) at each limit.
6. *The year's loss.* The frequency (25.29) on each draw, a Poisson count of losses, the loss (25.30); the risk curves, values at risk and tail means (25.34) of the three distributions.
7. *Designs.* Steps 3 to 5 for each alternative tree on the same draws, and the reductions (25.35) draw by draw.

Every answer is computed from the MAP state $\hat{\mathbf{z}}$ and the covariance $\mathbf{\Sigma}$ of step 3, by the second-order expectation, the delta method on the logarithm, the normal tail, or a sum over

the state, and that is the value the chapter reports; the exact Beta posteriors give the value set beside it, with the gap.

25.3.1 The expected annual risk and the plug-in risk

Engineering questions. What is the expected consequence: what is the expected loss per year of the hall's cooling, and how well is it known? What is the plug-in risk, and why is it below the expected risk? The first asks for the cost the records imply with its band; the second for the figure of the first reading, the cost at the fractions found down, and for the reason it falls short of the first.

General form. Both are the expected consequence of Chapter 2, equation (2.23): the loss of an action given the state, probability times consequence summed over the events, and its posterior expectation,

$$\mathcal{R}(a) = \mathbb{E}[L(a,\mathbf{z}) \mid \mathbf{y}] = \int L(a,\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \qquad L(a,\mathbf{z}) = c_a + \sum_e C_e \Pr(e \mid \mathbf{z}, a),$$

reported with its central 95 % interval, the estimation form of equation (2.11). The plug-in value is the loss at one state, $L(a,\hat{\mathbf{z}})$, which by equation (2.24) misses the curvature term $\frac{1}{2}\operatorname{tr}(\mathbf{H}_a\boldsymbol{\Sigma})$, with $\mathbf{H}_a$ the Hessian of the loss in the state.

Specialized form. The questions are equations (2.23) and (2.11) with the state $\mathbf{z}$ the three logits, the events e the configurations, C_e the year's cost of each, $c_a = 0$ for the present design, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of equation (25.8). The four modeling choices of Section 25.1.5 enter here.

From a cost rate to a loss per year. Let $\mathbf{S}(t)$ be the configuration of the units at time t and $C(\mathbf{S}(t))$ the cost rate it incurs. Over a year of $t_{\mathrm{yr}} = 8760\,\mathrm{h}$ the cost incurred is the integral of the rate,

$$L_{\mathrm{yr}} = \int_0^{t_{\mathrm{yr}}} C(\mathbf{S}(t))\, \mathrm{d}t,$$

a random quantity even when the unavailabilities are known. Its expectation given $\mathbf{q}$ follows in three lines:

$$\begin{aligned}
\mathbb{E}[L_{\mathrm{yr}} \mid \mathbf{q}] &= \int_0^{t_{\mathrm{yr}}} \mathbb{E}[C(\mathbf{S}(t)) \mid \mathbf{q}]\, \mathrm{d}t && \text{linearity of the expectation} \\
&= \int_0^{t_{\mathrm{yr}}} \sum_{\mathbf{s}} \Pr(\mathbf{S}(t) = \mathbf{s} \mid \mathbf{q})\, C(\mathbf{s})\, \mathrm{d}t && \text{expectation of a function of } \mathbf{S}(t) \\
&= t_{\mathrm{yr}} \sum_{\mathbf{s}} \Pr(\mathbf{s} \mid \mathbf{q})\, C(\mathbf{s}). && \text{steady state} \qquad (25.9)
\end{aligned}$$

The last line uses the meaning of the unavailabilities: each q_i is the long-run fraction of time unit i is down, and under independence given $\mathbf{q}$ the probability $\Pr(\mathbf{s} \mid \mathbf{q}) = \prod_i q_i^{1-s_i}(1-q_i)^{s_i}$ of Chapter 13 is the long-run fraction of time spent in configuration $\mathbf{s}$, the same at every instant of a plant in steady state.

The risk. The sum in equation (25.9) is the risk per hour,

$$\mathcal{R}(\mathbf{q}) = \sum_{\mathbf{s} \in \{0,1\}^n} \Pr(\mathbf{s} \mid \mathbf{q})\, C(\mathbf{s}) \qquad [\mathrm{R\$\,h^{-1}}], \qquad (25.10)$$

the expected cost rate at a random instant. The consequence $C(\mathbf{s})$ and the risk $\mathcal{R}(\mathbf{q})$ are rates, in reais per hour; the quantities per year carry the subscript yr, and $t_{\mathrm{yr}} = 8760\,\mathrm{h}$ is the length of the year. The annual risk is

$$\mathcal{R}_{\mathrm{yr}}(\mathbf{q}) = t_{\mathrm{yr}}\, \mathcal{R}(\mathbf{q}) = \mathbb{E}[L_{\mathrm{yr}} \mid \mathbf{q}] \qquad [\mathrm{R\$\,yr^{-1}}], \qquad (25.11)$$

the expected loss of a year given the unavailabilities. Both are functions of the state through $\mathbf{q}$, with C pinned. Equation (25.10) is the probabilistic risk assessment of the plant: the configurations are its scenarios, the enumeration gives their probabilities, and the schedule their consequences. In the terms of the general form the events are the configurations, C_e is the year's cost $t_{\mathrm{yr}}C(\mathbf{s})$ of configuration $\mathbf{s}$, and the action is the present design with no certain cost, $c_a = 0$, so that the loss given the state and the expected consequence are

$$L(a, \mathbf{z}) = \mathcal{R}_{\mathrm{yr}}(\mathbf{q}(\mathbf{z}); a), \qquad \mathcal{R}(a) = \mathbb{E}\big[\mathcal{R}_{\mathrm{yr}}(\mathbf{q}; a) \mid \mathbf{y}\big], \tag{25.12}$$

both in reais per year. The loss $L(a, \mathbf{z})$ is not the loss L_{yr} of one particular year: L_{yr} stays random when the state is known, and $L(a, \mathbf{z})$ is its expectation given the state, equation (25.11); Section 25.3.4 gives L_{yr} its distribution. The letter $\mathcal{R}$ with the argument $\mathbf{q}$ names the expected cost given the state, a derived quantity with a posterior, and with the argument a its posterior expectation; the design argument is dropped where only the present design is meant.

The sum by class. Splitting the sum by the classes of Table 25.1, the loss configurations all cost C_{L}, and their probabilities sum to the top event:

$$\mathcal{R}(\mathbf{q}) = C_{\mathrm{L}}\, U_{\mathrm{sys}}(\mathbf{q}) + \mathcal{R}_{\mathrm{deg}}(\mathbf{q}), \qquad \mathcal{R}_{\mathrm{deg}}(\mathbf{q}) = \sum_{\mathbf{s}:\ \phi(\mathbf{s})=1} \Pr(\mathbf{s} \mid \mathbf{q})\, C(\mathbf{s}). \tag{25.13}$$

The first term is the loss-state term, the cost rate of a lost hall times the probability of the top event; the second is the degraded term. For the hall the degraded term has a closed form. The cost of a cooled configuration is the sum of a CRAC part, $C_{\mathrm{C}}m_{\mathrm{C}}$, and a pump part, $C_{\mathrm{P}}\mathbf{1}(m_{\mathrm{P}} \geq 1)$, and the expectation of a sum is the sum of the expectations over the cooled configurations. The CRAC part is nonzero only when exactly one CRAC unit is down, with probability $4q_C(1-q_C)^3$, the pumps are not both down, $1-q_P^2$, and the controller is up, $1-q_K$; the pump part only when exactly one pump is down, $2q_P(1-q_P)$, at most one CRAC unit is down, $R_{\mathrm{C}}(q_C)$, and the controller is up. Hence

$$\mathcal{R}_{\mathrm{deg}}(\mathbf{q}) = (1-q_K)\Big[4\,C_{\mathrm{C}}\,q_C(1-q_C)^3(1-q_P^2) + 2\,C_{\mathrm{P}}\,q_P(1-q_P)\,R_{\mathrm{C}}(q_C)\Big], \tag{25.14}$$

and with equation (25.2) the risk of the hall is explicit. A configuration with one CRAC unit and one pump down receives both parts, R\$ 200 h^{-1}, as the schedule requires. Each configuration probability is multilinear in the seven units' unavailabilities, equation (13.22), and $C(\mathbf{s})$ does not depend on $\mathbf{q}$, so the risk is multilinear in the seven units and linear in the consequence rates. Because units share an entry, a product of a shared unavailability with itself appears: the risk of the hall is a polynomial of degree four in q_C, two in q_P and one in q_K, as equations (25.2) and (25.14) show.

The exact expectation, for the value beside the answer. By linearity of the expectation,

$$\mathbb{E}[\mathcal{R}(\mathbf{q}) \mid \mathbf{y}] = \sum_{\mathbf{s}} C(\mathbf{s})\, \mathbb{E}\big[\Pr(\mathbf{s} \mid \mathbf{q}) \mid \mathbf{y}\big], \tag{25.15}$$

and, the three group posteriors being independent, each expected configuration probability is a product of moments of the three posteriors. For the hall, equations (25.13) and (25.14) give

$$\begin{aligned}\mathbb{E}[\mathcal{R} \mid \mathbf{y}] = {} & C_{\mathrm{L}}\, \mathbb{E}[U_{\mathrm{sys}} \mid \mathbf{y}] \\ & + (1 - \mathbb{E}[q_K])\Big[4\,C_{\mathrm{C}}\,\mathbb{E}\big[q_C(1-q_C)^3\big](1-\mathbb{E}[q_P^2]) + 2\,C_{\mathrm{P}}\,\mathbb{E}\big[q_P(1-q_P)\big]\,\mathbb{E}[R_{\mathrm{C}}]\Big],\end{aligned} \tag{25.16}$$

every expectation over the posterior: the first is equation (14.28), and the others are combinations of the Beta moments of equation (14.20), such as $\mathbb{E}[q_C(1-q_C)^3] = \mathbb{E}[q_C] - 3\mathbb{E}[q_C^2] + 3\mathbb{E}[q_C^3] - \mathbb{E}[q_C^4]$. No draw is needed for the mean. By the law of total probability over $\mathbf{q}$, as in Chapter 13 for the unavailability, this expectation is also the predictive expected cost rate at a random instant,

with the aleatory status of the units and the epistemic uncertainty of their unavailabilities averaged together; times t_{yr} it is the expected loss of a year given the records.

The gap. The plug-in risk is $\mathcal{R}(\bar{\mathbf{q}})$, with $\bar{\mathbf{q}}$ the posterior means, which are the unavailabilities at the MAP state, $\mathbf{q}(\hat{\mathbf{z}})$ (Section 25.2.1). Expanding $\mathcal{R}$ to second order about $\bar{\mathbf{q}}$ and taking the expectation, the first-order term vanishes because $\mathbb{E}[\mathbf{q}-\bar{\mathbf{q}} \mid \mathbf{y}] = \mathbf{0}$, and what remains is the curvature weighted by the posterior covariance of the unavailabilities:

$$\mathbb{E}[\mathcal{R} \mid \mathbf{y}] - \mathcal{R}(\bar{\mathbf{q}}) \approx \tfrac{1}{2}\,\mathbb{E}\big[(\mathbf{q}-\bar{\mathbf{q}})^{\mathsf{T}}\mathbf{H}_{\mathbf{q}}(\mathbf{q}-\bar{\mathbf{q}}) \mid \mathbf{y}\big] = \tfrac{1}{2}\,\mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{q}}\mathbf{H}_{\mathbf{q}}), \tag{25.17}$$

with $\mathbf{H}_{\mathbf{q}} = \partial^2\mathcal{R}/\partial\mathbf{q}\,\partial\mathbf{q}^{\mathsf{T}}$ the 3×3 Hessian of the risk in the shared unavailabilities at $\bar{\mathbf{q}}$ and $\boldsymbol{\Sigma}_{\mathbf{q}} = \mathrm{Cov}(\mathbf{q} \mid \mathbf{y})$ their exact posterior covariance: the curvature term of equation (2.24), written in the unavailabilities about their posterior means, with their exact covariance in place of the Laplace covariance of the logits. The three posteriors are independent, so $\boldsymbol{\Sigma}_{\mathbf{q}} = \mathrm{diag}(\mathrm{Var}(q_j \mid \mathbf{y}))$ and the trace keeps only the diagonal of the Hessian, $\frac{1}{2}\sum_j (\partial^2\mathcal{R}/\partial q_j^2)\,\mathrm{Var}(q_j \mid \mathbf{y})$. The risk is multilinear in the units but not in a shared entry: q_C enters four units, and its second derivative has three parts, the loss-state part $C_{\mathrm{L}}(1-q_P^2)(1-q_K)\,(-R''_{\mathrm{C}})$ with $-R''_{\mathrm{C}} = 12-48q_C+36q_C^2$ the curvature of the CRAC factor, equation (13.31); the CRAC degraded part $4C_{\mathrm{C}}(1-q_P^2)(1-q_K)$ times $\mathrm{d}^2[q(1-q)^3]/\mathrm{d}q^2 = (1-q)(12q-6)$; and the pump degraded part through R''_{C}. The risk is linear in q_K, so the controller adds nothing. The loss-state term is convex in q_C, because the hall is lost through coincidences of CRAC units, and its expectation exceeds its plug-in value. The degraded term is concave: the probability of exactly one CRAC unit down, $4q_C(1-q_C)^3$, rises with q_C across the band, but less than in proportion, because at a higher q_C a second unit is more often down as well and the configuration becomes a loss; its expectation therefore falls short of its plug-in value. The sign of the gap is the sign of the curvature of the dominant term. The terms of third and fourth order, which the expansion leaves out, are those of the third and fourth central moments of the posterior of q_C.

The expectation from the state. The route reported is equation (2.24) itself, in the logits about $\hat{\mathbf{z}}$ with the Laplace covariance:

$$\mathcal{R}(a) \approx L(a,\hat{\mathbf{z}}) + \tfrac{1}{2}\,\mathrm{tr}(\mathbf{H}_a\boldsymbol{\Sigma}), \qquad \mathbf{H}_a = \left.\frac{\partial^2\mathcal{R}_{\mathrm{yr}}}{\partial\mathbf{z}\,\partial\mathbf{z}^{\mathsf{T}}}\right|_{\hat{\mathbf{z}}}, \tag{25.18}$$

with $L(a,\hat{\mathbf{z}}) = \mathcal{R}_{\mathrm{yr}}(\bar{\mathbf{q}})$ the plug-in risk and, $\boldsymbol{\Sigma}$ being diagonal, only the diagonal of $\mathbf{H}_a$ entering the trace. The risk is linear in q_K but not in its logit ξ_K, so in this route the controller's entry contributes a curvature term of its own, and the term of ξ_C carries the convexity of $q_C = 1/(1+\mathrm{e}^{-\xi_C})$ in the logit on top of that of the risk in q_C.

The band. Under the Laplace posterior the band is that of the delta method on the logarithm, the rule of Chapter 13 for a positive quantity uncertain by a factor, equation (14.29). The gradient of $\ln\mathcal{R}$ with respect to the state passes through the factors of equation (14.30), the derivative of each shared unavailability with respect to its logit and the sum over the units that share it, which $\mathcal{R}(\mathbf{q})$ already carries:

$$\begin{aligned} \mathbf{g} &= \frac{\partial \ln\mathcal{R}}{\partial\mathbf{z}} = \frac{1}{\mathcal{R}(\hat{\mathbf{z}})}\,\mathbf{D}\,\frac{\partial\mathcal{R}}{\partial\mathbf{q}}, \qquad \mathbf{D} = \mathrm{diag}(q_j(1-q_j)), \\ \mathrm{sd}(\ln\mathcal{R} \mid \mathbf{y}) &\approx \sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}, \qquad \mathcal{R}_{\mathrm{yr}} \in \mathcal{R}_{\mathrm{yr}}(\hat{\mathbf{z}})\,\exp\!\Big(\pm 1.96\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}\Big), \end{aligned} \tag{25.19}$$

with $\boldsymbol{\Sigma}$ the Laplace covariance of the three logits. The derivatives $\partial\mathcal{R}/\partial q_j$ are those of equations (25.2) and (25.14); their loss-state parts are C_{L} times the gradient of the top event, equation (14.44). With $\boldsymbol{\Sigma}$ diagonal the product $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g} = \sum_j g_j^2\Sigma_{jj}$ is a variance budget, one term per entry. The center of this band is the plug-in risk, which, by the gap (25.17) and the skew of the posterior of q_C, need not be the exact median. The exact band is read from draws of the three Beta posteriors, each draw passed through equations (25.2) and (25.14).

Building blocks. The state $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$ of equation (25.8); the closed forms (25.2) and (25.14) at $\hat{\mathbf{z}}$ for the plug-in risk; the diagonal of $\mathbf{H}_a$ in the logits and the gradient $\mathbf{g}$ (25.19). For the exact values beside them, the Beta posteriors of the three groups and their moments to order four, equation (14.20), the second derivatives for $\mathbf{H}_{\mathbf{q}}$, and draws of the three posteriors for the band.

The expected annual risk and its band. At $\hat{\mathbf{z}}$ the plug-in risk is R$ 1 252 736 per year (Section 25.3.1). Equation (25.18), with the diagonal (3 285 728, 75 967, 173 599) of $\mathbf{H}_a$ and the diagonal (0.34364, 0.25253, 1.00100) of $\boldsymbol{\Sigma}$, adds $\frac{1}{2}\operatorname{tr}(\mathbf{H}_a\boldsymbol{\Sigma}) =$ R$ 661 036:

$$\mathcal{R}(a) \approx 1\,252\,736 + 661\,036 = \text{R\$}\,1\,913\,773 \text{ per year.} \tag{25.20}$$

With the gradient $\mathbf{g} = (1.478,\ 0.034,\ 0.139)^{\mathsf{T}}$ of equation (25.19), $\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}} = 0.878$ and the central band is

$$1\,252\,736\,\exp(\pm 1.96 \times 0.878) = [\text{R\$}\,224\,298,\ \text{R\$}\,6\,996\,702], \tag{25.21}$$

a factor of 31, of median R$ 1 252 736. The value beside them comes from the exact posterior. The Beta moments of equation (14.20) give $\mathbb{E}[q_C^r \mid \mathbf{y}]$ equal to 0.03, 1.1881×10^{-3}, 5.8241×10^{-5} and 3.3927×10^{-6} for $r = 1$ to 4, and with them

$$\mathbb{E}[q_C(1-q_C)^3] = 0.03 - 3 \times 1.1881 \times 10^{-3} + 3 \times 5.8241 \times 10^{-5} - 3.3927 \times 10^{-6} = 0.026607,$$

$\mathbb{E}[q_P^2] = 1.247 \times 10^{-4}$, $\mathbb{E}[q_P(1-q_P)] = 0.01 - 1.247 \times 10^{-4} = 9.875 \times 10^{-3}$, $\mathbb{E}[R_{\mathrm{C}}] = 0.993327$ and $\mathbb{E}[U_{\mathrm{sys}} \mid \mathbf{y}] = 7.790 \times 10^{-3}$, equation (14.36). The three terms of equation (25.16) are $20\,000 \times 7.790 \times 10^{-3} =$ R$ 155.80 h^{-1} for the loss states, R$ 15.95 h^{-1} for one CRAC unit down and R$ 0.98 h^{-1} for one pump down, and

$$\mathbb{E}[\mathcal{R}_{\mathrm{yr}} \mid \mathbf{y}] = 8760 \times 172.7266 = \text{R\$}\,1\,513\,085 \text{ per year,} \tag{25.22}$$

R$ 400 688 below the reported value: the normal on the logit of q_C has a heavier upper tail than Beta$(3, 97)$.

Table 25.3 collects the posterior of the annual risk. Each row is a summary; the columns give its value from the state, its value over the draws of the exact posterior and, for the mean, the exact value of equation (25.22).

Table 25.3: The posterior of the annual risk of the hall's cooling. State: equations (25.20) and (25.21) and the delta spread $\mathcal{R}_{\mathrm{yr}}(\hat{\mathbf{z}})\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}$. Draws: 2×10^5, seed 24; exact: equation (25.16). Amounts in reais per year.

Summary	State	Draws	Exact
mean $\mathbb{E}[\mathcal{R}_{\mathrm{yr}} \mid \mathbf{y}]$ [R$/yr]	1 913 773	1 514 242	1 513 085
median [R$/yr]	1 252 736	1 085 231	–
standard deviation [R$/yr]	1 099 409	1 403 554	–
skewness [–]	–	2.6	–
central 95 % band [R$/yr]	[224 298, 6 996 702]	[192 552, 5 343 835]	–

Figure 25.2 draws the posterior. The horizontal axis is the annual risk in reais per year on a logarithmic scale; the vertical axis is the posterior density per decade, the density of $\log_{10}\mathcal{R}_{\mathrm{yr}}$. The dotted blue curve is the normal on $\ln\mathcal{R}_{\mathrm{yr}}$ of equation (25.19), the solid curve the density of the exact draws. The shaded area under the solid curve is the exact probability that the risk is below R$ 1.5 million, the acceptance probability of Section 25.3.3. The three gray dashed lines mark the limits of R$ 1.0, R$ 1.5 and R$ 2.0 million; the green dashed line the median, the black dotted line the plug-in risk of Section 25.3.1 and the red line the mean.

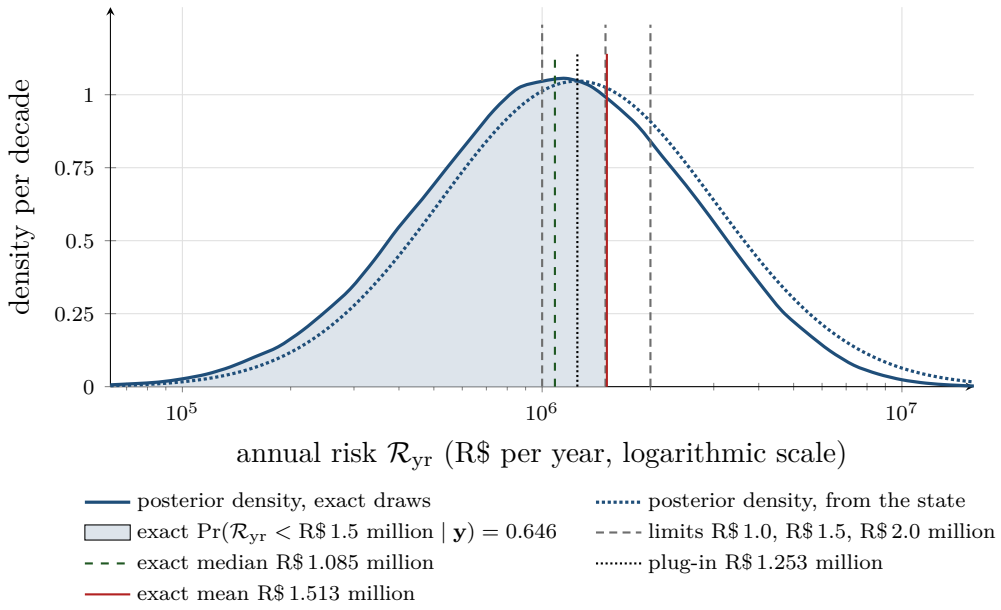


Figure 25.2: Posterior of the annual risk of the hall's cooling, $\mathcal{R}_{\text{yr}} = 8760\,\text{h} \times [C_{\text{L}}U_{\text{sys}} + \mathcal{R}_{\text{deg}}]$, equations (25.2) and (25.14), with $q_C \sim \text{Beta}(3, 97)$, $q_P \sim \text{Beta}(4, 396)$, $q_K \sim \text{Beta}(1, 999)$ and the schedule of Table 25.1: kernel density of $\log_{10} \mathcal{R}_{\text{yr}}$ over 10^5 of the 2×10^5 draws (seed 24), against the normal on $\ln \mathcal{R}_{\text{yr}}$ at $\hat{\mathbf{z}}$ with standard deviation 0.878 (dotted). The exact mean, R$ 1.51 million, lies just above the R$ 1.5 million limit, and the state's, R$ 1.91 million, further above it; the shaded area is the exact acceptance probability at that limit.

The state's density is symmetric on the logarithmic scale by construction, with its peak at R$ 1.25 million and a central 95 % band from R$ 0.22 to R$ 7.0 million, a factor of 31; the exact density is close to symmetric, with its peak near R$ 1.1 million and a band from R$ 0.19 to R$ 5.3 million, a factor of 28. On the linear scale the exact posterior is strongly skewed to the right, with a skewness of 2.6, and its median, R$ 1.09 million, lies well below its mean, R$ 1.51 million. The answer to the first question is the band as much as the mean: the expected loss is R$ 1.91 million per year (R$ 1.51 million exactly), and the evidence allows anything from a fifth of a million to seven million (five exactly). The width is that of the posterior of q_C, which spans a factor of eleven and enters the loss-state term squared. The draws' mean is set against the exact mean in Section 25.4.2.

The plug-in risk against the expected risk. At the plug-in values $q_C = 0.030$, $q_P = 0.010$, $q_K = 0.001$ the five classes of Table 25.1 have the probabilities 6.281×10^{-3} (loss of cooling), 0.1072 (one CRAC unit down), 0.0175 (one pump down), 0.0022 (one of each) and 0.8668 (nominal), and the three terms of equations (25.13) and (25.14) are

$$\begin{aligned} C_{\text{L}}U_{\text{sys}} &= 20\,000 \times 6.2806 \times 10^{-3} = \text{R\$ } 125.61\,\text{h}^{-1}, \\ \text{CRAC part} &= 0.999 \times 4 \times 150 \times 0.03 \times 0.97^3 \times 0.9999 = \text{R\$ } 16.41\,\text{h}^{-1}, \\ \text{pump part} &= 0.999 \times 2 \times 50 \times 0.01 \times 0.99 \times 0.994814 = \text{R\$ } 0.98\,\text{h}^{-1}, \end{aligned}$$

a total of $\mathcal{R}(\bar{\mathbf{q}}) = \text{R\$ } 143.01\,\text{h}^{-1}$ and

$$\mathcal{R}_{\text{yr}}(\bar{\mathbf{q}}) = 8760 \times 143.0064 = \text{R\$ } 1\,252\,736 \text{ per year}, \tag{25.23}$$

the plug-in risk, the figure of the first reading. It is the 50th percentile of the state's posterior of Figure 25.2, whose second-order mean is its 69th; of the exact posterior it is the 57th, and

the exact mean the 65th. Table 25.4 sets it against the expected risk term by term. Each row is a part of the risk; the columns give its plug-in value, its expectation from the state by equation (25.18) with the difference and ratio, then its exact expectation with the same two.

Table 25.4: Plug-in against expected annual risk of the hall's cooling, by part. Plug-in at $\bar{q}_C = 0.030$, $\bar{q}_P = 0.010$, $\bar{q}_K = 0.001$; state, equation (25.18); exact, from the Beta moments, equation (25.16).

Part	Plug-in [R$/yr]	State [R$/yr]	Difference [R$/yr]	Ratio [–]	Exact [R$/yr]	Difference [R$/yr]	Ratio [–]
loss states	1 100 366	1 744 236	+643 870	1.585	1 364 812	+264 446	1.240
degraded, one CRAC unit down	143 752	159 925	+16 173	1.113	139 689	−4 063	0.972
degraded, one pump down	8 619	9 611	+992	1.115	8 584	−35	0.996
annual risk	1 252 736	1 913 773	+661 037	1.528	1 513 085	+260 349	1.208

The plug-in risk is 35 % below the expected risk from the state and 17 % below the exact one. In the logits every term has positive curvature, the degraded ones included, because q_j is convex in its logit at these values; in the unavailabilities the degraded terms are concave. In the exact expectation the whole difference comes from the loss-state term, whose ratio, 1.240, is the ratio of the expected to the plug-in top event of Chapter 14, 7.790/6.281, since the factor $C_{\mathrm{L}} t_{\mathrm{yr}}$ is common. The degraded terms have ratios just below one, the concave case of Section 25.3.1.

The gap (25.17) explains the difference. The curvature of the CRAC factor at $q_C = 0.03$ is $-R''_{\mathrm{C}} = 12 - 48 \times 0.03 + 36 \times 0.03^2 = 10.592$, and the three parts of $\partial^2 \mathcal{R} / \partial q_C^2$ are $20\,000 \times 0.9999 \times 0.999 \times 10.592 = \mathrm{R\$}\,211\,615\,\mathrm{h}^{-1}$ for the loss states, $-\mathrm{R\$}\,3\,279\,\mathrm{h}^{-1}$ for the CRAC degraded part, through $(1-q)(12q-6) = -5.471$, and $-\mathrm{R\$}\,10\,\mathrm{h}^{-1}$ for the pump degraded part; the convex part is more than sixty times the concave part. The diagonal of the Hessian and of the posterior covariance of the unavailabilities are

$$\operatorname{diag} \mathbf{H_q} = (208\,326,\ 39\,521,\ 0)\ \mathrm{R\$\,h^{-1}}, \qquad \operatorname{diag} \mathbf{\Sigma_q} = (2.881 \times 10^{-4},\ 2.469 \times 10^{-5},\ 9.98 \times 10^{-7}),$$

so that $\frac{1}{2} \operatorname{tr}(\mathbf{\Sigma_q H_q}) = \frac{1}{2}(208\,326 \times 2.881 \times 10^{-4} + 39\,521 \times 2.469 \times 10^{-5}) = 30.01 + 0.49 = \mathrm{R\$}\,30.50\,\mathrm{h}^{-1}$: R$ 262 899 per year from the CRAC units, R$ 4 274 from the pumps and nothing from the controller, R$ 267 172 in all against the exact R$ 260 349. The remaining −R$ 6 823, 2.6 % of the gap, comes almost entirely from the cubic term of the CRAC unavailability: the third derivative of the risk in q_C, dominated by its loss-state part through $(-R_{\mathrm{C}})''' = -48 + 72 q_C$, is negative, and the third central moment of the right-skewed posterior is positive, so the term contributes −R$ 7 020 per year; the quartic term adds +R$ 202. A plant whose risk were dominated by degraded states, or by a single unit in series, would have a plug-in risk at or above its expectation; a redundant plant, whose loss needs a coincidence among units of one shared posterior, has it below.

25.3.2 The parts of the risk

Engineering question. Where does the risk come from: the loss states or the degraded states, and which cut sets? The owner wants to know which kind of hour carries the cost, and which combinations of failed units carry the hours without cooling.

General form. The estimation form again, equation (2.10), applied to each part f_k of a sum $f = \sum_k f_k$ of derived quantities. The expectation is linear, so the expected parts add to the expected whole,

$$\mathbb{E}[f \mid \mathbf{y}] = \sum_k \mathbb{E}[f_k \mid \mathbf{y}]; \tag{25.24}$$

their bands do not add, and a share f_k/f is a derived quantity with a posterior of its own.

Specialized form. The question is equation (25.24) with f the annual risk, each f_k a class of configuration or a group of cut sets, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \mathbf{\Sigma})$ of equation (25.8): each part has its own second-order expectation (25.18) and its own band. One choice enters here: the loss-state term, a single probability times a single cost, is attributed to the minimal cut sets by a rule, in proportion to their probabilities or by a fixed order of the cut sets. Either rule divides the term into parts that sum to it exactly, and neither is unique where cut sets occur together; Section 25.3.2 sets the two side by side. The exact values beside them are read from the exact posterior, the expectations from the moments and the bands and shares from the draws. The gradient $\mathbf{g}$ of equation (25.19) with $\mathbf{\Sigma}$ splits the band rather than the level: with $\mathbf{\Sigma}$ diagonal its terms $g_j^2 \Sigma_{jj}$ attribute the variance of $\ln \mathcal{R}$ to the three entries, the counterpart for the spread of what the split by class and by cut set does for the expectation, and a part f_k has its own gradient $\mathbf{D}\,(\partial f_k/\partial \mathbf{q})/f_k$ and Laplace band in the same way.

By state class. Equation (25.13) is already the decomposition by class: the loss-state term $C_{\mathrm{L}} U_{\mathrm{sys}}$ and the degraded term, which equation (25.14) splits further into its CRAC and pump parts. The decomposition is exact, because the classes partition the configurations, and it holds for every $\mathbf{q}$, so that each class has its own posterior and its share of the risk is a derived quantity.

By cut set, from above. The loss-state term is the cost rate times the top event, and Chapter 14 bounded the top event by the rare-event sum of the minimal cut sets, equation (14.24). Multiplying by $C_{\mathrm{L}} > 0$,

$$C_{\mathrm{L}}\, U_{\mathrm{sys}}(\mathbf{q}) \;\le\; \sum_{c \in \mathcal{M}} C_{\mathrm{L}}\, P_c(\mathbf{q}) = C_{\mathrm{L}}\, S_1(\mathbf{q}) = C_{\mathrm{L}}(6q_C^2 + q_P^2 + q_K), \tag{25.25}$$

with $\mathcal{M}$ the minimal cut sets, P_c the probability that all units of cut set c are down and S_1 their rare-event sum, as in Chapter 14: one term per minimal cut set, each the cost rate of a lost hall times the probability that all units of the cut set are down. The right side overcounts the loss-state term by the same fraction as the rare-event sum overcounts the top event, since the factor C_{L} is common. Attributing the exact loss-state term in proportion to the cut-set terms uses the share $\pi_{\mathcal{B}}$ of a group of cut sets, equation (14.37): the part of the loss-state term attributed to group $\mathcal{B}$ is $\pi_{\mathcal{B}}\, C_{\mathrm{L}} U_{\mathrm{sys}}$, and the parts sum to the term.

By cut set, exactly. An exact attribution without overcount partitions the loss configurations by a fixed order of the cut sets. Rewriting equation (25.2),

$$U_{\mathrm{sys}} = q_K + (1 - q_K)(1 - R_{\mathrm{C}}(q_C)) + (1 - q_K)\, R_{\mathrm{C}}(q_C)\, q_P^2, \tag{25.26}$$

which follows from $1 - R_{\mathrm{C}}(1 - q_P^2)(1 - q_K) = q_K + (1 - q_K)\big[(1 - R_{\mathrm{C}}) + R_{\mathrm{C}} q_P^2\big]$. The three terms are disjoint events: the controller down; the controller up and two or more CRAC units down; the controller up, at most one CRAC unit down and both pumps down. The partition is exact, and it depends on the order: a configuration with the controller and two CRAC units down is assigned to the controller. The two attributions, proportional and ordered, differ by the probability of the cut sets occurring together.

Linearity in the consequence. Every term of the decomposition is a probability times a consequence rate. Scaling every rate by a factor scales the risk and every part of it by that factor and leaves every share unchanged; scaling only C_{L} scales the loss-state term and leaves the shares of the cut sets within it unchanged, because they all carry the same factor. The attribution of the loss-state term to cut sets therefore does not depend on the consequence schedule at all. A limit on the risk in reais does.

Building blocks. The state $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$, the diagonal of each part's Hessian in the logits and each part's gradient. For the exact values beside them, the moments of equation (14.20) and the shares $\pi_{\mathcal{B}}$ of Chapter 14, and draws of the three posteriors for the bands of the parts and for the posterior of a share.

Where the risk comes from. Table 25.5 decomposes the expected annual risk. Each row is a class of configuration or a group of cut sets; the columns give its expected annual amount from the state and exactly, its share of the expected risk from the state and exactly, and the central 95 % band of the amount, by the delta method on its logarithm and over the draws. The upper block is the decomposition by class, equations (25.13) and (25.14), exact. The middle block attributes the loss-state term to the minimal cut sets in proportion to their expected probabilities, 84.3, 14.3 and 1.4 % from the state (the exact shares 86.4, 12.1 and 1.5 % of Table 14.6). The third block gives the rare-event terms of equation (25.25) before that scaling, and the last the ordered partition of equation (25.26). One row of the arithmetic, a CRAC pair in the rare-event sum with the exact moment: $20\,000 \times 8760 \times 1.1881 \times 10^{-3} = \text{R\$}\,208\,158$ per year, and six of them R$ 1 248 950.

Table 25.5: The expected annual risk of the hall's cooling by class of configuration and by minimal cut set. State: second-order expectations and delta bands on the logarithm at $\hat{\mathbf{z}}$. Exact: Beta moments, and bands from 2×10^5 draws, seed 24. Shares of the expected risk, R$ 1 913 773 from the state and R$ 1 513 085 exactly.

Part	State [R$/yr]	Exact [R$/yr]	Share, state (exact)	95 % band, state [R$/yr]	95 % band, draws [R$/yr]
loss states, $C_{\mathrm{L}} t_{\mathrm{yr}} U_{\mathrm{sys}}$	1 744 236	1 364 812	91.1 % (90.2 %)	[176 865, 6 845 925]	[143 385, 5 036 661]
degraded, one CRAC unit down	159 925	139 689	8.4 % (9.2 %)	[52 300, 395 119]	[32 631, 299 365]
degraded, one pump down	9 611	8 584	0.5 % (0.6 %)	[3 283, 22 629]	[2 380, 18 587]
loss states attributed to the six CRAC pairs	1 470 073	1 178 829	76.8 % (77.9 %)	–	–
loss states attributed to the controller	249 337	165 364	13.0 % (10.9 %)	–	–
loss states attributed to the pumps	24 826	20 619	1.3 % (1.4 %)	–	–
rare-event term, six CRAC pairs	1 548 418	1 248 950	–	[101 832, 8 789 679]	[42 192, 5 322 200]
rare-event term, controller	262 625	175 200	–	[24 703, 1 242 583]	[4 389, 643 621]
rare-event term, pumps	26 149	21 845	–	[2 492, 123 164]	[1 323, 83 685]
rare-event sum $C_{\mathrm{L}} t_{\mathrm{yr}} S_1$	1 837 191	1 445 996	–	[174 397, 7 436 280]	[144 280, 5 530 536]
ordered partition, controller down	262 625	175 200	13.7 % (11.6 %)	[24 703, 1 242 583]	[4 389, 643 621]
ordered partition, two CRAC units down, controller up	1 455 688	1 167 934	76.1 % (77.2 %)	[102 241, 8 059 518]	[41 779, 4 826 266]
ordered partition, both pumps down, the rest cooled	25 923	21 678	1.4 % (1.4 %)	[2 477, 122 407]	[1 311, 83 053]

Figure 25.3 draws the decompositions as stacked bars, the state classes from the state above the three exact ones. The horizontal axis is the expected annual amount in millions of reais. The top bar is the decomposition by class: the loss states in red, the degraded configurations with a CRAC unit down in light gray and with a pump down in dark gray. The middle bar replaces the loss-state term by the rare-event terms of the cut sets, the CRAC pairs in green, the controller in blue and the pumps in gold, with the same degraded parts; the bottom bar

uses the same terms scaled to the exact loss-state term. The vertical dashed line marks the expected risk.

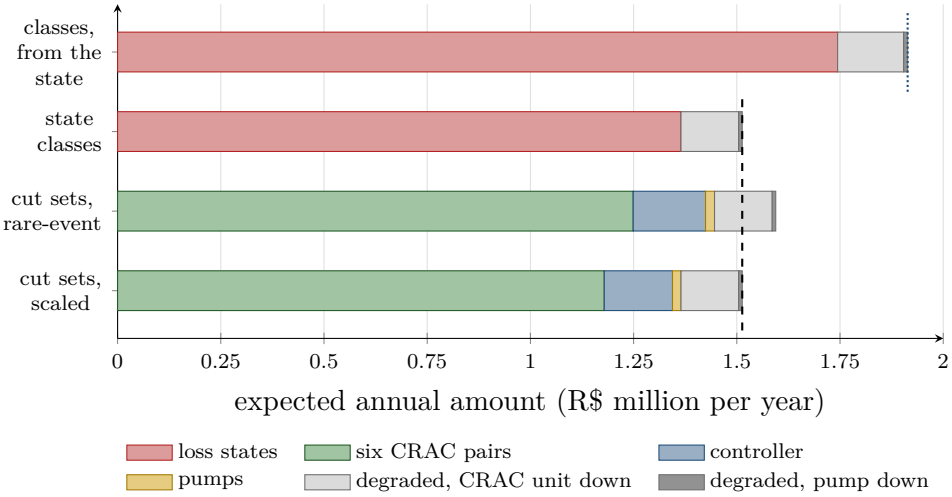


Figure 25.3: The expected annual risk of the hall's cooling decomposed by class of configuration (top), by minimal cut set with the rare-event terms $C_{\mathrm{L}} t_{\mathrm{yr}} P_c$ (middle) and by minimal cut set scaled to the exact loss-state term (bottom), each with the two degraded parts of equation (25.14). Exact expectations from the Beta moments; dashed line, the exact expected risk R$ 1.513 million per year. Top bar: the state classes from the state, second-order expectations at $\hat{\mathbf{z}}$; dotted line, R$ 1.914 million.

Loss states against degraded states. The loss states carry 91 % of the expected risk, R$ 1.74 million, and the degraded states 9 %, R$ 0.17 million (90 %, R$ 1.36 million and 10 %, R$ 0.15 million exactly). The two classes reach that result from opposite ends: at the plug-in values the hall spends 0.63 % of its time without cooling at R$ 20 000 h^{-1}, and 12.7 % of its time with a unit down at R$ 137 h^{-1} on average, twenty times the time at a hundred and fifty times less per hour. The ratio of expectations is itself a point value of an uncertain share: at $\hat{\mathbf{z}}$ the loss states' share is 0.878, and the delta method on its logit gives the band [0.75, 0.95] and a probability 0.32 of exceeding 0.9; over the draws it has median 0.88 and band [0.73, 0.94], and it exceeds 0.9 with probability 0.30. That the loss states dominate is certain; by how much is not. The degraded term alone is better known than the loss-state term, a band of a factor of seven and a half against thirty-nine (nine against thirty-five exactly), because it is nearly linear in q_C where the loss-state term is nearly quadratic.

The cut sets. Scaled to the loss-state term, the six CRAC pairs carry R$ 1.47 million per year, the controller R$ 0.25 million and the pumps R$ 0.02 million from the state, 77, 13 and 1.3 % of the whole risk; exactly, R$ 1.18, R$ 0.17 and R$ 0.02 million, 78, 11 and 1.4 %. The ordered partition of equation (25.26) gives R$ 1.17, R$ 0.18 and R$ 0.02 million. The CRAC pairs receive within one percent of their proportional amount; the controller and the pumps receive nearly their whole rare-event terms, 6 and 5 % more than their proportional amounts. The overcount of the rare-event sum lies almost entirely among the CRAC pairs themselves, in the twelve pairs of pairs that share a unit, Section 14.3.2, and the proportional rule spreads it over all three groups. The per-draw bands of the rare-event terms span two orders of magnitude each, the attribution question of Chapter 14 in reais: by the delta method on the logit of its share, the CRAC part of the ordered partition carries more than half of the loss-state term with probability 0.86, and over the draws with probability 0.85, the probability Chapter 14 found for the share of the CRAC pairs in the top event.

25.3.3 Risk-informed acceptance

Engineering question. Is the annual loss below R$ 1.5 million with an assurance of 0.9? The owner must state that the hall's cooling costs less than a limit, and the operator of the load asks how sure that statement is.

General form. The threshold form of Chapter 2, equation (2.12): the posterior probability of the event that the requirement holds, compared with an assurance level κ. A requirement that a quantity stay below an upper limit is the same form with $f = -\mathcal{R}_{\mathrm{yr}}$ and $f^* = -\mathcal{R}^*$, and the rule is applied to the requirement, never to the exceedance.

Specialized form. The question is equation (2.12) with $f = -\mathcal{R}_{\mathrm{yr}}$, $f^* = -\mathrm{R\$}\,1.5$ million and $\kappa = 0.9$, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of equation (25.8). One choice enters here: the requirement is read on the annual risk, the expected loss of a year given the state, and not on the loss of one particular year. This makes acceptance a statement about the plant's loss rate, a tail of one derived quantity, and it is the wrong reading when the limit is a budget that a single year must not overrun; that reading is the predictive of Section 25.3.4. The event and its probability are

$$\Pr\bigl(\mathcal{R}_{\mathrm{yr}}(\mathbf{q}) < \mathcal{R}^* \mid \mathbf{y}\bigr) = \int \mathbf{1}\bigl[\mathcal{R}_{\mathrm{yr}}(\mathbf{q}) < \mathcal{R}^*\bigr]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \tag{25.27}$$

the posterior probability of the event. Risk-informed acceptance with the assurance κ accepts the plant when this probability is at least κ. Since the probability is the posterior distribution function of $\mathcal{R}_{\mathrm{yr}}$ at $\mathcal{R}^*$, the rule is equivalent to $\mathcal{R}^* \geq \mathcal{R}_{\mathrm{yr},\kappa}$, where $\mathcal{R}_{\mathrm{yr},\kappa}$ is the κ-quantile of the risk's posterior: a limit is met with assurance κ exactly when it lies above the κ-quantile. The rule is not the comparison of the expected risk with the limit. The mean of a right-skewed posterior lies above its median, so a plant whose expected risk lies just below the limit has an acceptance probability above one half, and it can still fall well short of the required assurance.

Under the Laplace posterior the tail is the normal tail of $\ln \mathcal{R}$ with the gradient and the covariance of equation (25.19),

$$\Pr(\mathcal{R}_{\mathrm{yr}} < \mathcal{R}^* \mid \mathbf{y}) \approx \Phi\Bigl(\frac{\ln \mathcal{R}^* - \ln \mathcal{R}_{\mathrm{yr}}(\hat{\mathbf{z}})}{\sqrt{\mathbf{g}^{\mathsf{T}} \boldsymbol{\Sigma} \mathbf{g}}}\Bigr), \tag{25.28}$$

with Φ the standard normal distribution function; it inherits the center of the band, the plug-in risk. The exact value beside it is the fraction of the draws of the Beta posteriors in which the event holds.

Building blocks. $\mathcal{R}_{\mathrm{yr}}(\hat{\mathbf{z}})$ and $\sqrt{\mathbf{g}^{\mathsf{T}} \boldsymbol{\Sigma} \mathbf{g}}$ for the normal tail and its quantiles; and, for the exact values beside them, draws of the three posteriors passed through the closed forms.

Risk-informed acceptance. By equation (25.28), with $\mathcal{R}_{\mathrm{yr}}(\hat{\mathbf{z}}) = \mathrm{R\$}\,1\,252\,736$ and $\sqrt{\mathbf{g}^{\mathsf{T}} \boldsymbol{\Sigma} \mathbf{g}} = 0.878$, the acceptance probability at the owner's limit is

$$\Phi\Bigl(\frac{\ln(1\,500\,000/1\,252\,736)}{0.878}\Bigr) = \Phi(0.205) = 0.581,$$

with 0.399 at R$ 1.0 million and 0.703 at R$ 2.0 million. The values beside them are the fractions of the exact draws: 0.646 at R$ 1.5 million, the shaded area of Figure 25.2, 0.463 at R$ 1.0 million and 0.762 at R$ 2.0 million, 0.06 higher each, because the exact posterior's center lies below the plug-in risk. None reaches the assurance of 0.9. The limit that would be met with that assurance is the 90th percentile, $1\,252\,736 \exp(1.2816 \times 0.878) = \mathrm{R\$}\,3.86$ million (R$ 3.18 million exactly), with R$ 2.62 million at an assurance of 0.8 (R$ 2.23) and R$ 5.31 million at

0.95 (R$ 4.22). On the present evidence the plant is not accepted at any of the three limits with any assurance above 0.70 (0.76 exactly). The gap does not change the verdict.

Two shortcuts would decide otherwise, in opposite directions. The exact expected risk, R$ 1.513 million, is just above the R$ 1.5 million limit and would reject the plant at that limit while accepting it at a limit of R$ 1.52 million, where the acceptance probability is still only 0.65; the state's, R$ 1.91 million, would reject it by a wider margin. The plug-in risk, R$ 1.25 million, would accept it at R$ 1.5 million outright. Neither carries the information that the band spans a factor of 31 (28 exactly): the risk is below the limit with probability 0.58, and above three million with probability 0.16 (in about one draw in nine exactly). The band is wide because q_C is wide, and the operator can narrow the acceptance question in two ways: by buying evidence, more checks of the CRAC units, which narrows the posterior of q_C and with it the band of the risk; or by buying a fifth CRAC unit, which moves the whole posterior of the risk down. Section 25.3.5 values the second and Section 25.6 sets the first against it.

25.3.4 The loss of one year and the bad year

Engineering question. What is the loss in a bad year? The owner asks for the cost that a year exceeds with a small probability, and the answer depends on whether the year's cost or its rate is meant.

General form. The predictive form of Chapter 2, equation (2.13), for a future observation that stays random when the state is known, $p(y_{\text{new}} \mid \mathbf{y}) = \int p(y_{\text{new}} \mid \mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}$, read at a quantile, a tail of the threshold form (2.12).

Specialized form. The question is equation (2.13) with $y_{\text{new}} = L_{\text{yr}}$, read at a quantile by equation (2.12), and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of equation (25.8). Three choices enter here. Given $\mathbf{q}$, the losses of cooling are taken to begin as a Poisson process: this makes the year's count of losses Poisson and the year's loss a compound Poisson sum, and it fails if losses cluster, one restart provoking the next. Their mean duration is pinned at MDT = 12 h, an illustrative value: this fixes the cost of one loss, and it changes the spread of the year's loss but not the risk, which Section 25.4.4 confirms by varying it. The degraded term of a year is taken at its mean: the hall enters its degraded configurations many times a year for short stays at a small cost, so that their year's cost varies far less than the loss term, whose every stay costs C_{L} MDT. The fit enters through the unavailabilities: at the MAP state, $\mathbf{q}(\hat{\mathbf{z}}) = \bar{\mathbf{q}}$, they give the year's loss at a known rate; the posterior of the annual risk is the normal on its logarithm of equation (25.19), with the delta variance $\mathcal{R}_{\text{yr}}(\hat{\mathbf{z}})^2\, \mathbf{g}^{\mathsf{T}} \boldsymbol{\Sigma} \mathbf{g}$; and the predictive itself is a sum over the state, a Gauss–Hermite rule of forty nodes on each of the three entries, its tail mean from 2×10^6 draws of the Laplace normal, seed 24. The exact value beside each comes from the draws of the exact posterior, one count per draw.

The annual risk $\mathcal{R}_{\text{yr}}(\mathbf{q})$ is an expectation given the unavailabilities, a rate; the loss L_{yr} of one year is the random integral it is the expectation of. Its distribution needs a model of the losses of cooling in time, which the unavailability alone does not give.

Losses of cooling in time. Given $\mathbf{q}$, let the losses of cooling begin with frequency ν_{loss} per year and last MDT each on average. The long-run unavailability is the outage frequency times the mean down time, equation (16.23) of Chapter 16, so

$$\nu_{\text{loss}}(\mathbf{q}) = \frac{t_{\text{yr}}\, U_{\text{sys}}(\mathbf{q})}{\text{MDT}} \quad \text{per year.} \tag{25.29}$$

The mean down time is not determined by the checks of Chapter 14, which see only whether a unit is down. It is the mean duration of a loss of cooling itself, the outage of the system; it is not a unit's repair time: a loss through a CRAC pair ends at the first of the two repairs, so

that with unit repairs of twelve hours it would last about six. A model with repair rates would derive the system's mean down time, as Chapter 16 did for the pump pair. With the durations set at their mean, each loss costs C_{L} MDT, and the loss of a year is

$$L_{\mathrm{yr}} = C_{\mathrm{L}}\,\mathrm{MDT}\,K_f + \mathcal{R}_{\mathrm{deg,yr}}(\mathbf{q}), \qquad K_f \mid \mathbf{q} \sim \mathrm{Poisson}\big(\nu_{\mathrm{loss}}(\mathbf{q})\,\tau\big), \quad \tau = 1\,\mathrm{yr}, \tag{25.30}$$

with K_f the number of losses of cooling in the year and $\mathcal{R}_{\mathrm{deg,yr}} = t_{\mathrm{yr}}\mathcal{R}_{\mathrm{deg}}$ the degraded term.

The mean and the variance given the unavailabilities. By equation (25.29),

$$\mathbb{E}[L_{\mathrm{yr}} \mid \mathbf{q}] = C_{\mathrm{L}}\,\mathrm{MDT}\,\nu_{\mathrm{loss}}\tau + \mathcal{R}_{\mathrm{deg,yr}} = C_{\mathrm{L}}\,t_{\mathrm{yr}}\,U_{\mathrm{sys}} + \mathcal{R}_{\mathrm{deg,yr}} = \mathcal{R}_{\mathrm{yr}}(\mathbf{q}), \tag{25.31}$$

so the mean of a year's loss is the annual risk whatever the mean down time; MDT cancels. The variance does not cancel:

$$\mathrm{Var}(L_{\mathrm{yr}} \mid \mathbf{q}) = (C_{\mathrm{L}}\,\mathrm{MDT})^2\,\nu_{\mathrm{loss}}\tau = C_{\mathrm{L}}\,\mathrm{MDT} \times C_{\mathrm{L}}\,t_{\mathrm{yr}}\,U_{\mathrm{sys}}, \tag{25.32}$$

proportional to the cost of one loss. Fewer, longer losses give the same mean and a larger variance. If the durations d_j of the losses are themselves random and independent, exponential with mean MDT, the loss term is a compound Poisson sum $C_{\mathrm{L}}\sum_{j=1}^{K_f} d_j$ with variance $\nu_{\mathrm{loss}}\tau\,C_{\mathrm{L}}^2\mathbb{E}[d_j^2] = 2(C_{\mathrm{L}}\,\mathrm{MDT})^2\nu_{\mathrm{loss}}\tau$, twice the fixed-duration value.

The predictive and the split of its variance. Averaging over the posterior gives the predictive distribution of the count,

$$\Pr(K_f = j \mid \mathbf{y}) = \int \frac{(\nu_{\mathrm{loss}}(\mathbf{q})\tau)^j\,\mathrm{e}^{-\nu_{\mathrm{loss}}(\mathbf{q})\tau}}{j!}\,p(\mathbf{z} \mid \mathbf{y})\,\mathrm{d}\mathbf{z},$$

the form of Chapter 11's predictive count, equation (11.16), with the frequency a derived quantity of the state rather than a rate with a gamma posterior; the integral has no closed form here and is evaluated by the sum over the state. The law of total variance splits the predictive variance of the year's loss into two parts,

$$\mathrm{Var}(L_{\mathrm{yr}} \mid \mathbf{y}) = \underbrace{\mathbb{E}[\mathrm{Var}(L_{\mathrm{yr}} \mid \mathbf{q}) \mid \mathbf{y}]}_{\text{aleatory: the count of losses}} + \underbrace{\mathrm{Var}(\mathcal{R}_{\mathrm{yr}}(\mathbf{q}) \mid \mathbf{y})}_{\text{epistemic: the loss rate}}, \tag{25.33}$$

by equation (25.31); the first is $(C_{\mathrm{L}}\mathrm{MDT})^2\mathbb{E}[\nu_{\mathrm{loss}}]\tau$ under fixed durations, and the second is the posterior variance of the risk. It is the split of equation (11.21), the Poisson part and the part added because the rate is not known.

Three distributions and their tails. Three distributions of a loss in reais answer the question: the posterior of $\mathcal{R}_{\mathrm{yr}}(\mathbf{q})$, which says what the loss rate is and how well it is known; the distribution of L_{yr} given $\bar{\mathbf{q}}$, the year's loss at the plug-in rate, which says how much a year varies when the rate is taken as known; and the predictive distribution of L_{yr} given $\mathbf{y}$, which carries both. For any of them, written X, the risk curve, the value at risk and the tail mean are

$$\begin{aligned} \bar{F}_X(x) = \Pr(X > x \mid \mathbf{y}), \qquad \mathrm{VaR}_p = \min\{x : \Pr(X \le x \mid \mathbf{y}) \ge p\}, \\ \mathrm{CVaR}_p = \frac{1}{1-p}\int_p^1 \mathrm{VaR}_u\,\mathrm{d}u, \end{aligned} \tag{25.34}$$

the tail mean being the average of the worst fraction $1-p$ of the distribution. For a continuous X it equals $\mathbb{E}[X \mid X \ge \mathrm{VaR}_p,\ \mathbf{y}]$; for a count with steps, as the year's loss at a known rate, the conditional expectation would average a tail heavier than $1-p$, and the integral form takes only the share of the step at VaR_p that the fraction $1-p$ needs. The value at risk of the posterior of the risk at $p = 0.95$ is a loss rate that the evidence leaves a five percent chance of exceeding: a bad state of knowledge, not a bad year. The value at risk of the predictive of L_{yr} is a year's loss with a five percent chance of being exceeded, which is what the phrase "a bad year" means.

Principle 25.1. *The risk is the expected loss given the state, a derived quantity whose posterior describes the epistemic uncertainty of the loss rate. The loss of one year scatters around it by an aleatory count of losses whose variance grows with the cost of a single loss. The predictive of a year's loss carries both, its variance is the sum of the two parts, and a tail of either distribution must be reported with the distribution it is a tail of.*

Building blocks. The state $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$; the frequency (25.29) at each node of the sum over the state; Poisson probabilities at the plug-in frequency for the year at a known rate; the quantiles and tail means (25.34) of the three distributions; and, for the exact values beside them, the frequency on each posterior draw and one Poisson count per draw.

The bad year and the risk curve. At the plug-in values the hall is without cooling 55 h a year, so equation (25.29) gives $\nu_{\text{loss}} = 55.0/12 = 4.585$ losses a year of $C_{\text{L}}\,\text{MDT} = \text{R\$}\,240\,000$ each, with the degraded term R$ 152 371 a year; the predictive mean frequency over the state is 7.983 a year ($68.2/12 = 5.687$ exactly). Table 25.6 gives the three distributions of Section 25.3.4. Each row is a summary; the columns are the posterior of the annual risk, the loss of one year at the plug-in unavailabilities, and the predictive loss of one year, each from the state with the exact value in parentheses.

Table 25.6: The bad year: the posterior of the annual risk (epistemic), the loss of one year at the plug-in unavailabilities (aleatory) and the predictive loss of one year (both), with losses of cooling as a Poisson count of losses of MDT = 12 h, equation (25.30). Risk: normal on the logarithm at $\hat{\mathbf{z}}$, its mean the second-order value of equation (25.18). Predictive: Gauss–Hermite over the state, so its mean is the full average over the same normal, a different route to the same expectation. In parentheses: draws of the exact posterior, 2×10^5, seed 24.

Summary	Posterior of $\mathcal{R}_{\text{yr}}$	L_{yr} given $\bar{\mathbf{q}}$	L_{yr} given $\mathbf{y}$
mean [R$/yr]	1 913 773 (1 513 085)	1 252 736	2 084 301 (1 513 085)
median [R$/yr]	1 252 736 (1 085 231)	1 112 371	1 370 724 (1 089 077)
standard deviation [R$/yr]	1 099 409 (1 403 554)	513 895	2 368 336 (1 513 872)
$\Pr(X <$ R$ 1.0 million) [–]	0.399 (0.463)	0.328	0.351 (0.461)
$\Pr(X <$ R$ 1.5 million) [–]	0.581 (0.646)	0.688	0.524 (0.633)
$\Pr(X <$ R$ 2.0 million) [–]	0.703 (0.762)	0.906	0.649 (0.747)
value at risk $\text{VaR}_{0.95}$ [R$/yr]	5 306 139 (4 222 725)	2 072 371	6 286 156 (4 377 500)
tail mean $\text{CVaR}_{0.95}$ [R$/yr]	8 155 387 (5 916 448)	2 433 471	9 683 118 (6 202 920)

The loss of a year at the plug-in rate. With the rate known, the year's loss is R$ 152 371 + R$ 240 000 K_f with $K_f \sim \text{Poisson}(4.585)$. One row of the arithmetic: the year stays below R$ 1.5 million when $K_f \leq 5$, since five losses give R$ 1 352 371 and six R$ 1 592 371, and $\Pr(K_f \leq 5) = 0.688$. The standard deviation is $240\,000\sqrt{4.585} =$ R$ 513 895, equation (25.32); the 95th percentile is eight losses, R$ 2.07 million; a year without a single loss of cooling has probability $\text{e}^{-4.585} = 0.010$. With exponential durations instead of fixed ones the standard deviation rises to R$ 727 000 and the 95th percentile to R$ 2.62 million, the doubling of the variance derived below equation (25.32).

The predictive loss. With the rate uncertain, the count of losses is overdispersed: over the state its predictive mean is 7.98 and its variance 92.0, and a year without a loss of cooling has probability 0.050. By equation (25.33) the variance of the year's loss is $(240\,000)^2 \times 7.983 = 4.60 \times 10^{11}$ from the count plus 5.15×10^{12} from the posterior of the risk over the same sum:

8 % aleatory and 92 % epistemic, a predictive standard deviation of R$ 2.37 million. The delta variance alone, $1\,099\,409^2 = 1.21 \times 10^{12}$, misses the long logit tails the sum carries. The values beside them come from the exact posterior: a mean of 5.69 and a variance of 36.7, six and a half times the mean, a probability of 0.093 for a year without a loss, and $(240\,000)^2 \times 5.687 = 3.28 \times 10^{11}$ from the count plus $1\,403\,554^2 = 1.97 \times 10^{12}$ from the posterior of the risk: 14 % aleatory and 86 % epistemic, a predictive standard deviation of R$ 1.52 million, as the draws give. The year's loss is dominated by what is not known about the rate, not by the chance of the count.

Figure 25.4 draws the three risk curves. The horizontal axis is a loss x in reais per year on a logarithmic scale; the vertical axis is the probability of exceeding it, also on a logarithmic scale. The solid blue curve is the exact posterior of the annual risk, the dotted blue curve the same from the state; the black staircase is the loss of a year at the plug-in rate, which jumps at each additional loss of cooling; the purple dashed curve is the exact predictive loss of a year. The red dotted lines mark the exact 95th percentile of the posterior of the risk, where its curve crosses 0.05.

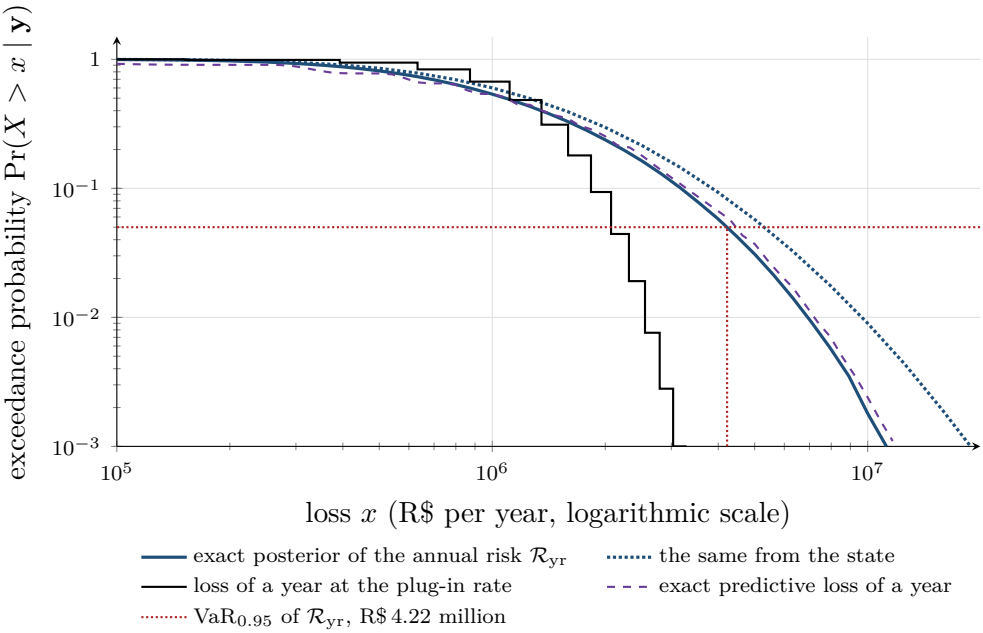


Figure 25.4: Risk curves of the hall's cooling: the exceedance probability of the posterior of the annual risk (solid), of the loss of a year at the plug-in unavailabilities, R$ 152 371 + R$ 240 000 K_f with $K_f \sim \text{Poisson}(4.585)$ (staircase), and of the predictive loss of a year, the same with $K_f \sim \text{Poisson}(\nu_{\text{loss}}(\mathbf{q}))$ averaged over the posterior (dashed); MDT = 12 h, 2×10^5 draws, seed 24. Dotted blue: the posterior of the annual risk from the state, the normal on its logarithm at $\hat{\mathbf{z}}$ with standard deviation 0.878.

Reading the curves. The staircase of the known rate is steep: it falls from 0.94 beyond one loss of cooling to 0.04 beyond eight, and no year at that rate costs more than R$ 3 million with any appreciable probability. The curve of the risk's posterior is far less steep, and the predictive curve lies close to it in the tail, slightly above, with the steps of the count still visible in the middle. The state's curve lies above the exact one in the tail and crosses 0.05 near R$ 5.3 million. The exact upper curves cross 0.05 near R$ 4.2 and R$ 4.4 million; the staircase at R$ 2.1 million. The 95th percentile of the risk's posterior, R$ 5.31 million with a tail mean of R$ 8.16 million (R$ 4.22 and R$ 5.92 million exactly), therefore answers the question in the sense of a bad state of knowledge: if q_C is in the upper part of its band, the hall's expected loss is five to

eight million a year (four to six exactly). The predictive, R$ 6.29 million with a tail mean of R$ 9.68 million (R$ 4.38 and R$ 6.20 million exactly), answers it in the sense of a bad year: a year's loss exceeds R$ 6.3 million with probability 0.05 on the present records (R$ 4.4 million exactly). The plug-in staircase, R$ 2.07 million, understates the bad year by half or more by either route, because it treats the loss rate as known.

25.3.5 The risk reduction of a design change

Engineering question. How much risk does each design change remove? The two changes are those of Chapter 14: a redundant controller, which replaces the basic event K by the AND gate K1 ∧ K2, and a fifth CRAC unit, which makes the CRAC group 3-out-of-5; and both together.

General form. The expected consequence of Chapter 2, equation (2.23), now over several actions: the actions are ranked by $\mathcal{R}(a) = \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$, the best one minimizing it, and the firmness of the ranking is the posterior probability $\Pr(L(a, \mathbf{z}) < L(a', \mathbf{z}) \mid \mathbf{y})$ of equation (2.24). A change to the plant brings no new evidence, so the design form (2.16) collapses to an expectation on the present posterior, as in equation (14.40).

Specialized form. The question is equation (2.23) with each action a a design, $L(a, \mathbf{z}) = \mathcal{R}_{\text{yr}}(\mathbf{q}(\mathbf{z}); a)$, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of equation (25.8), the firmness of the ranking by equation (2.24). Two choices enter here. The new units share the posteriors of the old, the second controller Beta(1, 999) and the fifth CRAC unit q_C, and fail independently given them: this values the changes without records of their own and keeps the state at three entries, and it fails if a new unit differs from the old ones or a common cause takes both controllers down (Section 25.6.6). The schedule is extended to the new configurations: one controller down with the cooling kept is given no consequence, and with five CRAC units the cooling is kept with up to two of them down, each costing R$ 150 h^{-1}; this keeps the schedule's rule, a cost per unit down, and it would understate a fifth unit's cost if two units down left the hall without margin at all.

Each design a gives a new enumeration, over 2^8 configurations (2^9 for both changes), and a new risk $\mathcal{R}_{\text{yr}}(\mathbf{q}; a)$ on the same posterior, with its second-order expectation (25.18) and its band and tail of equations (25.19) and (25.28). Its reduction is the derived quantity

$$\Delta\mathcal{R}_{\text{yr}}(\mathbf{q}; a) = \mathcal{R}_{\text{yr}}(\mathbf{q}; \text{present}) - \mathcal{R}_{\text{yr}}(\mathbf{q}; a), \tag{25.35}$$

with its own gradient, so that its band carries the correlation between the two risks: both are large when q_C is large. The redundant controller replaces q_K by q_K^2 and the fifth unit replaces the CRAC factor by the probability that at most two of five units are down; the exact expectation follows from the moments, and the exact band draw by draw. With a certain cost the loss of equation (25.12) becomes $L(a, \mathbf{z}) = c_a + \mathcal{R}_{\text{yr}}(\mathbf{q}(\mathbf{z}); a)$, with c_a the annualized cost of the change, so $\mathcal{R}(a) = c_a + \mathbb{E}[\mathcal{R}_{\text{yr}}(\mathbf{q}; a) \mid \mathbf{y}]$, and the action is worth taking when its expected reduction exceeds c_a. Which of two changes removes more is, draw by draw, the threshold event $\Delta\mathcal{R}_{\text{yr}}(\mathbf{q}; a) > \Delta\mathcal{R}_{\text{yr}}(\mathbf{q}; a')$, the event $L(a, \mathbf{z}) < L(a', \mathbf{z})$ with the certain costs left out. Its probability is the second line of equation (2.24), with the gradient $\mathbf{g}_a - \mathbf{g}_{a'}$ of the difference of the two annual risks in the logits at $\hat{\mathbf{z}}$ and the covariance $\boldsymbol{\Sigma}$; the exact value beside it comes from the draws.

Building blocks. The enumerations of the new trees; $\hat{\mathbf{z}}$, $\boldsymbol{\Sigma}$, the diagonal Hessians and the gradients $\mathbf{g}_a$ of each design and each reduction; and, for the exact values beside them, the moments $\mathbb{E}[q_C^h]$ to $h = 5$ and $\mathbb{E}[q_K^2]$ from equation (14.20) and draws of the three posteriors, common to all designs.

Risk reduction by design changes. Table 25.7 compares the four designs. Each design has two rows, from the state and exact; the columns give its expected annual risk, its median and central 95 % band, its acceptance probability at R$ 1.5 million, and the expected reduction from the present design with its central 95 % band.

Table 25.7: The annual risk of the hall's cooling under the present design and three alternatives. State: second-order expectations, medians at $\hat{\mathbf{z}}$, delta bands and normal tails. Exact: Beta moments for the expectations, and 2×10^5 draws, seed 24, for the rest. Amounts in reais per year.

Design	Expected risk [R$/yr]	Median [R$/yr]	95 % band [R$/yr]	Pr(< R$ 1.5 M) [–]	Expected reduction [R$/yr]	Band of the reduction [R$/yr]
present, state	1 913 773	1 252 736	[224 298, 6 996 702]	0.581	–	–
exact	1 513 085	1 085 231	[192 552, 5 343 835]	0.646	–	–
redundant controller, state	1 653 861	1 078 789	[149 806, 7 768 605]	0.628	259 912	[24 573, 1 231 329]
exact	1 339 572	895 930	[101 814, 5 158 615]	0.696	173 513	[4 353, 636 364]
fifth CRAC unit, state	631 411	442 377	[142 143, 1 376 770]	0.982	1 282 361	[92 061, 7 133 117]
exact	493 677	416 023	[121 499, 1 320 805]	0.985	1 019 408	[32 405, 4 105 556]
both changes, state	369 813	267 620	[67 777, 1 056 697]	0.993	1 543 959	[159 314, 6 091 459]
exact	319 145	241 677	[64 565, 1 052 222]	0.992	1 193 940	[110 563, 4 298 730]

Figure 25.5 draws the four posteriors. The vertical axis lists the designs; the horizontal axis is the annual risk on a logarithmic scale. For each design the solid horizontal line is the exact central 95 % band, the filled circle the exact expected risk and the open circle the exact median; the dotted blue line just below it is the band from the state, with its expected risk as a blue square; the vertical dashed line marks the R$ 1.5 million limit.

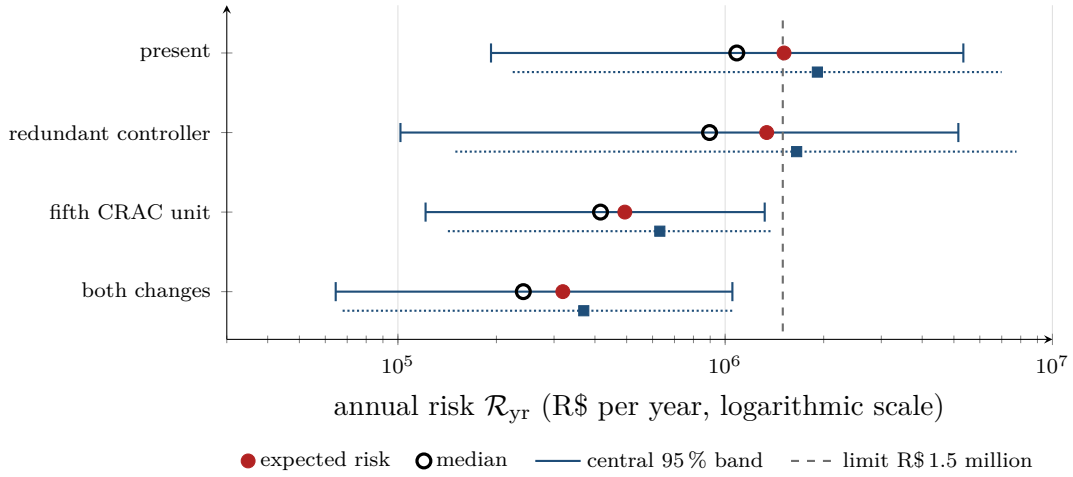


Figure 25.5: The posterior of the annual risk of the hall's cooling under the present design and three alternatives: expected risk (filled circles, exact from the Beta moments), median (open circles) and central 95 % band (lines) from 2×10^5 draws, seed 24, with $q_C \sim \text{Beta}(3, 97)$, $q_P \sim \text{Beta}(4, 396)$, $q_K \sim \text{Beta}(1, 999)$ shared by the units of each group. Dotted blue, with squares: the band and the second-order expected risk from the state.

The fifth CRAC unit. The fifth unit lowers the expected risk from R$ 1.91 to R$ 0.63 million per year, a reduction of R$ 1.28 million, 67 %; its loss-state term falls from R$ 1 744 236 to R$ 396 553, by 77 %, and its degraded terms rise by R$ 65 322. The values beside them come from the exact posterior: from R$ 1.51 to R$ 0.49 million, a reduction of R$ 1.02 million, 67 %. Its exact loss-state term falls from R$ 1 364 812 to R$ 290 281, a reduction of R$ 1 074 531 or 78.7 %, the 79 % by which Chapter 14 found it lowers the top event. Its degraded CRAC term rises from R$ 139 689 to R$ 194 759, by R$ 55 070: five units are down one at a time more often than four, and the configurations with two CRAC units down, which were losses, are now degraded at R$ 300 h^{-1}. Its degraded pump term rises from R$ 8 584 to R$ 8 638, by R$ 53, because the CRAC group now keeps the cooling more often, so that more of the configurations with one pump down are cooled rather than lost. The net reduction is the loss-state reduction less both increases, $1\,074\,531 - 55\,070 - 53 = \text{R\$}\,1\,019\,408$. Its band is wide, [R$ 0.09, R$ 7.1] million ([R$ 0.03, R$ 4.1] exactly), because it is large exactly when q_C is large; in 0.04 % of the draws, those with $q_C < 0.0013$, the fifth unit raises the risk, since the hall then almost never loses its cooling and the extra unit only adds degraded hours. With the fifth unit the band of the risk lies almost entirely below the limit: the acceptance probability at R$ 1.5 million rises from 0.581 to 0.982 (0.65 to 0.985 exactly), and at R$ 1.0 million to 0.92 (0.93 exactly).

The redundant controller. The second controller lowers the expected risk by R$ 259 912 per year, 14 %, with the band [R$ 24 573, R$ 1 231 329]; exactly, by R$ 173 513 per year, 11 %; its exact loss-state term falls by R$ 173 662, 12.7 %, and its degraded terms rise by R$ 148, since the cooling is now kept in the configurations with one controller down. The reduction is nearly the controller's whole attributed share, because the change replaces q_K by q_K^2 and the risk is linear in q_K. Its band, [R$ 4 353, R$ 636 364], is that of q_K. The acceptance probability rises only to 0.63 (0.70 exactly): the band of the risk is set by q_C, which the second controller does not touch.

The comparison. By the second line of equation (2.24), the difference of the two expected risks is $1\,653\,861 - 631\,411 = \text{R\$}\,1\,022\,449$ and the spread of the gradient of their difference is R$ 917 100, so that the fifth unit removes more risk than the second controller with posterior probability $\Phi(1.115) = 0.868$, and on expected values 4.9 times as much; draw by draw over the exact posterior the probability is 0.852 and the ratio 5.9. Both changes together remove R$ 1.54 million and bring the expected risk to R$ 0.37 million (R$ 1.19 and R$ 0.32 million exactly). At expected values, the fifth unit is worth buying if its annualized cost is below about R$ 1.28 million a year (R$ 1.02 million exactly) and the second controller if its cost is below about R$ 0.26 million (R$ 0.17 million); the prices, their annualization and the choice under uncertainty about both are the subject of a later chapter on redundancy and reliability investment.

25.4 The checks (step 5)

The hall has no physics row and no exact constraint, so no limit $\sigma_p \to 0$ of a model tolerance arises and no limit table is needed; the bound $0 < q < 1$ is not a row either, and the logit state handles it. What can fail is the Laplace route against the exact one, the enumeration and the draws, the choice of each record alone as the posterior, and the pinned inputs of the consequence model.

25.4.1 The Laplace route against the draws

The derivatives of the risk at the plug-in values, from equations (25.2) and (25.14), are $\partial\mathcal{R}/\partial q_C =$ R$ 7 263 h^{-1} for the shared CRAC entry, of which R$ 6 767 is the loss-state part $C_{\mathrm{L}} \times 0.3384$

and R$ 496 the degraded part; $\partial\mathcal{R}/\partial q_P = \text{R\$}\,494.6\,\text{h}^{-1}$; and $\partial\mathcal{R}/\partial q_K = \text{R\$}\,19\,877\,\text{h}^{-1}$. With the plug-in risk $\text{R\$}\,143.006\,\text{h}^{-1}$ the gradient (25.19) of $\ln\mathcal{R}$ in the logits and the variance budget with $\boldsymbol{\Sigma}$ of equation (25.8) are

$$\mathbf{g} = \frac{1}{143.006}\begin{bmatrix} 7262.9 \times 0.0291 \\ 494.6 \times 0.0099 \\ 19\,876.9 \times 0.000999 \end{bmatrix} = \begin{bmatrix} 1.478 \\ 0.034 \\ 0.139 \end{bmatrix},$$

$$\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g} = \sum_j g_j^2\,\Sigma_{jj} = 0.7506 + 0.0003 + 0.0193 = 0.7702 :$$

the CRAC units' entry carries 97.5 % of the variance of $\ln\mathcal{R}$, the controller 2.5 % and the pumps nothing that matters, and $\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}} = 0.878$.

For the expected risk, equation (25.18) needs the diagonal of the Hessian of the annual risk in the logits at $\hat{\mathbf{z}}$, $\operatorname{diag}\mathbf{H}_a = (3\,285\,728,\ 75\,967,\ 173\,599)$ reais per year, and with the diagonal $(0.34364,\ 0.25253,\ 1.00100)$ of $\boldsymbol{\Sigma}$

$$\tfrac{1}{2}\operatorname{tr}(\mathbf{H}_a\boldsymbol{\Sigma}) = 564\,558 + 9\,592 + 86\,887 = \text{R\$}\,661\,036,$$
$$\mathcal{R}(a) \approx 1\,252\,736 + 661\,036 = \text{R\$}\,1\,913\,773 \text{ per year},$$

26 % above the exact R$ 1 513 085. The controller's term, R$ 86 887, is the curvature of a risk that is linear in q_K but not in ξ_K. The error is in the Laplace posterior itself, not in the truncation of the expansion: the average of the annual risk over the Laplace normal of the logits is R$ 2 084 301, 38 % high, because a normal on the logit of q_C has a heavier upper tail than Beta(3, 97), and the second-order term moves the plug-in toward that average and past the exact value. The two routes move the plug-in in the same direction, up.

For the ranking of the two changes, the gradients of the two annual risks in the logits at $\hat{\mathbf{z}}$ give the gradient of their difference, $\mathbf{g}_{\text{ctrl}} - \mathbf{g}_{\text{fifth}} = (1\,535\,876,\ -159,\ -174\,410)$ reais per year, and $\sqrt{(\mathbf{g}_{\text{ctrl}} - \mathbf{g}_{\text{fifth}})^{\mathsf{T}}\boldsymbol{\Sigma}(\mathbf{g}_{\text{ctrl}} - \mathbf{g}_{\text{fifth}})} = \text{R\$}\,917\,100$. With the second-order expected risks of equation (25.18), R$ 1 653 861 for the redundant controller and R$ 631 411 for the fifth unit, the second line of equation (2.24) gives $\Phi(1\,022\,449/917\,100) = \Phi(1.115) = 0.868$ that the fifth unit removes more, against 0.852 from the draws; centered at the plug-in risks, the difference R$ 636 412 gives $\Phi(0.694) = 0.756$.

Table 25.8 sets the Laplace route beside the draws of the exact posterior. Each row is a summary of the annual risk or of the comparison of the designs; the columns give it from the exact posterior, over the draws or from the moments, and by the Laplace route: the second-order expectation (25.18), the delta method on the logarithm, equations (25.19) and (25.28), centered at the plug-in risk, and the comparison of equation (2.24).

The delta method gets the width and the budget right, a standard deviation of 0.878 against 0.852, and places the band too high: its center, the plug-in risk, lies 17 % above the draws' center $\text{e}^{13.883} = \text{R\$}\,1\,069\,853$, the result Chapter 14 found for the top event, since the plug-in value of a quantity that grows as q_C^2 sits above its median when the posterior of q_C is skewed to the right. Its acceptance probabilities are 0.06 lower at each limit for the same reason, and they do not change the verdict. The exact variances of the logits, the trigamma values of equation (14.9), widen the delta band to [R$ 191 683, R$ 8 187 191] and move it no closer. On the risk itself the delta method fails outright: its standard deviation, R$ 1 099 409, puts 12.7 % of a normal centered at the plug-in risk on negative losses. The Laplace comparison of the designs agrees with the draws to 0.02 and leaves the ranking as it was.

25.4.2 The enumeration, the cut-set sum and the draws

The sum over configurations. The direct sum of equation (25.10) over all 128 configurations at the plug-in values, each with its consequence from equation (25.3), gives $\text{R\$}\,143.006434\,\text{h}^{-1}$,

Table 25.8: The Laplace route for the annual risk of the hall's cooling, the second-order expectation, the delta method on $\ln \mathcal{R}$ and the comparison of the designs with $\mathbf{g}$, $\mathbf{H}_a$ and $\boldsymbol{\Sigma}$ at $\hat{\mathbf{z}}$, against the exact posterior: its moments for the expectation, its 2×10^5 draws, seed 24, for the rest.

Summary	Exact posterior	Laplace
expected annual risk $\mathcal{R}(a)$ [R$/yr]	1 513 085	1 913 773
center, exp of the mean of $\ln \mathcal{R}_{\text{yr}}$ [R$/yr]	1 069 853	1 252 736
standard deviation of $\ln \mathcal{R}_{\text{yr}}$ [–]	0.852	0.878
central 95 % band [R$/yr]	[192 552, 5 343 835]	[224 298, 6 996 702]
$\Pr(\mathcal{R}_{\text{yr}} <$ R$ 1.0 million $\mid \mathbf{y})$ [–]	0.463	0.399
$\Pr(\mathcal{R}_{\text{yr}} <$ R$ 1.5 million $\mid \mathbf{y})$ [–]	0.646	0.581
$\Pr(\mathcal{R}_{\text{yr}} <$ R$ 2.0 million $\mid \mathbf{y})$ [–]	0.762	0.703
Pr(the fifth unit removes more $\mid \mathbf{y}$) [–]	0.852	0.868

equal to the closed form of equations (25.2) and (25.14) to every printed digit; the probabilities of the five classes sum to one within 10^{-9}, and the printed values of Section 25.3.1 sum to 0.99998, to their rounding. The ordered partition (25.26) reproduces the loss-state term in every draw.

The state sum against the cut-set sum. The rare-event sum of equation (25.25) has expectation R$ 1 445 996 per year against the exact loss-state term R$ 1 364 812: an overcount of 5.95 %, the 5.9 % by which Chapter 14's rare-event sum overstated the expected top event, scaled by $C_{\text{L}} t_{\text{yr}}$, as equation (25.25) requires. A risk computed from the cut-set sum is an upper bound for the loss-state term; in expectation it is R$ 81 184 per year too high, more than half of the whole degraded term.

The draws against the exact mean. The draws give a mean of R$ 1 514 242 against the exact R$ 1 513 085. The standard deviation of the draws is R$ 1 403 554, so the Monte Carlo standard error of the mean is $1\,403\,554/\sqrt{2 \times 10^5} =$ R$ 3 138, and the difference is 0.37 standard errors.

25.4.3 Sensitivity to the prior

Each posterior was taken as its record alone, and the three records hold three, four and one units found down. Table 25.9 repeats the risk under the two priors of Section 14.4.3, the Jeffreys prior Beta$(\frac{1}{2}, \frac{1}{2})$ and the uniform prior Beta$(1, 1)$ added to every record, each refitted with its prior row. Each row is a prior; the columns give the expected annual risk, its median, the acceptance probability at R$ 1.5 million, the shares of the expected risk that the fifth CRAC unit and the redundant controller remove, and the probability that the fifth unit removes more, each from the state with the exact value in parentheses.

The level moves and the verdicts do not. The expected risk rises by 25 % under the Jeffreys prior and by 51 % under the uniform one (29 and 60 % exactly), as the expected top event of Chapter 14 did, because a pseudo-count of one half or one is not small beside three, four and one units found down. Every prior moves the risk up, so the acceptance probability at R$ 1.5 million falls, to 0.45 and 0.33 (0.51 and 0.38 exactly): the plant is not accepted at an assurance of 0.9 under any of the three, and under the uniform prior the expected risk itself is nearly twice the limit (more than half as large again exactly). The fifth CRAC unit removes two thirds of the expected risk and the second controller a seventh to a ninth under every prior, and the fifth unit removes more with probability 0.87 (0.85 to 0.86 exactly). A dispersion check, whether the

Table 25.9: Sensitivity of the hall's risk to the prior on each group's unavailability. State: as in Table 25.7, refitted under each prior. In parentheses, exact: expected risks and removals from the Beta moments, medians and probabilities from 2×10^5 draws, seed 24. Removals as fractions of the expected risk of the present design under the same prior.

Prior	Expected risk [R$/yr]	Median [R$/yr]	Pr(< R$ 1.5 M) [–]	Fifth unit removes [–]	Second controller removes [–]	Fifth removes more [–]
record alone	1 913 773	1 252 736	0.581	67.0 %	13.6 %	0.868
	(1 513 085)	(1 085 231)	(0.646)	(67.4 %)	(11.5 %)	(0.852)
Jeffreys,	2 385 398	1 660 614	0.449	66.6 %	14.5 %	0.869
Beta($\frac{1}{2}$, $\frac{1}{2}$)	(1 950 920)	(1 471 002)	(0.510)	(66.5 %)	(13.3 %)	(0.852)
uniform,	2 882 356	2 099 834	0.325	66.4 %	14.9 %	0.873
Beta(1, 1)	(2 417 104)	(1 891 301)	(0.375)	(66.1 %)	(14.3 %)	(0.861)

four CRAC units share one unavailability or one of them fails more often, would compare the units' own counts with the binomial spread of a common one; the records are kept as group totals, so the check cannot be made from them, as in Chapter 14.

25.4.4 Sensitivity to the consequence schedule and the down time

Table 25.10 changes the schedule. Each row is a schedule; the columns give the expected annual risk and its ratio to the present one, the acceptance probability at R$ 1.5 million, the share of the CRAC pairs in the loss-state term, and the expected reductions of the fifth CRAC unit and of the redundant controller of Section 25.3.5, each from the state with the exact value in parentheses.

Table 25.10: The risk of the hall's cooling under changes of the consequence schedule. State: second-order expectations and normal tails at $\hat{\mathbf{z}}$. In parentheses, exact: expected risks and expected reductions from the Beta moments, each linear in the rates; acceptance probabilities from 2×10^5 draws, seed 24. Amounts in reais per year.

Schedule	Expected risk [R$/yr]	Ratio [–]	Pr(< R$ 1.5 M) [–]	CRAC share [–]	Fifth unit [R$/yr]	Second controller [R$/yr]
present	1 913 773	1 (1)	0.581	0.843	1 282 361	259 912
	(1 513 085)		(0.646)	(0.864)	(1 019 408)	(173 513)
loss rate	3 658 008	1.911	0.309	0.843	2 630 044	520 069
doubled	(2 877 897)	(1.902)	(0.366)	(0.864)	(2 093 939)	(347 175)
loss rate halved	1 041 654	0.544	0.818	0.843	608 519	129 833
	(830 679)	(0.549)	(0.866)	(0.864)	(482 143)	(86 683)
degraded rates	2 083 309	1.089	0.531	0.843	1 217 039	259 667
doubled	(1 661 359)	(1.098)	(0.598)	(0.864)	(964 285)	(173 365)
every rate	3 827 545	2.000	0.279	0.843	2 564 722	519 824
doubled	(3 026 170)	(2.000)	(0.334)	(0.864)	(2 038 816)	(347 027)

Doubling the cost of a lost hall multiplies the expected risk by 1.91 (1.90 exactly), not by two, because the degraded tenth does not move; doubling every rate multiplies it by exactly two, as linearity requires. The share of the CRAC pairs in the loss-state term does not change in any row, and neither does the order of the two design changes: the fifth unit removes four and a half to five times as much as the second controller under every schedule (five and a half to six exactly), and the attribution and the design ranking are properties of the probabilities. The acceptance probability changes in every row, from 0.28 to 0.82 (0.33 to 0.87 exactly), because a limit in reais is compared with a risk that scales with the schedule. Linearity in the consequence makes the attribution robust to an error in the schedule and the acceptance sensitive to it: by

equation (25.27), doubling every rate gives $\Pr(2\mathcal{R}_{\mathrm{yr}} < \mathcal{R}^*) = \Pr(\mathcal{R}_{\mathrm{yr}} < \mathcal{R}^*/2)$, and in general an error of some factor in the schedule is an error of the same factor in the limit, read on a posterior that spans a factor of 31 (28 exactly).

The mean down time enters the year's loss and not the risk, equation (25.31). With MDT = 6 h or 24 h in place of 12 h the plug-in frequency of losses is 9.170 or 2.292 a year, the predictive probability that a year costs less than R$ 1.5 million is 0.540 or 0.493 against 0.524 (0.644 or 0.596 against 0.633 exactly), and the predictive 95th percentile of a year's loss is R$ 6.21 or R$ 6.51 million against R$ 6.29 million (R$ 4.35 or R$ 4.60 against R$ 4.38 million): fewer, longer losses spread the year more, as equation (25.32) requires, and the bad year moves by a few percent because its variance is mostly epistemic.

25.5 The report (step 6)

The records do not support the first reading. Priced at the fractions found down, the hall's cooling costs R$ 1.25 million a year, below the owner's R$ 1.5 million limit; the expected loss the records imply is R$ 1.91 million from the posterior of the state (R$ 1.51 million exactly), above it by both, and the probability that the annual risk lies below the limit is 0.58 (0.65 exactly), short of the assurance of 0.9 the statement requires, and lower still, 0.45 and 0.33 (0.51 and 0.38), under the Jeffreys and the uniform priors. The figure at the fractions is low because the loss of cooling needs two CRAC units down together and the CRAC unavailability, from three units found down, is uncertain by a factor of eleven. A bad year is costlier still: one year in twenty, on the present records, costs more than R$ 6.3 million (R$ 4.4 million exactly). The answers below are given in the units in which the questions were asked.

25.5.1 The answers and their bands

Table 25.11 gives the answers. Each row is a question of Table 25.2, in its order; the second column gives the answer from the posterior of the state, the records alone, the third the exact value beside it, and the fourth its band or comparison: the central 95 % posterior interval, a posterior probability, or the range over the three priors of Table 25.9 where the prior moves the answer.

The report line. The hall's cooling is expected to cost R$ 1.91 million a year from the posterior of the state (R$ 1.51 million from the exact posterior), above the R$ 1.5 million limit by both, nine tenths of it in hours without cooling and one tenth in hours with a unit down. The annual risk, the rate at which the cooling loses money, is known only within a band: on the present records it could be as low as R$ 0.2 million a year or as high as R$ 7.0 million (R$ 5.3 million exactly), because a hundred checks of the CRAC units leave their unavailability uncertain by a factor of eleven. The probability that the annual risk is below the R$ 1.5 million limit is 0.58 (0.65 exactly), which does not meet an assurance of 0.9. Any two of the four CRAC units failing together account for about four fifths of the cost, the controller for a tenth and the pumps for a little over one percent. A fifth CRAC unit would remove about R$ 1.3 million a year (R$ 1.0 million exactly) and bring the probability of meeting the limit to 0.982 (0.985); a second controller would remove about R$ 0.26 million (R$ 0.17 million). In one year in twenty, on the present records, the loss would exceed R$ 6.3 million (R$ 4.4 million exactly). The figures assume that units fail independently given their unavailabilities, that a loss of cooling lasts 12 hours on average, and that the costs per hour do not depend on the load.

Table 25.11: The answers to the six questions asked of the hall's cooling with its consequence schedule. Amounts in reais per year; bands are central 95 % posterior intervals, by the delta method for the state; exact expected values from the Beta moments, exact bands and probabilities from 2×10^5 draws.

Question	Answer (state route)	Exact	Band and comparison
What is the expected consequence?	expected annual loss R$ 1.91 million; median R$ 1.25 million	R$ 1.51 million; median R$ 1.09 million; band [R$ 0.19, R$ 5.34] million, a factor of 28	[R$ 0.22, R$ 7.00] million, a factor of 31; the CRAC units' entry carries 97.5 % of the variance of $\ln \mathcal{R}$; R$ 2.39 and R$ 2.88 million under the Jeffreys and uniform priors (R$ 1.95 and R$ 2.42 exact)
Plug-in against expected	plug-in R$ 1.25 million, 35 % below the expectation	17 % below; the gap R$ 0.26 million, in the loss states, convex in the shared q_C; $\frac{1}{2}\,\mathrm{tr}(\mathbf{\Sigma_q H_q})$ gives R$ 0.27 million	the gap R$ 0.66 million is $\frac{1}{2}\,\mathrm{tr}(\mathbf{H}_a\mathbf{\Sigma})$, in the logits; the exact degraded terms have a small negative gap
Where the risk comes from	loss states 91 % (R$ 1.74 million), degraded states 9 % (R$ 0.17 million)	90 % (R$ 1.36 million), 10 % (R$ 0.15 million); pairs R$ 1.18, controller R$ 0.17, pumps R$ 0.02 million; loss share [0.73, 0.94]; cut-set sum 5.95 % above the state sum	six CRAC pairs R$ 1.47 million, controller R$ 0.25 million, pumps R$ 0.02 million; loss share of the risk [0.75, 0.95]
Below R$ 1.5 million with assurance 0.9	Pr = 0.581; not accepted	Pr = 0.646; not accepted; 0.463 at R$ 1.0 million, 0.762 at R$ 2.0 million; limit at assurance 0.9, R$ 3.18 million; 0.51 and 0.38 under the priors	0.399 at R$ 1.0 million, 0.703 at R$ 2.0 million; the limit met with assurance 0.9 is R$ 3.86 million; 0.45 and 0.33 under the Jeffreys and uniform priors
The bad year	95th percentile of the risk R$ 5.31 million, tail mean R$ 8.16 million	R$ 4.22 and R$ 5.92 million; a year's loss: R$ 4.38 and R$ 6.20 million; variance 86 % epistemic	a year's loss: predictive 95th percentile R$ 6.29 million, tail mean R$ 9.68 million; at the plug-in rate R$ 2.07 million; variance 92 % epistemic
Risk removed by each change	fifth CRAC unit R$ 1.28 million (−67 %); redundant controller R$ 0.26 million (−14 %)	R$ 1.02 million (−67 %), R$ 0.17 million (−11 %); bands [R$ 0.03, R$ 4.1], [R$ 0.004, R$ 0.64] million; probability 0.852 (0.85 to 0.86 over the priors); both R$ 1.19 million	bands [R$ 0.09, R$ 7.1] and [R$ 0.02, R$ 1.2] million; fifth unit removes more with probability 0.868 (0.87 over the priors); both R$ 1.54 million

25.5.2 What the expected loss is and is not

The expected annual loss, R$ 1.91 million (R$ 1.51 million exactly), is the average over many years of the cost the hall's cooling incurs, averaged also over what is not known about its units. It is not the loss of a typical year: the median of the predictive is R$ 1.37 million (R$ 1.09 million exactly), and in about a third of the years (almost half exactly) the loss is below a million. It is not the loss rate either: the rate is uncertain by a factor of 31 (28 exactly), and the mean is a weighted average over that uncertainty in which the upper tail weighs heavily. It is the right number for a decision taken many times, or for a comparison of designs on expected cost, which is how Section 25.3.5 uses it. For a decision taken once against a limit, the posterior probability of meeting the limit is the right number, and for a budget that must hold in a bad year, the predictive value at risk. The three are different numbers, and a report that gives only the first gives the least of the three for acceptance.

The probabilistic risk assessments of large plants often present the risk as a family of risk curves, one aleatory exceedance curve for each draw of the uncertain parameters, summarized by the curves of their percentiles. The staircase of Figure 25.4 is one member of that family, the member at the posterior means; the family spreads to either side of it as q_C moves through its band, and the predictive curve is its average. The posterior of the risk is the distribution of the means of the family's members.

Common mistake

Accepting or rejecting on a point value of the risk. With the limit at R$ 1.5 million, the exact expected risk R$ 1.513 million rejects the plant by a margin of one percent, and the plug-in risk, R$ 1.25 million and 17 % below that expectation because the CRAC units share one posterior, accepts it by a margin of seventeen; the posterior probability of meeting the limit is 0.58 (0.65 exactly) whichever point is used, and that probability, set against a stated assurance, is the acceptance criterion.

Common mistake

Reading the 95th percentile of the risk's posterior as the loss of a bad year, or the loss of a year at the plug-in rate as the bad year. The first is a bad loss rate, R$ 5.31 million (R$ 4.22 exactly), and the second a bad year at a known rate, R$ 2.07 million; the bad year on the present records is the predictive percentile, R$ 6.29 million (R$ 4.38 exactly), and 92 % of its variance (86 % exactly) comes from the uncertainty of the rate.

25.6 The next measurement (step 7)

The acceptance answer is the one the report leaves open, and it is open because the posterior of q_C is wide. The design answer does not wait on a record: the fifth CRAC unit removes more than the second controller with posterior probability 0.868 (0.852 exactly), removes two thirds of the expected risk under every prior tried, and ranks first under every schedule tried. The next measurement is the record that would narrow the band of the risk and decide its acceptance.

25.6.1 The candidate records

Four candidates are weighed, as in Chapter 14: a hundred or four hundred further checks of the CRAC units, kept as the present record was, and a hundred or a thousand further checks of the controller. For N further CRAC checks with n_{d} units found down, the state is refitted, with the count row added, at $\hat{\xi}_C = \ln((3+n_{\mathrm{d}})/(97+N-n_{\mathrm{d}}))$ and $\Sigma_{CC} = 1/(3+n_{\mathrm{d}}) + 1/(97+N-n_{\mathrm{d}})$; each count is weighted by the Laplace predictive, the binomial probability of the count averaged by Gauss–Hermite over $\mathcal{N}(\hat{\xi}_C, \Sigma_{CC})$; and for each count the acceptance probability (25.28)

at R$ 1.5 million and the band of the risk are recomputed. A record of controller checks is treated in the same way on ξ_K. The values beside them come from the exact posterior, $\mathrm{Beta}(3 + n_\mathrm{d},\ 97 + N - n_\mathrm{d})$ with the Beta–binomial predictive of equation (14.47), where by the law of total probability, equation (14.48), the acceptance probability averaged over the counts is its present value for every candidate; the state route, whose tail is not a martingale under the Laplace predictive, moves it by up to 0.02. What a record can do is push it toward one end or the other. A record settles the question when, after it, the acceptance probability is at least 0.9, the plant accepted, or at most 0.1, the plant rejected. The fifth CRAC unit is set beside the records for comparison: it changes the plant rather than the evidence, and its acceptance probability needs no further record.

25.6.2 The preposterior covariance

Under the Laplace posterior the width of the risk's band after a record has a closed form. A record of N checks of group j is one more count row on entry j, of the form of equation (25.5), whose Jacobian row at $\hat{\mathbf{z}}$ is $\mathbf{J}_a = \sqrt{N q_j (1 - q_j)}\ \mathbf{e}_j^\mathsf{T}$. Information adds, so the preposterior covariance of equation (2.17), written for this record as in equation (14.49), and the band of $\ln \mathcal{R}$ after it are

$$\boldsymbol{\Sigma}_a = \left(\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^\mathsf{T} \mathbf{J}_a\right)^{-1}, \qquad \mathrm{sd}(\ln \mathcal{R} \mid \mathbf{y}, a) \approx \sqrt{\mathbf{g}^\mathsf{T} \boldsymbol{\Sigma}_a \mathbf{g}}, \qquad w_a = \frac{2 \times 1.96}{\ln 10} \sqrt{\mathbf{g}^\mathsf{T} \boldsymbol{\Sigma}_a \mathbf{g}}, \quad (25.36)$$

with $\mathbf{g}$ the gradient of equation (25.19) and w_a the width of the central 95 % band in decades; the outcome of the record does not enter $\boldsymbol{\Sigma}_a$, so no average over outcomes is needed. The Jacobian rows are (1.706, 0, 0) for a hundred CRAC checks, (3.412, 0, 0) for four hundred, (0, 0, 0.316) for a hundred controller checks and (0, 0, 0.9995) for a thousand. A hundred CRAC checks take Σ_{CC} from 0.344 to 0.172 and four hundred to 0.069; a hundred controller checks take Σ_{KK} from 1.001 to 0.910 and a thousand to 0.500; each record leaves the other entries where they were. Through the gradient, $\mathbf{g}^\mathsf{T} \boldsymbol{\Sigma}_a \mathbf{g}$ falls from 0.770 to 0.395 and 0.170 for the CRAC records and only to 0.768 and 0.761 for the controller records, band widths of 1.07, 0.70, 1.49 and 1.49 decades against 1.49 today. A CRAC check adds $q_C(1 - q_C) = 0.0291$ to the diagonal entry of $\mathbf{J}^\mathsf{T}\mathbf{J}$ it feeds and a controller check 0.000999, twenty-nine times less, and the gradient weights the CRAC entry ten times as heavily as the controller's, 1.478 against 0.139.

25.6.3 The ranking of the records

Table 25.12 gives the result. Each row is a course of action; the columns give the entry the record feeds, the information it adds, the Laplace band width after it by equation (25.36), the expected width of the band of the risk, the probability that the record ends with the plant accepted at an assurance of 0.9, the probability that it ends with the acceptance probability at most 0.1, and the expected acceptance probability, each of the last four from the state with the exact value in parentheses. The exact columns use 10^5 draws with common random numbers across the counts, seed 251, the CRAC or controller unavailability drawn through the Beta quantile function.

A hundred more CRAC checks settle the acceptance at 0.9 only if none of them finds a unit down, an outcome of predictive probability 0.088, after which the probability is 0.94; one unit found down leaves it at 0.87 just short, and three found, the expected count, leave it at 0.61 (exactly: 0.121, 0.96, 0.90 and 0.66). Four hundred checks settle it with probability 0.30 and reject it, with an acceptance probability below 0.1, with probability 0.21 (0.38 and 0.15 exactly); in about half the outcomes the question is still open. The checks narrow the band from 1.49 to 0.70 decades on average (1.44 to 0.74 exactly), and the Laplace widths at $\hat{\mathbf{z}}$ of equation (25.36) follow them closely. Controller checks settle nothing: even a thousand with no unit found down, an outcome of probability 0.38, leave the acceptance probability at 0.61 (0.50

Table 25.12: The preposterior of the acceptance probability at R$ 1.5 million per year for four candidate records, against the fifth CRAC unit, which changes the plant instead of the evidence. A CRAC check is one check of one unit. Added information and Laplace width from equation (25.36); the other columns from the state refitted after each count and weighted by the Laplace predictive, in parentheses from the exact posterior, by 10^5 draws. Widths are those of the central 95 % band of the annual risk. Settled, accepted: the probability that the acceptance probability after the record is at least 0.9; settled, rejected: that it is at most 0.1.

Course of action	Entry fed	Added information [–]	Laplace width [decades]	Expected width [decades]	Settled, accepted [–]	Settled, rejected [–]	Expected Pr(< R$ 1.5 M) [–]
present record	–	–	1.49	1.49 (1.44)	–	–	0.581 (0.646)
100 more CRAC checks	ξ_C	2.91	1.07	1.05 (1.07)	0.088 (0.121)	0.085 (0.052)	0.564 (0.646)
400 more CRAC checks	ξ_C	11.64	0.70	0.70 (0.74)	0.304 (0.377)	0.211 (0.153)	0.563 (0.646)
100 more controller checks	ξ_K	0.10	1.49	1.49 (1.43)	0 (0)	1×10^{-5} (0)	0.578 (0.646)
1000 more controller checks	ξ_K	1.00	1.49	1.45 (1.40)	0 (0)	0.003 (0)	0.561 (0.646)
fifth CRAC unit, present record	–	–	–	0.99 (1.04)	1 (1)	0 (0)	0.982 (0.985)

and 0.67 exactly), and they remove 0.04 decades of the width. The fifth unit narrows the band to about a decade, from 1.49 to 0.99 (1.44 to 1.04 exactly), and moves it below the limit, and it decides the acceptance question at once.

25.6.4 The recommendation

The reported answer whose band the record is to narrow is the annual risk, whose central band runs from R$ 0.22 to R$ 7.00 million, 1.49 decades wide, and with it the acceptance probability, 0.58. In expectation a hundred CRAC checks remove 0.44 decades of that width and four hundred remove 0.79; a hundred controller checks remove nothing visible and a thousand 0.04 (exactly, of the 1.44 decades from R$ 0.19 to R$ 5.34 million, 0.37, 0.70, 0.01 and 0.04). The next record is therefore four hundred further checks of the CRAC units, kept as the present record was, which halve the band in expectation and decide the acceptance one way or the other with probability 0.51, against 0.17 for a hundred (0.53 and 0.17 exactly); both routes rank it first; the record enters as one more count row on ξ_C, and the method returns to step 4 with it. Controller checks do nothing for this question, although they remain the record of the controller's own cut set. The record does not replace the design decision. The operator's choice is between buying assurance with evidence, which may come back unfavorable, and buying it with a unit, which changes the plant and settles the acceptance at once; the value of the first set against its price is the subject of a later chapter on the value of information.

25.6.5 What the analysis showed

The first reading priced the hall's cooling at the fractions found down and signed the statement at R$ 1.25 million a year; the analysis overturns it. The records imply an expected loss of R$ 1.91 million (R$ 1.51 million exactly), above the owner's limit by both, and leave the probability of meeting that limit at 0.58 (0.65), not the nine in ten the statement needs, and lower under any prior added to the records. The reason is the hall's redundancy. Its cooling is lost through coincidences of CRAC units, so the cost grows with the square of an unavailability known from three units found down, and the average over what the records leave open weighs the high

values heavily. Nine tenths of the cost is hours without cooling, and four fifths of the whole passes through the CRAC pairs; the controller, the hall's only single point of failure, carries a tenth. A bad year is not the year at the fractions: one year in twenty costs more than R$ 6.3 million (R$ 4.4 million exactly), two to three times what a known rate would give, and most of that spread is ignorance of the rate rather than the luck of the count. The engineer carries back to the site three things: the statement cannot be signed on the present records; a fifth CRAC unit removes two thirds of the expected cost and would let it be signed, where a second controller would not; and four hundred more CRAC checks narrow the band and are more likely to settle the acceptance than to confirm the rejection, 0.30 against 0.21 (0.38 against 0.15 exactly), but leave it open about half the time, so the cost is reported as a band and not as the figure at the fractions.

25.6.6 Limits of the model

The risk is exact for the drawing, the posterior and the schedule; the three rest on assumptions that the records do not test.

Pinned consequences. The schedule is taken as known. An error in it scales the risk and the acceptance limit together, Table 25.10, and an uncertain cost per hour of lost cooling adds its own spread to the band; Exercise 25.6 gives the size. A later chapter on repair and replacement decisions frees the costs as uncertain quantities.

A single hall. The risk is that of one hall's cooling. A plant of several halls on one chilled-water loop shares the pumps and the controller among them, and the loss of a shared unit costs as many halls' loads at once; the consequence of a configuration is then the sum over the halls it takes down, and the risk is still the sum of equation (25.10) over the configurations of all units.

No recovery dynamics. The consequence is a cost per hour, and the duration of a loss is carried by the unavailability, not by the consequence. A cost that grows with the length of a loss, because a hot hall damages equipment after some hours or because shed load must be restarted, is not a rate, and its expectation needs the distribution of the durations: the mean down time of equation (25.29) and its spread, which a model with repair rates, as in Chapter 16, provides. The same model would replace the pinned MDT = 12 h by a derived quantity, and would change the year's loss of Table 25.6 but not the risk, by equation (25.31).

Consequences that depend on the state of the load. A loss at full load costs more than a loss at night, and if the units also fail more often under full load, the cost and the probability of a loss are correlated across the year. The risk is then a sum over periods of the period's cost rate times the period's probability, not the average cost times the average probability; Exercise 25.5 shows that the difference can be a fifth of the risk. The extension to consequences that are functions of the state of the plant is made in the later chapter on repair and replacement decisions.

Independence given the unavailabilities. Every configuration probability of this chapter is a product over units. A common cause that fails two pumps at once, the subject of Chapter 15, adds a cut set of order one and raises the loss-state term by its probability times $C_{\mathrm{L}} t_{\mathrm{yr}}$. Had the pumps' cut set the pair unavailability of that chapter, 2.14×10^{-4} with the common cause, in place of $\mathbb{E}[q_P^2 \mid \mathbf{y}] = 1.25 \times 10^{-4}$, the pumps' rare-event term would rise from R$ 21 845 to R$ 37 493 per year, still a small part of the risk.

The level and the prior. With three, four and one units found down, the level of the expected risk depends on the prior by up to 60 % (Section 25.4.3), and every prior tried moves it up; the attribution and the design comparison do not move. A level that must be quoted in a signed statement needs a prior argued from the history of similar units, or longer records, the CRAC record first.

Common mistake

Taking the robustness of the attribution to the consequence schedule as robustness of the acceptance: across the schedules of Table 25.10 the attribution does not move and the acceptance probability moves from 0.28 to 0.82 (0.33 to 0.87 exactly).

The next chapter turns the gradients and shares that this chapter and the two before it have used into the importance measures of the basic events, Birnbaum, Fussell–Vesely and the risk-reduction and risk-achievement measures, each derived as a derivative or a ratio of the top event or of the risk, with its band.

25.7 What to retain

Definition 25.1 (Risk as an expected loss over enumerated configurations)**.** For units with unavailabilities $\mathbf{q}$, independent given $\mathbf{q}$, and a consequence rate $C(\mathbf{s})$ for each configuration $\mathbf{s}$, the risk is $\mathcal{R}(\mathbf{q}) = \sum_{\mathbf{s}} \Pr(\mathbf{s} \mid \mathbf{q})\, C(\mathbf{s})$ per hour and $\mathcal{R}_{\mathrm{yr}} = t_{\mathrm{yr}}\mathcal{R}$ per year, the expected loss of a year given $\mathbf{q}$. It splits exactly by class of configuration, $\mathcal{R} = C_{\mathrm{L}} U_{\mathrm{sys}} + \mathcal{R}_{\mathrm{deg}}$, and its loss-state term is bounded by the cut-set sum $C_{\mathrm{L}} \sum_c P_c$. Its expectation is $\sum_{\mathbf{s}} C(\mathbf{s}) \mathbb{E}[\Pr(\mathbf{s} \mid \mathbf{q}) \mid \mathbf{y}]$, and it exceeds the plug-in value by about $\frac{1}{2}\operatorname{tr}(\boldsymbol{\Sigma}_{\mathbf{q}}\mathbf{H}_{\mathbf{q}})$; with the unavailabilities carried as logits, its Laplace band on the logarithm is $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ with $\mathbf{g} = \mathbf{D}\,(\partial\mathcal{R}/\partial\mathbf{q})/\mathcal{R}$. Risk-informed acceptance at limit $\mathcal{R}^*$ with assurance κ requires $\Pr(\mathcal{R}_{\mathrm{yr}} < \mathcal{R}^* \mid \mathbf{y}) \geq \kappa$. The loss of one year has mean $\mathcal{R}_{\mathrm{yr}}(\mathbf{q})$ given $\mathbf{q}$ and predictive variance $\mathbb{E}[\operatorname{Var}(L_{\mathrm{yr}} \mid \mathbf{q})] + \operatorname{Var}(\mathcal{R}_{\mathrm{yr}})$.

Three conclusions carry beyond the hall. The risk is a derived quantity of the state, linear in the consequences and polynomial in the unavailabilities; under a shared posterior its expectation exceeds the plug-in value when the dominant losses need coincidences, by half the trace of the posterior covariance of the unavailabilities times the risk's curvature, and its band, its decomposition and its acceptance probability come from the same posterior as any other derived quantity. Acceptance is a tail probability of the risk, equivalent to comparing the limit with a quantile of the risk's posterior, and it is sensitive to the consequence schedule and to the prior while the attribution of the risk to cut sets and the ranking of design changes are not. And the posterior of the risk and the loss of one year are two distributions: the first is epistemic, the second adds the aleatory count of losses, and a bad year is a tail of the predictive, which for this hall is dominated by the uncertainty of the rate. The record worth buying for the acceptance question is the one whose added information the risk's gradient weights most, here the CRAC checks.

Table 25.13 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 25.13: The terms of risk and expected consequences and where each is defined.

Term	Expression	Units	Meaning	Defined in
Consequence rate	$C(\mathbf{s})$, equation (25.3)	R$/h	cost per hour in $\mathbf{s}$	25.1.3
Risk	$\mathcal{R} = \sum_{\mathbf{s}} \Pr(\mathbf{s} \mid \mathbf{q}) C(\mathbf{s})$	R$/h	expected cost rate	25.3.1
Annual risk	$\mathcal{R}_{\mathrm{yr}} = t_{\mathrm{yr}} \mathcal{R} = \mathbb{E}[L_{\mathrm{yr}} \mid \mathbf{q}]$, $t_{\mathrm{yr}} = 8760\,\mathrm{h}$	R$/yr	expected loss of a year	25.3.1
Loss of an action	$L(a, \mathbf{z}) = c_a + \mathcal{R}_{\mathrm{yr}}(\mathbf{q}(\mathbf{z}); a)$, $\mathcal{R}(a) = \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$	R$/yr	the general form for the hall	25.3.1
Split by class	$C_{\mathrm{L}} U_{\mathrm{sys}} + \mathcal{R}_{\mathrm{deg}}$	R$/h	loss and degraded terms	25.3.1
Expected risk	$\sum_{\mathbf{s}} C(\mathbf{s}) \mathbb{E}[\Pr(\mathbf{s} \mid \mathbf{q}) \mid \mathbf{y}]$	R$/h	products of moments	25.3.1
Plug-in gap	$\frac{1}{2} \operatorname{tr}(\boldsymbol{\Sigma}_{\mathbf{q}} \mathbf{H}_{\mathbf{q}})$	R$/h	sign of the curvature	25.3.1
Laplace band of the risk	$\mathbf{g} = \mathbf{D}(\partial \mathcal{R} / \partial \mathbf{q}) / \mathcal{R}$, $\mathbf{g}^{\mathsf{T}} \boldsymbol{\Sigma} \mathbf{g}$	–	variance budget of $\ln \mathcal{R}$	25.3.1
Cut-set bound	$C_{\mathrm{L}} U_{\mathrm{sys}} \le C_{\mathrm{L}} \sum_c P_c$	R$/h	rare-event attribution	25.3.2
Acceptance probability	$\Pr(\mathcal{R}_{\mathrm{yr}} < \mathcal{R}^* \mid \mathbf{y}) \ge \kappa$	–	tail of the risk	25.3.3
Loss frequency	$\nu_{\mathrm{loss}} = t_{\mathrm{yr}} U_{\mathrm{sys}} / \mathrm{MDT}$	yr^{-1}	losses of cooling per year	25.3.4
Loss of a year	$C_{\mathrm{L}} \mathrm{MDT}\, K_f + \mathcal{R}_{\mathrm{deg,yr}}$	R$/yr	aleatory around $\mathcal{R}_{\mathrm{yr}}$	25.3.4
Variance split	$\mathbb{E}[\mathrm{Var}(L_{\mathrm{yr}} \mid \mathbf{q})] + \mathrm{Var}(\mathcal{R}_{\mathrm{yr}})$	$(\mathrm{R\$/yr})^2$	aleatory plus epistemic	25.3.4
Risk curve	$\bar{F}_X(x) = \Pr(X > x \mid \mathbf{y})$	–	exceedance probability	25.3.4
Value at risk, tail mean	VaR_p, $\mathrm{CVaR}_p = (1-p)^{-1} \int_p^1 \mathrm{VaR}_u \, \mathrm{d}u$	R$/yr	quantile and tail average	25.3.4
Risk reduction	$\Delta \mathcal{R}_{\mathrm{yr}}(\mathbf{q}; a)$, draw by draw	R$/yr	value of a design change	25.3.5
Preposterior band	$w_a \propto \sqrt{\mathbf{g}^{\mathsf{T}} (\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}} \mathbf{J}_a)^{-1} \mathbf{g}}$	decades	width after a record	25.6.2

25.8 Exercises

Exercise 25.1. *Formulation.* (a) A pumping station has three pumps of which two must run, each with a known unavailability $q = 0.02$, independent. A loss of the station costs R$ 10 000 h^{-1} and a single pump down with the station running costs R$ 100 h^{-1}. Write the risk as a sum over the configurations and compute it per hour and per year. (b) The posterior of a plant's annual risk is lognormal with median R$ 0.8 million and a standard deviation of 0.9 on the natural logarithm. Compute the probability that the annual risk is below R$ 1.5 million, its mean, and the limit that would be met with an assurance of 0.9. ((a) $\mathcal{R} = 10\,000\,[3q^2(1-q) + q^3] + 100 \times 3q(1-q)^2 = 10\,000 \times 0.001184 + 100 \times 0.057624 = \mathrm{R\$}\,17.60\,\mathrm{h}^{-1}$, or R$ 154 197 per year, two thirds of it in the loss states. (b) $\Phi\big(\ln(1.5/0.8)/0.9\big) = \Phi(0.698) = 0.758$; mean $0.8\,\mathrm{e}^{0.9^2/2} = \mathrm{R\$}\,1.20$ million, below the limit although the acceptance probability is only 0.76; the limit met with assurance 0.9 is $0.8\,\mathrm{e}^{1.2816 \times 0.9} = \mathrm{R\$}\,2.54$ million.)

Exercise 25.2. *Derivation.* Two units in parallel share one unavailability q, with posterior mean $\bar{q}$ and variance $\mathrm{Var}(q \mid \mathbf{y})$. Losing both costs C_{L} per hour and one unit down with the service kept costs C_1 per hour. (a) Write $\mathcal{R}(q)$ and derive $\mathbb{E}[\mathcal{R} \mid \mathbf{y}] - \mathcal{R}(\bar{q})$ exactly, and compare it with $\frac{1}{2} \operatorname{tr}(\boldsymbol{\Sigma}_{\mathbf{q}} \mathbf{H}_{\mathbf{q}})$. (b) State when the gap is positive. (c) Evaluate it for Beta$(3, 97)$,

$C_{\mathrm{L}} = \mathrm{R\$}\,20\,000$ and $C_1 = \mathrm{R\$}\,150\,\mathrm{h}^{-1}$. ((a) $\mathcal{R}(q) = C_{\mathrm{L}}q^2 + 2C_1q(1-q) = 2C_1q + (C_{\mathrm{L}} - 2C_1)q^2$; the linear term has no gap, so $\mathbb{E}[\mathcal{R}] - \mathcal{R}(\bar{q}) = (C_{\mathrm{L}} - 2C_1)\,\mathrm{Var}(q \mid \mathbf{y})$, which is $\frac{1}{2}\mathcal{R}''\,\mathrm{Var}(q \mid \mathbf{y})$, the one-entry trace, exact because $\mathcal{R}$ is quadratic. (b) When a loss costs more than two units' degraded rates, $C_{\mathrm{L}} > 2C_1$. (c) $19\,700 \times 2.881 \times 10^{-4} = \mathrm{R\$}\,5.68\,\mathrm{h}^{-1}$, R\$ 49 721 per year, against $\mathcal{R}(0.03) = \mathrm{R\$}\,26.73\,\mathrm{h}^{-1}$.)

Exercise 25.3. *Derivation.* For the pair of the previous exercise, carry the unavailability as a one-entry state, the number ξ, its logit. (a) Derive the Laplace band of $\ln\mathcal{R}$ at the MAP state $\bar{q} = 0.03$, with the Laplace variance of the logit of $\mathrm{Beta}(3, 97)$, $\sigma_\xi^2 = 1/3 + 1/97$; (b) derive the exact band, using that $\mathcal{R}$ increases with q; (c) compare them. ((a) The slope is $g = \mathrm{d}\ln\mathcal{R}/\mathrm{d}\xi = \mathcal{R}'(q)\,q(1-q)/\mathcal{R}$ with $\mathcal{R}' = 2C_{\mathrm{L}}q + 2C_1(1-2q) = 1482$, so $g = 1482 \times 0.0291/26.73 = 1.613$; $\sigma_\xi = (1/3 + 1/97)^{1/2} = 0.586$ and $\mathrm{sd}(\ln\mathcal{R} \mid \mathbf{y}) \approx g\,\sigma_\xi = 0.946$; band $26.73\,\mathrm{e}^{\pm 1.854} = [4.19,\ 170.6]\,\mathrm{h}^{-1}$. (b) The quantiles map through $\mathcal{R}$: $[\mathcal{R}(0.0063),\ \mathcal{R}(0.0711)] = [2.67,\ 120.8]\,\mathrm{h}^{-1}$. (c) The Laplace band is too high, by a factor of 1.57 at the lower end and 1.41 at the upper. Centering on $\mathcal{R}(\bar{q}) = 26.7$ where the median is $\mathcal{R}(0.0269) = 22.4$ explains a factor of 1.20; the rest is the asymmetry of the exact band, which reaches 2.13 below the median in logarithm and 1.69 above it, where the Laplace band is symmetric. The widths in logarithm, 3.7 and 3.8, are close.)

Exercise 25.4. *Derivation.* Losses of a service occur as a Poisson process of frequency ν per year, each of mean duration MDT, at a cost C_{L} per hour. (a) Derive the expected loss of a year and show that it equals $C_{\mathrm{L}}t_{\mathrm{yr}}U$ with U the unavailability. (b) Derive its variance for fixed durations and for exponential durations. (c) Evaluate both for the hall at the plug-in values, $\nu = 4.585$ per year and $\mathrm{MDT} = 12\,\mathrm{h}$. ((a) $\mathbb{E}[L] = C_{\mathrm{L}}\,\mathrm{MDT}\,\mathbb{E}[K] = C_{\mathrm{L}}\,\mathrm{MDT}\,\nu\tau$, and $U = \nu\,\mathrm{MDT}/t_{\mathrm{yr}}$ gives $C_{\mathrm{L}}t_{\mathrm{yr}}U\tau$. (b) Fixed: $\mathrm{Var} = (C_{\mathrm{L}}\mathrm{MDT})^2\nu\tau$; exponential: $\mathrm{Var} = \nu\tau\,C_{\mathrm{L}}^2\mathbb{E}[D^2] = 2(C_{\mathrm{L}}\mathrm{MDT})^2\nu\tau$. (c) Mean $240\,000 \times 4.585 = \mathrm{R\$}\,1.100$ million; standard deviations R\$ 0.514 and R\$ 0.727 million.)

Exercise 25.5. *Application.* The cost of a lost hall is $\mathrm{R\$}\,30\,000\,\mathrm{h}^{-1}$ in the four summer months, 2920 h, and $\mathrm{R\$}\,15\,000\,\mathrm{h}^{-1}$ in the other 5840 h. (a) With the posterior of the hall unchanged through the year, compute the loss-state term by season and in total. (b) In summer the CRAC units work harder, and their unavailability is twice the year's, $2q_C$, with the same posterior draw; compute the summer term with $\mathbb{E}[U_{\mathrm{sys}}] = 2.605 \times 10^{-2}$ for summer, and compare the total with the one obtained from the average cost $\mathrm{R\$}\,20\,000\,\mathrm{h}^{-1}$. ((a) Each season contributes R\$ 682 406, since $30\,000 \times 2920 = 15\,000 \times 5840$; the total, R\$ 1 364 812, is that of the year-round $\mathrm{R\$}\,20\,000\,\mathrm{h}^{-1}$. (b) Summer $30\,000 \times 2.605 \times 10^{-2} \times 2920 = \mathrm{R\$}\,2.28$ million, total R\$ 2.96 million; with the average cost the total would be $20\,000 \times (2.605 \times 10^{-2} \times 2920 + 7.790 \times 10^{-3} \times 5840) = \mathrm{R\$}\,2.43$ million. The difference, R\$ 0.53 million, is the covariance of cost and probability across the seasons, which an average cost hides.)

Exercise 25.6. *Application.* The cost of an hour without cooling is uncertain: a single value for the year, lognormal with median R\$ 20 000 and a standard deviation of 0.5 on its natural logarithm, independent of the units. (a) Derive the expected annual risk. (b) State how the band and the acceptance probability at R\$ 1.5 million change. (c) State which conclusions of the chapter are unchanged. ((a) By independence, $\mathbb{E}[\mathcal{R}_{\mathrm{yr}}] = \mathbb{E}[\mathrm{factor}] \times 1\,364\,812 + 148\,273$ with $\mathbb{E}[\mathrm{factor}] = \mathrm{e}^{0.5^2/2} = 1.133$: R\$ 1.69 million. (b) From 2×10^5 draws, the standard deviation of $\ln\mathcal{R}_{\mathrm{yr}}$ rises from 0.85 to 0.95, the band widens to [R\$ 0.17, R\$ 6.91] million, the median hardly moves, R\$ 1.09 million, and the acceptance probability falls from 0.646 to 0.631. (c) The attribution to cut sets, the ranking of the designs and the factor by which each design scales the loss-state term, since the uncertain factor multiplies the loss-state term of every configuration alike.)

Chapter 26

Reliability Importance Measures

26.1 The importance of each unit to the loss of cooling

The owner of the data hall of Chapters 14 and 25 has one maintenance budget and three requests on it: a closer watch on the units whose failures the losses of cooling run through, an improvement program for one group of units, and a rule for which unit may never be taken out of service while the hall runs. The facilities engineer is asked to rank the seven units. The records are those of Chapter 14: of a hundred checks of the four CRAC units, three found a unit down; of four hundred checks of the two chilled-water pumps, four; of a thousand checks of the control system, one. Chapter 14 left a table of the slopes of the probability of losing cooling in each unit's unavailability, at the fractions found down, and the first reading takes it as the ranking: the controller's slope is 0.995, twelve times a CRAC unit's 0.085 and a hundred times a pump's 0.0099, so the controller is the hall's most important unit, to be watched most closely, improved first and never let fail.

The engineer who must defend the ranking asks what "most important" means in each of the three requests. A slope says how far the loss of cooling moves per unit of a unit's unavailability. It does not say how much of the loss of cooling actually passes through the unit, which is what a watch is for; nor how much worse the hall is with the unit certainly down, which is what an outage rule is for; nor whether the order would survive what the records leave open, when the slopes were read at fractions from three, four and one units found down. The question of this chapter is the owner's, stated in full:

which unit matters most to the loss of cooling, in each sense in which the word is used (the slope of the top event, the share of it that runs through the unit, the top event with the unit pinned down and pinned up), with the band that the evidence leaves on each measure and the posterior probability that each ranking holds.

What changes from Chapters 14 and 25 is the derived quantity: the state, its rows and its posterior are those of Chapter 14, and the new quantities are the top event evaluated with one unit pinned, and ratios and differences of those evaluations.

Jargon: importance measure, Birnbaum importance, critical unit, Fussell–Vesely importance, criticality importance, risk achievement worth (RAW), risk reduction worth (RRW), pinned down, pinned up, pivotal decomposition, uncertainty importance

An *importance measure* ranks the basic events of a system by their effect on the probability of the top event. The *Birnbaum importance* of a unit is the derivative of the top event's probability with respect to that unit's unavailability; it equals the probability that the other units leave the system in a configuration in which this unit decides the outcome, the unit then being *critical*. The *Fussell–Vesely importance* is the fraction of the top event's probability that the unit's failure accounts for; the *criticality importance* is the probability that the unit is down and critical, given that the top event has occurred. The *risk achievement worth* (RAW) is the ratio of the top event's probability with the unit certainly down to its present value; the *risk reduction worth* (RRW) is the ratio of the present value to the probability with the unit never failing. A unit is *pinned down* when its unavailability is set to one, *pinned up* when it is set to zero. The *pivotal decomposition* writes the top event's probability as the average of its two pinned values. In the names RAW and RRW, "risk" is the top event's probability, the standard usage of the measures, not the expected loss in reais of Chapter 25. The *uncertainty importance* of a unit ranks it by the part of the top event's band that its record leaves.

Key idea

Every importance measure of a unit is built from three numbers: the top event at the present unavailabilities and with the unit pinned down and pinned up. The Birnbaum importance is their difference, the gradient of the top event in the unit's unavailability; the Fussell–Vesely importance is the share of the top event that pinning the unit up removes, equal here to the criticality importance; the two worths are ratios of the same numbers, and two identities tie all four to the Birnbaum importance and the top event. Each measure is a derived quantity of the state with a band, and each ranking is an event with a posterior probability, to be reported beside the order. The gradient of the logarithm of the top event in the state is the Fussell–Vesely vector summed over the units that share an entry and scaled by one minus the entry's unavailability, so the band of the top event and the importance measures are one computation; the budget of that band ranks the records, which none of the units' measures does.

26.1.1 Three evaluations of one top event

A ranking of units needs a number per unit, and the number depends on what the ranking is for. The companion volume *Reliability Engineering* develops the importance measures in its chapter on importance measures, with their names, their definitions for a fault tree, their reading for maintenance and their limits, and those names are used here. The measures share one construction. Hold every unavailability at its present value and evaluate the top event three times: as it is, U_{sys}; with unit i certainly down, U_i^+; and with unit i never down, U_i^-. A difference of the pinned values is a slope, a ratio of the present value to a pinned value is a factor, and a difference over the present value is a share. For a small tree the three evaluations are three passes of the enumeration of Chapter 13, and every measure of every unit costs two passes per unit beyond the one already made.

Because the unavailabilities are known only through their records, the three evaluations are functions of the state, and so is every measure built from them: each measure has a posterior, and the order of two units under a measure is an event of that posterior. A ranking read at the plug-in values reports one draw of that event as if it were the event's probability.

26.1.2 The physical system and its graph (step 0)

The plant is that of Chapter 14, drawn in Figure 14.1: four CRAC units C1 to C4, of which three must run; two chilled-water pumps P1 and P2, of which one must run; and one control system K, which must run. Units of one design checked under one record share one entry, so the state holds the three logits of equation (14.2),

$$\mathbf{z} = \begin{bmatrix} \xi_C \\ \xi_P \\ \xi_K \end{bmatrix}, \qquad q_j = \frac{1}{1 + \mathrm{e}^{-\xi_j}}, \quad j \in \{C, P, K\}, \tag{26.1}$$

while the enumeration and the measures run over the seven units. The top event is that of the fault tree of Figure 14.2, and its probability given the unavailabilities is equation (14.26),

$$U_{\mathrm{sys}}(\mathbf{q}) = 1 - R_{\mathrm{C}}(q_C)\,(1 - q_P^2)\,(1 - q_K), \qquad R_{\mathrm{C}}(q) = (1-q)^3(1+3q), \tag{26.2}$$

with R_{C} the probability that at most one CRAC unit is down. In this chapter $\mathbf{q} = (q_1, \ldots, q_7)^{\mathsf{T}}$ lists the units' unavailabilities in the order C1 to C4, P1, P2, K, and the units of a group take the value of their group's entry; a measure is computed for one unit with its own entry of $\mathbf{q}$ changed and the others held.

Figure 26.1 draws the tree of Figure 14.2 again with the questions of this chapter attached. The box at the top is the top event; under it the OR gate combines the voting gate 2/4 on the four CRAC units, the AND gate on the two pumps and the controller's basic event. Under each group a gray box names the three evaluations every measure of that group's units needs: the top event with one unit pinned down ($q_i = 1$) and pinned up ($q_i = 0$), beside the present value. The graph itself is not extended. A pinned unit is an evaluation of the derived quantity at a changed argument, not a row and not an entry, and the plant graph of Chapter 14 serves unchanged.

26.1.3 The probabilistic questions (step 1)

The owner asks five questions of the tree and the three records, and step 1 writes each as a quantity over the state, with the unknown on the left of the bar and the records on the right. Table 26.1 lists them: the first column is the question in engineering language; the second the general form of its category in the notation of Chapter 2, for a quantity $f(\mathbf{z})$ derived from the state and a level f^* of it; the third its mathematical form for the hall; the fourth its category; and the last the subsection where it is answered. The first three are estimation of derived quantities, one per family of measures; the slope of the first is a gradient, the quantity Chapters 13 and 14 named and left for this one. The fourth asks whether the orders those measures give are firm, a prediction under a threshold on the difference of two units' measures. The fifth is estimation again, of the band of the top event, and it asks why that band is built from the same numbers. Which record to extend, for the order of the units or for that band, is the next measurement of step 7, not a question of this table. The atlas of Appendix A lists the same forms.

26.1.4 The evidence as rows (step 2)

Step 2 attaches the rows of Chapter 14: each check record is one count row on the logit of its group, the hundred CRAC checks on ξ_C, the four hundred pump checks on ξ_P and the thousand controller checks on ξ_K, in the working form below. Which row acts on which entry is a property of the graph, and it fixes where the Jacobian of the rows can be nonzero, marked $*$:

$$\mathbf{J} = \begin{bmatrix} * & 0 & 0 \\ 0 & * & 0 \\ 0 & 0 & * \end{bmatrix} \begin{matrix} \text{CRAC checks} \\ \text{pump checks} \\ \text{controller checks} \end{matrix}, \qquad \text{columns } \xi_C, \xi_P, \xi_K, \tag{26.3}$$

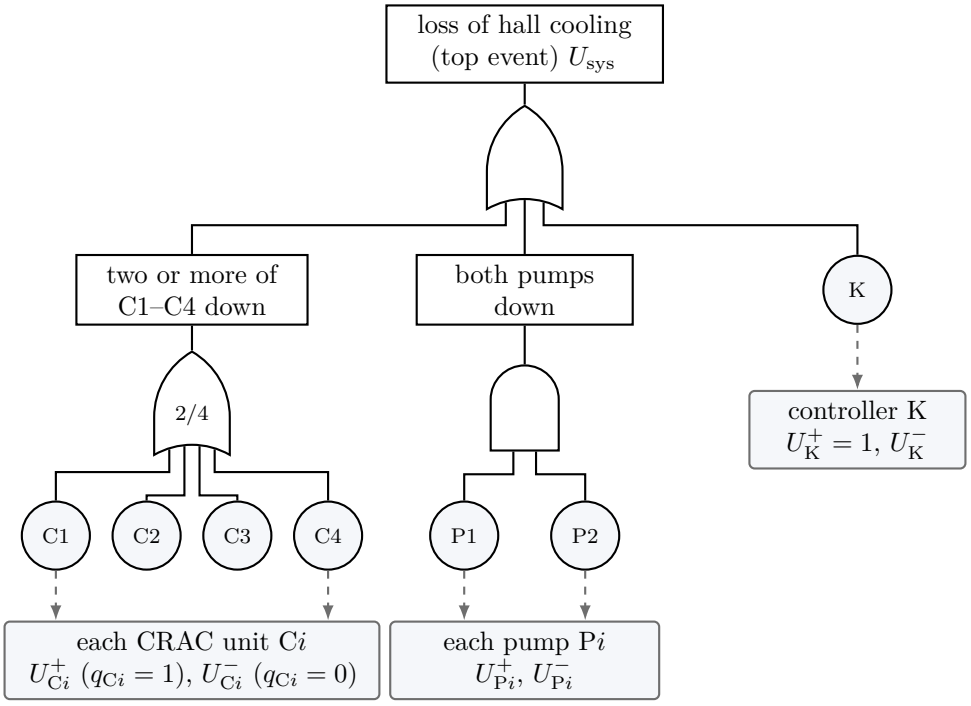


Figure 26.1: Fault tree of the loss of hall cooling, as in Figure 14.2: top event (≥ 2 of C1–C4 down) $\vee$ (P1 $\wedge$ P2) $\vee$ K. Gray boxes: the evaluations of the top event that the importance measures of each group's units need, with one unit pinned down (U_i^{+}, $q_i = 1$) and pinned up (U_i^{-}, $q_i = 0$) and the others at their present unavailabilities, the units of a group sharing q_C or q_P. The controller's own cut set makes $U_{\text{K}}^{+} = 1$.

Table 26.1: The questions asked of the importance of the hall's units, the general form of their category, their mathematical form for the hall, their category and where each is answered. $\mathbf{y}$ is the three check records; U_i^{+} and U_i^{-} the top event with unit i pinned down and pinned up; $I_{\text{B},i}$, $I_{\text{FV},i}$, RAW_i and RRW_i the four measures of unit i; $\mathbf{g}$ the gradient of the top event in the state.

Engineering question	General form	Mathematical form	Category	Answered in
Which unit's unavailability matters most to the loss of cooling?	$\Pr(f(\mathbf{z}) \leq f^{*} \mid \mathbf{y})$, location and quantiles	$I_{\text{B},i} = \partial U_{\text{sys}}/\partial q_i = U_i^{+} - U_i^{-}$, with bands (Birnbaum)	estimation of a derived quantity (a gradient)	26.3.1
What share of the loss of cooling runs through each unit?	$\Pr(f(\mathbf{z}) \leq f^{*} \mid \mathbf{y})$, location and quantiles	$I_{\text{FV},i} = (U_{\text{sys}} - U_i^{-})/U_{\text{sys}}$, with bands (Fussell–Vesely)	estimation of a derived quantity	26.3.2
How much worse if the unit were certainly down, how much better if it never failed?	$\Pr(f(\mathbf{z}) \leq f^{*} \mid \mathbf{y})$, location and quantiles	$\text{RAW}_i = U_i^{+}/U_{\text{sys}}$, $\text{RRW}_i = U_{\text{sys}}/U_i^{-}$, with bands	estimation of derived quantities	26.3.3
Are the rankings firm?	$\Pr(f(\mathbf{z}) \geq f^{*} \mid \mathbf{y})$	$\Pr(I_{\text{K}} > I_{\text{C1}} \mid \mathbf{y})$ for each measure I, and the other pairs	prediction under a threshold	26.3.4
Why does the band of the loss of cooling use the same numbers?	$\Pr(f(\mathbf{z}) \leq f^{*} \mid \mathbf{y})$, central interval	central 95 % band of $\ln U_{\text{sys}}$; its Laplace spread $\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}/U_{\text{sys}}$ with $\mathbf{g}$ built from the $I_{\text{B},i}$	estimation of a derived quantity (its band)	26.3.5

the pattern of equation (14.3). No row acts on two entries and no row is a physics row, so no model tolerance σ_p arises. The questions add no rows: the pinned evaluations and every measure are derived quantities of the state through equation (26.2), and the shared columns of the pattern are what makes the measures of the four CRAC units, and of the two pumps, move together under the posterior.

The modeling choices. Two choices are made for the whole hall, and each buys something and can fail. The units are taken as down independently given their unavailabilities: this is what makes the top event multilinear in the units' unavailabilities and the pivotal decomposition exact, and it fails under a cause that takes two units down at once, which pins the second unit when the first is pinned (Section 26.6.5). A unit whose group shares one entry is given back its own unavailability, set to the shared value, and the measure is computed with that one value changed: this lets the measures speak of a unit, which is what a watch or an outage rule acts on, and it is exact only if the units of a group are alike, which the group totals of the records cannot test (Section 26.6.5).

The count rows. The rows are those of Chapter 14, equation (14.5): a record alone is a flat prior on the logit, and each record is one count row in the working form of Chapter 4, equation (4.43), the binomial pair in the logit, with k_j units found down in n_j checks,

$$r_j(\mathbf{z}) = \frac{n_j q_j - k_j}{\sqrt{n_j q_j (1-q_j)}}, \qquad \mathbf{J}(\mathbf{z}) = \operatorname{diag}\Big(\sqrt{n_j q_j (1-q_j)}\Big), \qquad j \in \{C, P, K\}, \tag{26.4}$$

three rows on three entries with the pattern (26.3).

26.2 The posterior of the state (step 3)

The five questions of Table 26.1 rest on the posterior of Chapter 14, the posterior of the state that the three count rows of step 2 give: the MAP state $\hat{\mathbf{z}}$, the information matrix $\mathbf{\Lambda} = \mathbf{J}^\mathsf{T}\mathbf{J}$ and the covariance $\mathbf{\Sigma} = \mathbf{\Lambda}^{-1}$, with the gradient $\mathbf{g}$ of a derived quantity. Its exact form is three independent Beta posteriors, one per record. This section derives the posterior and computes it for the hall; Section 26.3 computes every answer from it.

26.2.1 The posterior from the three rows

The logit is the canonical parameter of the binomial row, so the curvature term of the count rows vanishes and the curvature of the objective is $\mathbf{J}^\mathsf{T}\mathbf{J}$ exactly; with the rows normalized, as here, the information matrix is $\mathbf{\Lambda} = \mathbf{J}^\mathsf{T}\mathbf{J}$. The normal equations, equation (14.6), are three separate equations, and their solution and the Laplace covariance are those of equation (14.7), beside the exact posterior of equation (14.8):

$$\hat{z}_j = \ln \frac{k_j}{n_j - k_j}, \qquad \mathbf{\Sigma} = \operatorname{diag}\Big(\frac{1}{k_j} + \frac{1}{n_j - k_j}\Big), \qquad q_j \mid \mathbf{y} \sim \operatorname{Beta}(k_j,\ n_j - k_j), \tag{26.5}$$

independent across the groups, a result of the pattern and not an assumption, with $\mathbb{E}[q_j \mid \mathbf{y}] = k_j/n_j = q(\hat{z}_j)$: the plug-in values of every measure are its values at the MAP state and at the posterior means. The exact variance of a logit is the trigamma sum of equation (14.9), the counterpart of Σ_{jj}.

26.2.2 The posterior of the hall

The hall's inputs are the three check records, three CRAC units found down in 100 checks, four pumps in 400 and one controller in 1000. The fit is that of Chapter 14, whose Table 14.2 runs the normal equations from $\mathbf{z}_0 = (-4,\ -4,\ -4)^{\mathsf{T}}$ to the MAP state in six iterations. By equation (14.10),

$$\hat{\mathbf{z}} = \begin{bmatrix} -3.4761 \\ -4.5951 \\ -6.9068 \end{bmatrix}, \qquad \boldsymbol{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J} = \begin{bmatrix} 2.9100 & 0 & 0 \\ 0 & 3.9600 & 0 \\ 0 & 0 & 0.9990 \end{bmatrix},$$

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda}^{-1} = \begin{bmatrix} 0.3436 & 0 & 0 \\ 0 & 0.2525 & 0 \\ 0 & 0 & 1.0010 \end{bmatrix},$$

and the exact posteriors are $q_C \sim \text{Beta}(3,\ 97)$, $q_P \sim \text{Beta}(4,\ 396)$ and $q_K \sim \text{Beta}(1,\ 999)$, Table 14.3.

The pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ is what step 4 uses: every answer of the chapter is computed from it. Because this problem also admits the exact posterior, each answer has an exact value beside it, from the draws and the quadrature, and the gap between the two is reported with it.

26.3 The answers (step 4)

Every measure is a rational function of the three unavailabilities, so at the MAP state it is a number, and its band and the probability of every ranking follow from the gradient of its logarithm and the covariance of Section 26.2. Each question of Table 26.1 is answered from that posterior: its engineering question, the general form of its category, the form it takes for the hall, the quantities of the posterior it uses, and the hall's numbers. The questions use the tree of Figure 26.1 and, for the exact values, 10^5 triples (q_C, q_P, q_K) drawn from the three Beta posteriors with the generator numpy `default_rng(26)`, each passed through the closed form (26.2) and through the measures as each question defines them. The plug-in values are at the MAP state, which maps to the posterior means $\bar{q}_C = 0.030$, $\bar{q}_P = 0.010$ and $\bar{q}_K = 0.001$, where the top event is $U_{\text{sys}} = 6.2806 \times 10^{-3}$, the value of Table 14.11. Every answer is computed from the MAP state $\hat{\mathbf{z}}$ and the covariance $\boldsymbol{\Sigma}$ of step 3, and that is the value the chapter reports; the draws and the quadrature of the exact posterior give the value set beside it, with the gap.

26.3.1 The Birnbaum importance

Engineering question. Which unit's unavailability matters most to the loss of cooling?

General form. The estimation form of Chapter 2, equations (2.9) to (2.11), for the derived quantity $f(\mathbf{z}) = \partial U_{\text{sys}}/\partial q_i$: a function of the state like any other, summarized by a location and a central 95 % interval,

$$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) \le f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}.$$

Specialized form. The question is equations (2.9) to (2.11) with the state $\mathbf{z}$ the three logits, $f(\mathbf{z}) = I_{\text{B},i}$, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 26.2.2. No new choice enters.

The pivotal decomposition. Conditioning on the status of unit i by the law of total probability, as Chapter 13 did for the service availability, equation (13.23), the top event is the average of

its two pinned values,

$$U_{\text{sys}}(\mathbf{q}) = q_i\, U_i^{+} + (1 - q_i)\, U_i^{-}, \qquad U_i^{+} = U_{\text{sys}}(\mathbf{q} \mid q_i = 1), \quad U_i^{-} = U_{\text{sys}}(\mathbf{q} \mid q_i = 0), \tag{26.6}$$

and neither pinned value depends on q_i. The top event is therefore a straight line in q_i, and its slope is the difference of the two ends,

$$I_{\text{B},i} = \frac{\partial U_{\text{sys}}}{\partial q_i} = U_i^{+} - U_i^{-}, \tag{26.7}$$

the same at every value of q_i: the Birnbaum importance of unit i, equation (13.24) written for the top event, here derived for any coherent tree of units independent given $\mathbf{q}$ rather than for one structure. The difference $U_i^{+} - U_i^{-}$ is the probability of the configurations of the other units in which unit i decides the outcome, the hall failing with i down and cooled with i up: by coherence $U_i^{+} \geq U_i^{-}$, and the configurations in which both pinned values agree cancel. A unit's Birnbaum importance is thus the probability that it is critical, and it is a property of the other units, not of the unit.

For the hall the three kinds of unit have closed forms. Pinning C1 down leaves the CRAC factor failed if any of the other three is down; pinning it up, if two or more of them are; pinning a pump down leaves the pump factor $1 - q_P$, pinning it up makes it one; pinning the controller down fails the hall, pinning it up removes its factor. Subtracting the two pinned forms of equation (26.2),

$$\begin{aligned} I_{\text{B,C}} &= 3q_C(1 - q_C)^2\,(1 - q_P^2)\,(1 - q_K), \\ I_{\text{B,P}} &= q_P\, R_{\text{C}}(q_C)\,(1 - q_K), \\ I_{\text{B,K}} &= R_{\text{C}}(q_C)\,(1 - q_P^2), \end{aligned} \tag{26.8}$$

each the probability that the rest of the hall leaves that unit critical: a CRAC unit when exactly one other CRAC unit is down and the pumps and the controller hold; a pump when the other pump is down and the CRAC units and the controller hold; the controller whenever the rest of the hall would cool it, which is nearly always.

The vector and the shared entries. The seven Birnbaum importances form the gradient $\mathbf{I}_{\text{B}} = \nabla_{\mathbf{q}} U_{\text{sys}}$, one entry per unit, which equation (14.43) wrote for the same tree. The state carries one logit per group, so the gradient of the top event in the state collects the units that share an entry through the incidence matrix $\mathbf{A}$ of equation (14.30) and converts each unavailability to its logit through $\mathbf{D} = \text{diag}(q_j(1 - q_j))$:

$$\mathbf{A}^{\mathsf{T}}\mathbf{I}_{\text{B}} = \begin{bmatrix} 4\, I_{\text{B,C}} \\ 2\, I_{\text{B,P}} \\ I_{\text{B,K}} \end{bmatrix}, \qquad \mathbf{g} = \frac{\partial U_{\text{sys}}}{\partial \mathbf{z}} = \mathbf{D}\,\mathbf{A}^{\mathsf{T}}\mathbf{I}_{\text{B}}, \tag{26.9}$$

equation (14.44). The entry $4I_{\text{B,C}}$ is the slope of the top event in the shared q_C, all four units moving together; the entry $I_{\text{B,C}}$ is the slope in one unit's unavailability, the others held. Which of the two is meant depends on whether the change acts on one unit or on the design.

The band. The importance is a derived quantity of the state, so its value is that of equation (26.8) at $\hat{\mathbf{z}}$, the median under the Laplace normal, and its band that of the delta method on its logarithm, equation (14.29), with its gradient in the logits; the exact band beside it comes from the three Beta posteriors, by the draws. The importance of a unit does not depend on its own unavailability, by equation (26.6), but through a shared entry it depends on its group's: $I_{\text{B,C}}$ moves with q_C, which the other three CRAC units share.

Building blocks. The enumeration with one unit pinned down and then up, or the closed forms (26.8), at $\hat{\mathbf{z}}$; $\mathbf{A}$ and $\mathbf{D}$ at $\hat{\mathbf{z}}$; $\boldsymbol{\Sigma}$ with each measure's logarithmic gradient for the bands; and, for the exact values beside them, draws of the three posteriors. The hall's values follow the third question, in Section 26.3.3.

26.3.2 The Fussell–Vesely and criticality importances

Engineering question. What share of the loss of cooling runs through each unit?

General form. The estimation form again, equation (2.9), for $f(\mathbf{z})$ a share of the top event, a quantity between zero and one.

Specialized form. The question is equation (2.9) with the state $\mathbf{z}$ the three logits, $f(\mathbf{z}) = I_{\mathrm{FV},i}$, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 26.2.2. One choice enters here: the share is defined by the exact top event with the unit pinned up, the part of the top event that would vanish if the unit never failed, rather than by the minimal cut sets that contain the unit. This makes the share an exact function of three evaluations and ties it to the Birnbaum importance by an identity, and it differs from the cut-set definitions where the unit's cut sets occur together with others; Section 26.4.1 sets the three definitions side by side. The Fussell–Vesely importance is

$$I_{\mathrm{FV},i} = \frac{U_{\mathrm{sys}} - U_i^-}{U_{\mathrm{sys}}}. \tag{26.10}$$

By the pivotal decomposition (26.6), $U_{\mathrm{sys}} - U_i^- = q_i\,(U_i^+ - U_i^-) = q_i\,I_{\mathrm{B},i}$, so

$$I_{\mathrm{FV},i} = \frac{q_i\,I_{\mathrm{B},i}}{U_{\mathrm{sys}}} = I_{\mathrm{CR},i}, \tag{26.11}$$

the criticality importance: the probability that unit i is down and critical, $q_i I_{\mathrm{B},i}$, as a fraction of the top event, which is the probability that i is down and critical given that the top event has occurred. The equality holds for every coherent tree whose units are independent given $\mathbf{q}$, because it needs only the pivotal decomposition; it fails when pinning one unit changes the probabilities of others, under a common cause, where $U_{\mathrm{sys}} - U_i^-$ and $q_i I_{\mathrm{B},i}$ are computed from different joint distributions. Where the Birnbaum importance weighs a unit by how often it decides, the Fussell–Vesely importance weighs that by how often it is down.

The cut-set forms. The standard definition through the minimal cut sets takes the probability of the union of the cut sets that contain i,

$$I^{\mathrm{c}}_{\mathrm{FV},i} = \frac{\Pr\bigl(\bigcup_{c\ni i} E_c \mid \mathbf{q}\bigr)}{U_{\mathrm{sys}}}, \qquad I^{\mathrm{r}}_{\mathrm{FV},i} = \frac{\sum_{c\ni i} P_c}{S_1}, \tag{26.12}$$

with E_c the event that every unit of cut set c is down, P_c its probability and S_1 the rare-event sum of equation (14.24); the second is the rare-event form, the share of S_1 that Chapter 14 reported, and its equation (14.37) is the same share for a group of cut sets. The three forms differ by the overlaps. Write $E_i^{\mathrm{rest}} = \bigcup_{c\ni i} E_{c\setminus i}$ for the event that the rest of some cut set of i is down and $E_i^{\mathrm{other}} = \bigcup_{c\not\ni i} E_c$ for the event that a cut set without i is down; neither involves unit i, so both are independent of its status given $\mathbf{q}$. Then $\Pr(\bigcup_{c\ni i} E_c) = q_i \Pr(E_i^{\mathrm{rest}})$, while $U_i^+ = \Pr(E_i^{\mathrm{rest}} \cup E_i^{\mathrm{other}})$ and $U_i^- = \Pr(E_i^{\mathrm{other}})$, so that

$$I^{\mathrm{c}}_{\mathrm{FV},i} - I_{\mathrm{FV},i} = \frac{q_i\bigl[\Pr(E_i^{\mathrm{rest}}) - \Pr(E_i^{\mathrm{rest}} \cup E_i^{\mathrm{other}}) + \Pr(E_i^{\mathrm{other}})\bigr]}{U_{\mathrm{sys}}} = \frac{q_i\,\Pr(E_i^{\mathrm{rest}} \cap E_i^{\mathrm{other}})}{U_{\mathrm{sys}}} \;\geq\; 0 : \tag{26.13}$$

the cut-set form also counts the configurations in which a cut set without i is down at the same time, which pinning i up would not cure. The gap is of the order of the joint occurrences, the S_2 terms of equation (14.25), and under the rare-event approximation, where joint occurrences are neglected, all three forms coincide (Exercise 26.3).

A group of units. The share of a group whose units are all pinned up together, $(U_{\mathrm{sys}} - U_{\mathrm{sys}}(\mathbf{q} \mid q_j = 0 \text{ for the group}))/U_{\mathrm{sys}}$, is not the sum of its units' shares. For the CRAC units, each unit's

share counts the pair configurations it belongs to, and each pair has two members, so the sum over the four units counts the pair term about twice. The measure that does add over the units of a group is the logarithmic slope, Section 26.3.5.

Building blocks. The pinned evaluations U_i^- and the top event, or the closed forms (26.8) times q_i; the cut-set probabilities of equation (14.18) for the cut-set forms; $\mathbf{\Sigma}$ and each share's logarithmic gradient for the bands, and for the exact values beside them, draws. The hall's values follow the third question, in Section 26.3.3.

26.3.3 The risk achievement and risk reduction worths

Engineering question. How much worse if the unit were certainly down, how much better if it never failed?

General form. The estimation form again, equation (2.9), for two positive ratios of evaluations of the top event.

Specialized form. The question is equation (2.9) with the state $\mathbf{z}$ the three logits, $f(\mathbf{z})$ in turn RAW_i and RRW_i, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \mathbf{\Sigma})$ of Section 26.2.2. One choice enters here: the unit certainly down stands for the unit out of service for a period, with the other units' unavailabilities unchanged. This makes the achievement worth the factor by which an outage of the unit multiplies the loss of cooling while it lasts, and it fails if the outage changes the others, as when the remaining CRAC units run harder while one is out (Section 26.6.5). The two worths are

$$\mathrm{RAW}_i = \frac{U_i^+}{U_{\mathrm{sys}}}, \qquad \mathrm{RRW}_i = \frac{U_{\mathrm{sys}}}{U_i^-}, \tag{26.14}$$

both at least one for a coherent tree. With the pivotal decomposition (26.6), $U_i^+ - U_{\mathrm{sys}} = (1 - q_i)(U_i^+ - U_i^-)$, and the definitions give the identities

$$\mathrm{RAW}_i - 1 = \frac{(1 - q_i)\, I_{\mathrm{B},i}}{U_{\mathrm{sys}}} = \frac{1 - q_i}{q_i}\, I_{\mathrm{FV},i}, \qquad 1 - \frac{1}{\mathrm{RRW}_i} = I_{\mathrm{FV},i}, \tag{26.15}$$

so that the Birnbaum importance and the top event, with q_i, determine all four measures, and the reduction worth is a monotone function of the Fussell–Vesely importance, ranking the units in the same order. The achievement worth ranks them by $I_{\mathrm{B},i}(1 - q_i)$, nearly the Birnbaum order when the unavailabilities are small.

A cut set of order one. A unit that is a minimal cut set alone has $U_i^+ = 1$, so

$$\mathrm{RAW}_{\mathrm{K}} = \frac{1}{U_{\mathrm{sys}}}, \tag{26.16}$$

the reciprocal of the top event: the controller's achievement worth is as uncertain as the top event, and its band is the top event's band inverted. Its order against any other unit's is certain, since $U_i^+ < 1$ for a unit that is not a cut set alone.

The reduction worth and the risk reduction of a change. A change that lowers one unit's unavailability from q_i to q_i' removes, by the straight line of equation (26.6),

$$U_{\mathrm{sys}} - U_{\mathrm{sys}}(\mathbf{q} \mid q_i') = (q_i - q_i')\, I_{\mathrm{B},i} = \Big(1 - \frac{q_i'}{q_i}\Big)\, I_{\mathrm{FV},i}\, U_{\mathrm{sys}}, \tag{26.17}$$

at most the perfect unit's $I_{\mathrm{FV},i} U_{\mathrm{sys}} = (1 - 1/\mathrm{RRW}_i)\, U_{\mathrm{sys}}$. The redundant controller of Chapters 14 and 25 replaces q_K by q_K^2, so it removes the fraction $1 - q_K$ of what a perfect controller

would, which is why Chapter 14 found its reduction, equation (14.42), to be the expected coefficient of q_K times $\mathbb{E}[q_K] - \mathbb{E}[q_K^2]$, and why Chapter 25 found it removes nearly the controller's whole attributed share of the risk (Section 25.3.5). The equation also says how the measures rank improvements. A change that removes a fixed amount $q_i - q_i'$ of unavailability ranks the units by the Birnbaum importance; a change that removes a fixed fraction $1 - q_i'/q_i$ of it, the more common kind for a maintenance program, ranks them by the Fussell–Vesely importance. A change of design, such as a fifth CRAC unit, alters the tree and is not a change of one unit's unavailability; no measure of a unit values it, and it was valued by its own enumeration in Chapter 14. For the loss-state term $C_{\mathrm{L}} t_{\mathrm{yr}} U_{\mathrm{sys}}$ of Chapter 25 the three ratio measures are unchanged and the Birnbaum importance is multiplied by $C_{\mathrm{L}} t_{\mathrm{yr}}$, by the linearity in the consequence of Section 25.3.2.

Building blocks. The pinned evaluations U_i^+ and U_i^-, the top event, and the identities (26.15) as a check; $\mathbf{\Sigma}$ and each worth's logarithmic gradient for the bands, and for the exact values beside them, draws.

The Birnbaum importances of the hall. Table 26.2 gives the four measures of one unit of each kind at the plug-in values. Each row is a unit; the columns give its unavailability, the top event with the unit pinned down and pinned up, and the four measures of equations (26.7), (26.10) and (26.14). The pinned values are those of Table 14.9. One row of the arithmetic, a CRAC unit: by equation (26.8), $3 \times 0.03 \times 0.97^2 \times 0.9999 \times 0.999 = 0.0846$, the difference $0.08833 - 0.00374$ of its pinned values.

Table 26.2: The importance measures of the hall's units at the plug-in values $q_C = 0.030$, $q_P = 0.010$, $q_K = 0.001$, where $U_{\mathrm{sys}} = 6.2806 \times 10^{-3}$. Pinned values by the enumeration of 128 configurations with one unit's unavailability set to one or to zero.

Unit	$\bar{q}_i$	U_i^+	U_i^-	$I_{\mathrm{B},i}$	$I_{\mathrm{FV},i}$	RAW_i	RRW_i
each CRAC unit, Ci	0.030	0.08833	3.743×10^{-3}	0.0846	0.404	14.06	1.678
each pump, Pi	0.010	0.01612	6.181×10^{-3}	0.00994	0.0158	2.567	1.016
controller, K	0.001	1	5.286×10^{-3}	0.9947	0.158	159.2	1.188

Summed over the units of each entry, equation (26.9), the slopes in the shared unavailabilities are $\mathbf{A}^{\mathsf{T}}\mathbf{I}_{\mathrm{B}} = (0.338,\ 0.0199,\ 0.995)^{\mathsf{T}}$, and with $\mathbf{D} = \mathrm{diag}(0.0291,\ 0.0099,\ 0.000999)$ the gradient in the state is $\mathbf{g} = 10^{-3}(9.846,\ 0.197,\ 0.994)^{\mathsf{T}}$, the vector of Section 14.3.5.

The Fussell–Vesely importances of the hall. By equation (26.11), a CRAC unit's share is $0.03 \times 0.0846 / 6.2806 \times 10^{-3} = 0.404$, the controller's $0.001 \times 0.9947 / 6.2806 \times 10^{-3} = 0.158$ and a pump's 0.0158, which are also the criticality importances. The cut-set forms of equation (26.12) give 0.417, 0.159 and 0.0159 for the union of each unit's cut sets, and the shares of the rare-event sum that Chapter 14 reported are 0.415, 0.154 and 0.015 (Table 14.9); the gap (26.13) is 0.013 for a CRAC unit, 0.0008 for the controller and 0.0001 for a pump. Pinned up as groups, the four CRAC units remove 0.825 of the top event, the two pumps 0.0158 and the controller 0.158, against the sum 1.616 of the four CRAC units' own shares.

The worths of the hall. The achievement worths of Table 26.2 are 159.2 for the controller, $1/U_{\mathrm{sys}}$ by equation (26.16), 14.06 for a CRAC unit and 2.567 for a pump; the reduction worths are 1.678 for a CRAC unit, 1.188 for the controller and 1.016 for a pump. Table 26.3 checks the identities (26.6) and (26.15). Each row is an identity; the columns give its form and its value for the three kinds of unit, each side computed separately from the pinned values.

Table 26.3: The identities among the importance measures at the plug-in values. Each side is computed from the pinned values of Table 26.2 and the top event; the two sides agree to twelve digits.

Identity	Form	CRAC unit	Pump	Controller
pivotal decomposition	$q_i U_i^+ + (1-q_i)U_i^- = U_{\text{sys}}$	6.2806×10^{-3}	6.2806×10^{-3}	6.2806×10^{-3}
share and criticality	$(U_{\text{sys}} - U_i^-)/U_{\text{sys}} = q_i I_{\text{B},i}/U_{\text{sys}}$	0.40404	0.015824	0.15838
achievement worth	$\text{RAW}_i - 1 = (1-q_i)I_{\text{B},i}/U_{\text{sys}}$	13.064	1.5665	158.22
achievement and share	$\text{RAW}_i - 1 = (1-q_i)I_{\text{FV},i}/q_i$	13.064	1.5665	158.22
reduction worth	$1 - 1/\text{RRW}_i = I_{\text{FV},i}$	0.40404	0.015824	0.15838

Figure 26.2 draws the four measures side by side. It has four panels, one per measure; in each the horizontal axis lists the three kinds of unit, a CRAC unit in green, a pump in gold and the controller in blue, and the vertical axis is the measure, on a logarithmic scale for the Birnbaum importance and the achievement worth and on a linear scale for the share and the reduction worth, whose vertical axis starts at one.

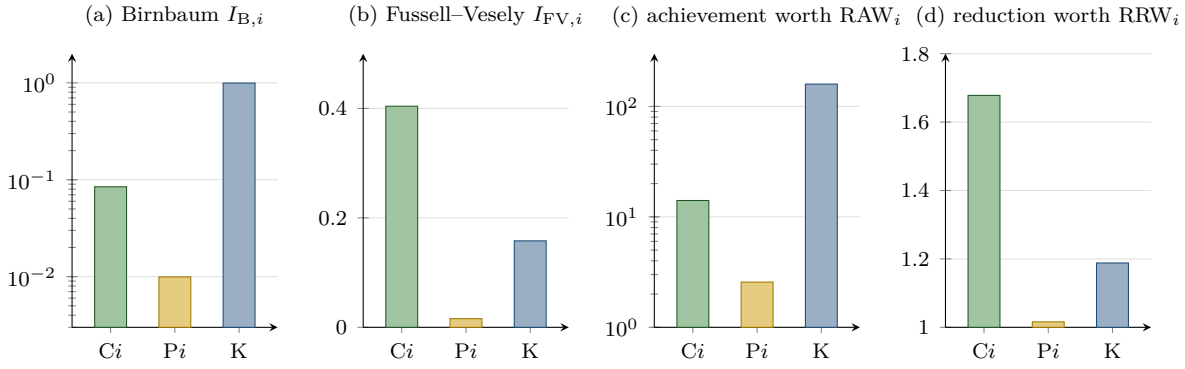


Figure 26.2: The four importance measures of a CRAC unit (green), a pump (gold) and the controller (blue) at the plug-in values, from Table 26.2: (a) the Birnbaum importance, logarithmic scale; (b) the Fussell–Vesely importance; (c) the risk achievement worth, logarithmic scale; (d) the risk reduction worth, from one.

The Birnbaum importance and the achievement worth order the units K, CRAC, pump; the share and the reduction worth order them CRAC, K, pump.

26.3.4 The firmness of the rankings

Engineering question. Are the rankings firm?

General form. The threshold form of Chapter 2, equation (2.12), for the event that one unit's measure exceeds another's, $f(\mathbf{z}) = \ln I_{\text{K}}(\mathbf{z}) - \ln I_{\text{C1}}(\mathbf{z})$ and $f^* = 0$:

$$\Pr(I_{\text{K}} > I_{\text{C1}} \mid \mathbf{y}) = \int \mathbf{1}\left[\ln I_{\text{K}}(\mathbf{z}) - \ln I_{\text{C1}}(\mathbf{z}) > 0\right] p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}.$$

The probability itself is the answer to "is it firm"; the question sets no assurance level.

Specialized form. The question is equation (2.12) with $f(\mathbf{z})$ the logarithm of the ratio of two units' measures, $f^* = 0$, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$

of Section 26.2.2, so that each probability is a normal tail of that logarithm. No new choice enters. For each measure the event reduces to an inequality between the unavailabilities.

Birnbaum. By equation (26.8), $I_{\mathrm{B,K}} > I_{\mathrm{B,C}}$ is $(1-q_C)^3(1+3q_C) > 3q_C(1-q_C)^2(1-q_K)$, the pump factor cancelling, which holds for every q_K when $(1-q_C)(1+3q_C) > 3q_C$, that is when $1-q_C-3q_C^2>0$, or $q_C < (\sqrt{13}-1)/6 = 0.434$:

$$\Pr(I_{\mathrm{B,K}} > I_{\mathrm{B,C}} \mid \mathbf{y}) \;\geq\; \Pr(q_C < 0.434 \mid \mathbf{y}) \approx \Phi\Big(\frac{-0.264-\hat{\xi}_C}{\sqrt{\Sigma_{CC}}}\Big), \tag{26.18}$$

a tail of the CRAC entry alone, -0.264 being the logit of 0.434; the exact value beside it is $F_{\mathrm{Beta}}(0.434;\; a_C, b_C)$, with $F_{\mathrm{Beta}}(x;\; a, b)$ the distribution function of $\mathrm{Beta}(a,b)$, the regularized incomplete Beta function of Chapter 5.

Achievement and reduction worths. By equation (26.16), $\mathrm{RAW_K} > \mathrm{RAW}_i$ for every other unit at every state, with probability one. By equation (26.15) the reduction worth is an increasing function of the Fussell–Vesely importance, so its rankings and their probabilities are those of the Fussell–Vesely importance.

Fussell–Vesely. The two shares have the same denominator, so by equation (26.11) the event is $q_K I_{\mathrm{B,K}} > q_C I_{\mathrm{B,C}}$, and with equation (26.8) the pump factor cancels again: $q_K/(1-q_K) > 3q_C^2/((1-q_C)(1+3q_C))$. In the logits of the state the event is the sign of the contrast d of equation (14.33), a difference of two logarithms (not the deviation $\mathbf{d}$ of Chapter 20 nor the held-out departure d_k of Chapter 21), here written for two shares,

$$d(\mathbf{z}) = \ln\frac{I_{\mathrm{FV,K}}}{I_{\mathrm{FV,C}}} = \xi_K - 2\xi_C - \ln 3 - \ln(1-q_C) + \ln(1+3q_C), \qquad I_{\mathrm{FV,K}} > I_{\mathrm{FV,C}} \iff d(\mathbf{z}) > 0, \tag{26.19}$$

nearly linear in the state: the controller's share overtakes a CRAC unit's when its logit exceeds twice the CRAC logit plus $\ln 3$. Under the Laplace posterior d has the gradient

$$\mathbf{g}_d = \frac{\partial d}{\partial \mathbf{z}} = \begin{bmatrix} -2 + q_C + 3q_C(1-q_C)/(1+3q_C) \\ 0 \\ 1 \end{bmatrix}, \qquad \Pr(I_{\mathrm{FV,K}} > I_{\mathrm{FV,C}} \mid \mathbf{y}) \approx \Phi\Big(\frac{d(\hat{\mathbf{z}})}{\sqrt{\mathbf{g}_d^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}_d}}\Big), \tag{26.20}$$

whose two terms $g_{d,j}^2\Sigma_{jj}$ say which record leaves the order open; the pumps' entry has no term. The value beside it is exact: given q_C the event is a tail of the controller's posterior, $\mathrm{Beta}(a_K, b_K)$ with $a_K = 1$, whose tail at x is $(1-x)^{b_K}$; at the threshold $x = \eta/(1+\eta)$, $\eta(q_C) = 3q_C^2/((1-q_C)(1+3q_C))$, that is $(1+\eta)^{-b_K}$, and only the average over q_C remains,

$$\Pr(I_{\mathrm{FV,K}} > I_{\mathrm{FV,C}} \mid \mathbf{y}) = \int_0^1 \big(1+\eta(q)\big)^{-b_K}\,\mathrm{Beta}(q;\; a_C, b_C)\,\mathrm{d}q, \tag{26.21}$$

a one-dimensional integral against the CRAC posterior.

The spread of the two shares. Near the plug-in values $I_{\mathrm{FV,K}} \approx q_K/(q_K + 6q_C^2 + q_P^2)$, so its odds are nearly $q_K/(6q_C^2+q_P^2)$ and its logit moves one for one with ξ_K: the controller's share carries the whole spread of the controller's logit, which a record with k_K units found down sets at $\Sigma_{KK} \approx 1/k_K$ by equation (14.7), whatever the number of checks. A CRAC unit's share is bounded above by one half, the value it takes when the CRAC pairs carry everything, since each unit belongs to half of the six pairs (Exercise 26.4).

The order of all three units under a measure is a joint event of two such inequalities, the bivariate normal tail of the two contrasts under $\boldsymbol{\Sigma}$; its exact value is read from the same draws.

Building blocks. The normal tails of equations (26.18) and (26.20), with $\mathbf{g}_d$ and $\boldsymbol{\Sigma}$, and the same for the other pairs and the joint orders; and, for the exact values beside them, the incomplete Beta tail, the quadrature of equation (26.21) and draws of the three posteriors.

The bands and the rankings of the hall. Table 26.4 gives the posterior of each measure. Each row is a measure of one unit; the columns give its value at $\hat{\mathbf{z}}$, the median of the Laplace normal, its central 95 % band by the delta method on its logarithm, and then, for the exact values beside them, its median, mean and central 95 % band over the 10^5 draws. The exact Birnbaum bands are those of Table 14.9 to the printed digits.

Table 26.4: Posterior bands of the importance measures: from the state, the value at $\hat{\mathbf{z}}$ and the delta band on the logarithm; exact, 10^5 draws of $q_C \sim \text{Beta}(3, 97)$, $q_P \sim \text{Beta}(4, 396)$, $q_K \sim \text{Beta}(1, 999)$, numpy `default_rng(26)`; plug-in at the posterior means.

Measure and unit	At $\hat{\mathbf{z}}$	95 % band, state	Median, exact	Mean, exact	95 % band, exact
I_{B}, CRAC unit	0.0846	[0.030, 0.241]	0.0762	0.0829	[0.019, 0.183]
I_{B}, pump	0.00994	[0.0037, 0.0264]	0.00912	0.00994	[0.0027, 0.0217]
I_{B}, controller	0.9947	[0.984, 1.006]	0.9957	0.9932	[0.973, 0.9997]
I_{FV}, CRAC unit	0.404	[0.25, 0.65]	0.410	0.370	[0.084, 0.479]
I_{FV}, pump	0.0158	[0.0011, 0.226]	0.0154	0.0340	[0.00085, 0.187]
I_{FV}, controller	0.158	[0.014, 1.84]	0.131	0.208	[0.0037, 0.771]
RAW, CRAC unit	14.06	[6.4, 31.0]	14.0	15.5	[6.7, 32.8]
RAW, pump	2.567	[0.73, 9.04]	2.65	3.80	[1.22, 13.4]
RAW, controller	159.2	[26, 991]	186	295	[35, 1215]
RRW, CRAC unit	1.678	[1.215, 2.318]	1.694	1.628	[1.092, 1.918]
RRW, pump	1.016	[0.974, 1.060]	1.016	1.040	[1.001, 1.229]
RRW, controller	1.188	[0.749, 1.884]	1.151	1.523	[1.004, 4.36]

A band on the logarithm can pass a bound the measure cannot: the controller's slope passes one, its share and a CRAC unit's pass one and one half, an achievement worth of a pump and every reduction worth but the CRAC unit's pass one from above. These are the skew and the bound of the state route (Section 26.4.2), and the exact bands beside them stay inside. The controller's exact Birnbaum band spans three percent and its share's a factor of two hundred; its achievement-worth band is the reciprocal of the top event's band, $[1.0 \times 10^{-3},\ 3.9 \times 10^{-2}]$ from the state and $[8.2 \times 10^{-4},\ 2.8 \times 10^{-2}]$ exactly. The controller's posterior, from one unit found down in a thousand checks, has $\sqrt{\Sigma_{KK}} = 1.000$ and a central band $[2.5 \times 10^{-5},\ 3.7 \times 10^{-3}]$, a factor of about one hundred and fifty (Table 14.3).

Table 26.5 gives the probabilities of the orders. Each row is a measure; the columns give the posterior probabilities that the controller exceeds a CRAC unit, that the controller exceeds a pump and that a CRAC unit exceeds a pump, then the order of the three at the plug-in values and the posterior probability of that order, each from the state with the exact value in parentheses. For the reduction worth the event is the same as for the share, by equation (26.15), and its probabilities are the share's.

By the share the controller ranks above a CRAC unit with probability, by equation (26.20) with $d(\hat{\mathbf{z}}) = -0.937$ and $\sqrt{\mathbf{g}_d^{\mathsf{T}} \boldsymbol{\Sigma} \mathbf{g}_d} = 1.493$,

$$\Phi\Big(\frac{-0.937}{1.493}\Big) = 0.265;$$

the value beside it comes from the exact posterior, 0.243: 0.2427 by the quadrature of equation (26.21), 0.2426 over the draws, 0.022 lower. Of the full Fussell–Vesely orders, CRAC, K, pump has probability 0.685; K, CRAC, pump 0.253; CRAC, pump, K 0.047; and the other three together 0.015 (exactly 0.649, 0.218, 0.104 and 0.029).

Figure 26.3 draws the contest of the shares. The left panel has the share on its horizontal axis and the posterior density on the vertical: the green curve is a CRAC unit's share and the

Table 26.5: Posterior probabilities of the orders of the units under each measure: normal tails of the log ratios under $\mathbf{\Sigma}$, with the bivariate normal for the orders; in parentheses from the 10^5 draws, numpy `default_rng(26)`. The Birnbaum entry against a CRAC unit is bounded by equation (26.18), the achievement worth's against any unit is one by equation (26.16), and the exact Fussell–Vesely entry against a CRAC unit is equation (26.21).

Measure	K above a CRAC unit	K above a pump	CRAC unit above a pump	Plug-in order	Probability of that order
Birnbaum	$1 - 2 \times 10^{-8}$ ($1 - 9.5 \times 10^{-22}$) or more	1 (1)	0.998 (0.993)	K, CRAC, pump	0.998 (0.993)
Fussell–Vesely	0.265 (0.243)	0.949 (0.889)	0.985 (0.971)	CRAC, K, pump	0.685 (0.649)
achievement worth	1 (1)	1 (1)	0.99999 (0.993)	K, CRAC, pump	0.99999 (0.993)
reduction worth	0.265 (0.243)	0.949 (0.889)	0.985 (0.971)	CRAC, K, pump	0.685 (0.649)

blue curve the controller's, both over the exact draws. The right panel has on its horizontal axis the logarithm of their ratio, d of equation (26.19), and its posterior density on the vertical axis, the black curve over the exact draws and the dotted blue curve the normal of the state, $\mathcal{N}(-0.937,\ 1.493^2)$; the shaded area to the right of zero under the black curve, where the controller's share is the larger, is the exact $\Pr(I_{FV,K} > I_{FV,C} \mid \mathbf{y})$.

The two marginals overlap from 0.05 to 0.5. The state's d has median -0.937 and central band $[-3.86,\ 1.99]$; the exact density has median -1.13 and central band $[-4.85,\ 2.19]$. Both cross zero.

26.3.5 The band of the top event

Engineering question. Why does the band of the loss of cooling use the same numbers?

General form. The estimation form, equation (2.11), for the central interval of $f(\mathbf{z}) = \ln U_{\text{sys}}$, with the spread under the Laplace posterior from the delta method, $\text{sd}(f \mid \mathbf{y}) \approx \sqrt{\mathbf{g}_f^{\mathsf{T}} \mathbf{\Sigma} \mathbf{g}_f}$, $\mathbf{g}_f = \mathbf{g}/U_{\text{sys}}$.

Specialized form. The question is equation (2.11) with $f(\mathbf{z}) = \ln U_{\text{sys}}$ and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \mathbf{\Sigma})$ of Section 26.2.2. No new choice enters. The band reported is the delta method on the logarithm of the top event, equation (14.29), whose gradient carries the budget; the value beside it is the central interval of the draws of the three Beta posteriors (Table 14.13). What this question adds is the reading of that gradient. By equation (26.9) the gradient of the top event in the state is the Birnbaum vector summed over the units that share an entry and scaled by $\mathbf{D}$; divided by the top event, its entry j is

$$\frac{g_j}{U_{\text{sys}}} = \frac{q_j(1 - q_j)}{U_{\text{sys}}} \sum_{i \in j} I_{B,i} = (1 - q_j) \sum_{i \in j} I_{CR,i} = (1 - q_j) \sum_{i \in j} I_{FV,i}, \tag{26.22}$$

by equation (26.11), with $i \in j$ the units whose unavailability is entry j. The gradient of the logarithm of the top event in the state is the Fussell–Vesely vector, summed over the units of each entry and scaled by $1 - q_j$: a logit is a relative change of an unavailability, and the share of a unit is the relative change of the top event per relative change of the unit's unavailability.

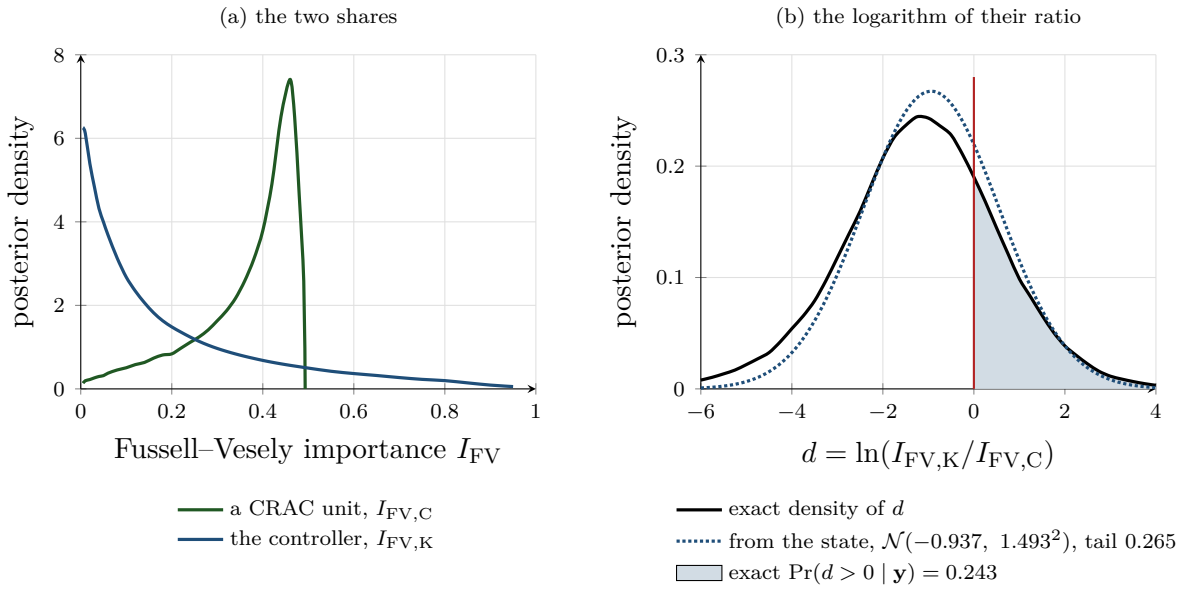


Figure 26.3: The Fussell–Vesely importances of a CRAC unit and of the controller under the posterior: (a) their marginal densities, kernel estimates on the logit of the share mapped back, the CRAC unit's closed at the largest share drawn, 0.493; (b) the density of the logarithm of their ratio, equation (26.19), kernel estimate, with the probability that the controller's share is the larger shaded (red line at $d = 0$), against the normal of the state (dotted). 10^5 draws, numpy `default_rng(26)`.

The Laplace variance of Chapter 14 is then

$$\operatorname{Var}(\ln U_{\mathrm{sys}} \mid \mathbf{y}) \approx \frac{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}{U_{\mathrm{sys}}^2} = \sum_j (1 - q_j)^2 \Big(\sum_{i \in j} I_{\mathrm{FV},i}\Big)^2 \Sigma_{jj}, \tag{26.23}$$

the diagonal sum because $\boldsymbol{\Sigma}$ is diagonal by the pattern (26.3): equation (14.46) with its terms named. On the unavailabilities themselves the same product is $\mathbf{I}_{\mathrm{B}}^{\mathsf{T}}\mathbf{A}\,\boldsymbol{\Sigma}_{\mathbf{q}}\,\mathbf{A}^{\mathsf{T}}\mathbf{I}_{\mathrm{B}}$, the Birnbaum vector against the Laplace covariance $\boldsymbol{\Sigma}_{\mathbf{q}} = \mathbf{D}\boldsymbol{\Sigma}\mathbf{D}$ of the three shared unavailabilities. The gradient that the band uses is the gradient the Birnbaum importance is; its logarithmic form is the Fussell–Vesely vector. The measures and the Laplace band are one computation, three evaluations per unit and one covariance; its terms $V_j = (g_j/U_{\mathrm{sys}})^2\Sigma_{jj}$ are the budget of the band by entry, the uncertainty importance of each group's record, which step 7 uses to rank the records (Section 26.6.1).

Building blocks. The Fussell–Vesely importances at $\hat{\mathbf{z}}$ and the diagonal of $\boldsymbol{\Sigma}$ for the gradient, the band and the budget; and, for the exact band beside them, the draws of the three posteriors, as in Chapter 14.

The band of the top event of the hall. By equation (26.22), the logarithmic gradient is

$$\frac{\mathbf{g}}{U_{\mathrm{sys}}} = \begin{bmatrix} 0.97 \times 4 \times 0.40404 \\ 0.99 \times 2 \times 0.015824 \\ 0.999 \times 0.15838 \end{bmatrix} = \begin{bmatrix} 1.568 \\ 0.031 \\ 0.158 \end{bmatrix},$$

the vector of Section 14.3.5. With the diagonal of $\boldsymbol{\Sigma}$, equation (26.23) gives the budget terms $V_C = 0.845$, $V_P = 2.5 \times 10^{-4}$ and $V_K = 0.0251$, 97.1 %, 0.03 % and 2.9 % of the variance 0.870 of

$\ln U_{\text{sys}}$, Chapter 14's budget. The standard deviation is 0.933, and about the median 6.28×10^{-3} the band is

$$6.2806 \times 10^{-3} \exp(\pm 1.96 \times 0.933) = [1.0 \times 10^{-3},\ 3.9 \times 10^{-2}].$$

The value beside it comes from the exact posterior: Chapter 14's 4×10^5 draws give the median 5.42×10^{-3} and the central band $[8.2 \times 10^{-4},\ 2.8 \times 10^{-2}]$ (Table 14.13), and the 10^5 draws of this chapter the same band to the printed digits; the state's upper end is higher by 0.011.

26.4 The checks (step 5)

The hall has no physics row and no exact constraint, so no limit $\sigma_p \to 0$ of a model tolerance arises; the bound $0 < q < 1$ is carried by the logits. What can fail is the arithmetic of the pinned evaluations, the definition of the share, the Laplace route against the exact one, and the choice of each record alone as the posterior.

26.4.1 The enumeration, the closed forms and the identities

Table 26.6 collects the checks. Each row is a quantity; the columns give its value by the closed form, the cut-set formula or the Laplace route, its value by the full calculation it is checked against, and the agreement.

Table 26.6: The checks of step 5. Draws: 10^5, numpy `default_rng(26)`; bands are central 95 % intervals.

Quantity	Closed form, cut sets or Laplace	Full calculation	Agreement
pinned values $U_i^{\pm}$	closed forms of equation (26.2) with one factor pinned	enumeration of 128 configurations with q_i set to 1 or 0	to 10^{-15}
$I_{\text{B},i}$	equation (26.8)	$U_i^+ - U_i^-$; central differences of (26.2)	to 10^{-15}; to eight digits
identities	equations (26.6), (26.11), (26.15)	each side from the pinned values	to 10^{-12}
share, CRAC unit	cut-set union 0.417; rare-event 0.415	exact 0.404	gap (26.13) 0.013
share, controller	cut-set union 0.159; rare-event 0.154	exact 0.158	within 0.004
$\Pr(I_{\text{FV,K}} > I_{\text{FV,C}})$	Laplace (26.20) 0.265	quadrature (26.21) 0.2427; draws 0.2426	Laplace 0.02 high
band of $I_{\text{B,C}}$	Laplace [0.030, 0.241]	draws [0.019, 0.183]	too high by 1.3 to 1.6
band of $I_{\text{FV,K}}$	Laplace [0.014, 1.84]	draws [0.0037, 0.771]	upper end above one
band of RAW_{K}	Laplace [26, 991]	draws [35, 1215]	$1/U_{\text{sys}}$ both ways
band of U_{sys}	Laplace on $\ln U$, sd 0.933	draws, sd of $\ln U$ 0.909	width right, center high

The pinned evaluations. Each of the six pinned values of Table 26.2 was computed twice, by the enumeration of the 128 configurations with the unit's unavailability set to one or to zero and by the closed form (26.2) with the unit's factor changed, and the two agree to 10^{-15}; the closed slopes (26.8) agree with the differences of the pinned values to the same digits and with central differences of the closed form in each unit's unavailability to eight. The enumeration with q_i set to an intermediate value lies on the straight line of equation (26.6), as multilinearity requires.

The identities. The five identities of Table 26.3 hold to 10^{-12} for each kind of unit, each side computed from the pinned values alone. They hold draw by draw as well, so that the bands of the reduction worths in the exact posterior are those of the shares mapped through $1/(1 - I_{\mathrm{FV}})$; the delta bands of the state route are not so mapped, a CRAC unit's share band giving [1.33, 2.86] against its reported [1.215, 2.318]. The probabilities of the reduction-worth orders equal those of the share orders exactly, as Table 26.5 shows.

The definition of the share. The three forms of the share order the units alike, a CRAC unit, the controller, a pump. The cut-set union exceeds the exact share by the gap of equation (26.13), 0.013 for a CRAC unit, whose cut sets overlap the other CRAC pairs, and by under 0.001 for the controller and a pump; the rare-event form of Chapter 14 lies within 0.011 of the exact one for each unit. The choice of definition does not move the answer to the second question.

26.4.2 The Laplace route against the draws

The ranking of the shares. At $\hat{\mathbf{z}}$, $d = \ln(0.15838/0.40404) = -0.937$, and the gradient of equation (26.20) is $\mathbf{g}_d = (-1.890,\ 0,\ 1)^{\mathsf{T}}$, so that $\mathbf{g}_d^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}_d = 1.227 + 1.001 = 2.228$, 55 % from ξ_C and 45 % from ξ_K, and $\Phi(-0.937/1.493) = 0.265$, against the exact 0.243. Over the draws d has the standard deviation 1.76 and the mean -1.18; with the exact variances of the logits, equation (14.9), in place of the diagonal of $\boldsymbol{\Sigma}$, $\mathbf{g}_d^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}_d$ becomes 3.09, the draws' spread, and the normal tail 0.297. The Laplace route has the spread too small and the center too high, and the two errors partly cancel; both routes call the order unsettled.

The measures' bands. Under the Laplace posterior each measure has the band of equation (14.29) on its logarithm, with its gradient by central differences in the logits, the bands of Table 26.4. For a CRAC unit's slope the band [0.030, 0.241] lies 1.3 to 1.6 times above the draws' [0.019, 0.183], the shift of the right-skewed CRAC posterior. For the controller's slope the Laplace band [0.984, 1.006] passes one, which a probability cannot. For the controller's share the band [0.014, 1.84] ends above one, and for a CRAC unit's, [0.25, 0.65], above its bound of one half: a share near a bound has no symmetric band on its logarithm. For the controller's achievement worth the Laplace band [26, 991] is the reciprocal of the Laplace band of the top event, by equation (26.16), against the draws' [35, 1215]. The other pinned measures pass their bounds as well, a pump's achievement worth below one and the reduction worths of a pump and of the controller below one. Chapter 14 found the Laplace preposterior close to the exact one for the CRAC entry and too optimistic for the controller's (Section 14.6.3).

The top event. The Laplace route of equation (14.29), with the gradient of equation (26.22), gives Chapter 14's standard deviation 0.933 of $\ln U_{\mathrm{sys}}$ and band $[1.0 \times 10^{-3},\ 3.9 \times 10^{-2}]$, against the draws' standard deviation 0.909 and band $[8.2 \times 10^{-4},\ 2.8 \times 10^{-2}]$ (Table 14.10). On the unavailabilities themselves the product $\mathbf{I}_{\mathrm{B}}^{\mathsf{T}}\mathbf{A}\,\boldsymbol{\Sigma}_{\mathbf{q}}\,\mathbf{A}^{\mathsf{T}}\mathbf{I}_{\mathrm{B}}$ is 3.43×10^{-5}, the standard deviation 5.86×10^{-3} that Chapter 14 printed as 5.9×10^{-3}, against 7.59×10^{-3} over the draws; a normal band on U_{sys} with that spread puts 14 % of its probability below zero.

Figure 26.4 draws the logarithmic Laplace band against the draws. The horizontal axis is the probability of the top event on a logarithmic scale; the vertical axis is the density per decade. The blue curve is the density of the draws; the dashed black curve is the Laplace normal on $\ln U_{\mathrm{sys}}$, centered at the plug-in value with the standard deviation 0.933; the two pairs of vertical lines mark the two central 95 % bands.

The two curves have nearly the same width, and the Laplace curve sits to the right by a factor of 1.16, centered on the plug-in value where the draws' median is 5.42×10^{-3}: the shift Chapter 14 traced to the right skew of the CRAC posterior.

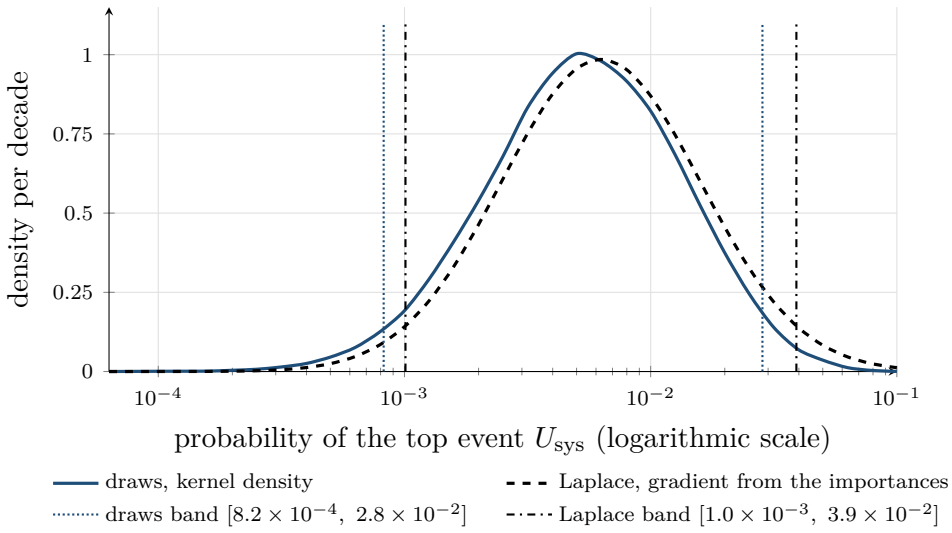


Figure 26.4: The band of the probability of losing cooling, the Laplace route against the draws: the posterior density of $\log_{10} U_{\text{sys}}$ over 10^5 draws of the three Beta posteriors (kernel estimate, numpy `default_rng(26)`), against the Laplace normal on $\ln U_{\text{sys}}$ with mean $\ln(6.2806 \times 10^{-3})$ and standard deviation 0.933 from the gradient of equation (26.22), both drawn per decade; vertical lines, the two central 95 % bands.

26.4.3 Sensitivity to the prior

Each posterior was taken as its record alone. Table 26.7 repeats the measures and the rankings under the two priors of Section 14.4.3, added to every record. Each row is a prior; the columns give the three posteriors, the top event at their means, the shares of a CRAC unit and of the controller and the controller's achievement worth at the means, which are the values at the MAP state refitted under that prior, the probability that the controller's share exceeds a CRAC unit's, by equation (26.20) with the exact value of equation (26.21) in parentheses, the probability of the plug-in order of the shares in the same way, and the probability that the controller's slope exceeds a CRAC unit's. One row of the arithmetic, the uniform prior: the controller's posterior is Beta(2, 1000), of mean 0.0020, twice the record's, and its share at the means rises from 0.158 to 0.182. The draws are 10^5, numpy `default_rng(26)`.

Table 26.7: Sensitivity of the importance measures and their rankings to the prior on each group's unavailability. Measures at the posterior means, the refitted MAP state; probabilities from the state, normal tails under $\boldsymbol{\Sigma}$; in parentheses Pr(K > C) for the shares by the quadrature of equation (26.21) and the other probabilities from 10^5 draws, numpy `default_rng(26)`.

Prior	Posteriors q_C, q_P, q_K	U_{sys} at means	I_{FV} CRAC, K	RAW_{K}	Pr(K > C), shares	Plug-in share order	Pr(K > C), slopes
record alone	Beta(3, 97), (4, 396), (1, 999)	6.28×10^{-3}	0.404, 0.158	159	0.265 (0.243)	0.685 (0.649)	1 (1)
uniform	Beta(4, 98), (5, 397), (2, 1000)	1.09×10^{-2}	0.391, 0.182	92	0.259 (0.261)	0.729 (0.715)	1 (1)
Jeffreys	Beta(3.5, 97.5), (4.5, 396.5), (1.5, 999.5)	8.49×10^{-3}	0.395, 0.175	118	0.267 (0.260)	0.711 (0.695)	1 (1)

The level moves and the orders do not. The controller's achievement worth falls from 159 to 92 under the uniform prior, because it is the reciprocal of a top event that rises by seventy

percent at the means; the shares move by at most 0.024; the probability that the controller's share exceeds a CRAC unit's stays between 0.26 and 0.27 (0.24 and 0.26 exactly), and the Birnbaum order stays certain. The report gives the orders with their probabilities, which no prior tried changes, and the achievement worth of the controller with its range over the priors.

26.5 The report (step 6)

The first reading ranked the controller first because its slope is twelve times a CRAC unit's, and that ranking answers only one of the owner's three requests. By the share of the loss of cooling that runs through each unit, the measure a watch follows, a CRAC unit ranks first with 0.404 against the controller's 0.158, and that order is itself unsettled: the controller's share exceeds a CRAC unit's with posterior probability 0.265 (0.243 exactly), because its record holds one unit found down. The slope's order and the achievement worth's, controller first, hold in every draw. And the record that most narrows the band of the loss of cooling is the CRAC record, with 97 % of that band.

26.5.1 The answers and their bands

Table 26.8 gives the answers. Each row is a question of Table 26.1, in its order; the second column gives the answer from the posterior of the state, at the MAP state with the records alone, with its band or probability; the third the exact value beside it; and the fourth the comparison: the central 95 % posterior interval, a posterior probability, or the range over the priors of Table 26.7.

The report line. The controller is the unit the hall can never let fail, and a change in its unavailability moves the loss of cooling most; the losses of cooling run mostly through the CRAC units, each carrying about 40 % of them, though whether the controller or a CRAC unit carries more is not settled by the records; the figures assume that units fail independently given their unavailabilities.

26.5.2 Three rankings for three decisions

The owner's three requests take three measures. The outage rule asks which unit may never be let fail, and the achievement worth answers it with certainty: the controller, whose outage makes the loss of cooling certain, 159 times its present probability, against fourteen for a CRAC unit. The watch asks where the losses come from, and the share answers it, a CRAC unit first but not firmly; a watch that must not miss the larger contributor therefore covers both. The improvement program asks which change buys most, and equation (26.17) splits it in two: a program that removes a fixed amount of unavailability follows the Birnbaum importance, the controller first; one that removes a fixed fraction of it follows the share, a CRAC unit first; a change of design is valued by its own enumeration, as Chapters 14 and 25 valued the fifth CRAC unit. None of the three measures says which record to extend; the budget of the band does.

Common mistake

Reading a list of Birnbaum importances as a ranking for maintenance attention. The Birnbaum importance of a unit is the probability that it is critical and does not depend on how often the unit is down; the controller's 0.995 is far above a CRAC unit's 0.085, yet a CRAC unit's failures carry two and a half times the controller's share of the loss of cooling. The ranking for a watch is by the share, and it is reported with the probability that it holds.

Table 26.8: The answers to the five questions asked of the importance of the hall's units. Bands are central 95 % posterior intervals: by the delta method on the logarithm from the state; exact from 10^5 draws, numpy `default_rng(26)`, and the exact band of the top event that of Chapter 14.

Question	Answer (state route)	Exact	Comparison
Which unit's unavailability matters most to the loss of cooling?	the controller, $I_{\mathrm{B}} = 0.9947$; a CRAC unit 0.0846; a pump 0.00994; bands [0.984, 1.006], [0.030, 0.241], [0.0037, 0.0264]	bands [0.973, 0.9997], [0.019, 0.183], [0.0027, 0.0217]	in the shared unavailabilities (0.338, 0.0199, 0.995)
What share of the loss of cooling runs through each unit?	a CRAC unit 0.404; the controller 0.158; a pump 0.0158; equal to the criticality importance; bands [0.25, 0.65], [0.014, 1.84], [0.0011, 0.226]	bands [0.084, 0.479], [0.0037, 0.771], [0.00085, 0.187]	cut-set forms 0.417, 0.159, 0.0159; the four CRAC units as a group 0.825
How much worse if the unit were certainly down, how much better if it never failed?	RAW: controller 159.2, CRAC unit 14.06, pump 2.567, bands [26, 991], [6.4, 31.0], [0.73, 9.04]; RRW: CRAC unit 1.678, controller 1.188, pump 1.016, bands [1.215, 2.318], [0.749, 1.884], [0.974, 1.060]	RAW bands [35, 1215], [6.7, 32.8], [1.22, 13.4]; RRW bands [1.092, 1.918], [1.004, 4.36], [1.001, 1.229]	$\mathrm{RAW_K}$ 92 to 159 at the means over the priors
Are the rankings firm?	by slope and achievement worth, the controller first, certainly; by share and reduction worth, a CRAC unit first; $\Pr(I_{\mathrm{FV,K}} > I_{\mathrm{FV,C}} \mid \mathbf{y}) = 0.265$; the full share order CRAC, K, pump with probability 0.685	0.243; 0.649	0.26 to 0.27 over the priors (0.24 to 0.26 exact)
Why does the band of the loss of cooling use the same numbers?	its gradient is the Birnbaum vector, and on the logarithm the shares: $\mathbf{g}/U_{\mathrm{sys}} = (1.568, 0.031, 0.158)$; band $[1.0 \times 10^{-3}, 3.9 \times 10^{-2}]$, median 6.28×10^{-3}	band $[8.2 \times 10^{-4}, 2.8 \times 10^{-2}]$, median 5.42×10^{-3} (Chapter 14)	the CRAC entry carries 97.1 % of the variance of $\ln U_{\mathrm{sys}}$

26.6 The next measurement (step 7)

The slopes and the achievement worths are settled; the order of the controller and a CRAC unit by their shares is not, and it decides whether a watch on the CRAC units alone would miss the larger contributor. The next measurement is the record that would settle this order, weighed beside the record that would narrow the band of the top event.

26.6.1 The candidate records

Four candidates are weighed, as in Chapters 14 and 25: a hundred or four hundred further checks of the CRAC units, kept as the present record was, and a hundred or a thousand further checks of the controller. Two criteria rank them.

The order of the shares. For n_a further CRAC checks with n_{d} units found down, the state is refitted, at $\hat{\xi}_C = \ln((3+n_{\mathrm{d}})/(97+n_a-n_{\mathrm{d}}))$ and $\Sigma_{CC} = 1/(3+n_{\mathrm{d}})+1/(97+n_a-n_{\mathrm{d}})$; each count is weighted by the Laplace predictive, the binomial probability averaged by Gauss–Hermite over the entry; and for each count the probability that the controller's share exceeds a CRAC unit's is recomputed by equation (26.20). A record of controller checks is treated in the same way on ξ_K. The value beside it comes from the exact posterior, $\mathrm{Beta}(3+n_{\mathrm{d}},\ 97+n_a-n_{\mathrm{d}})$ with the Beta–binomial predictive of equation (14.47) and the quadrature of equation (26.21), and $\mathrm{Beta}(1+n_{\mathrm{d}},\ 999+n_a-n_{\mathrm{d}})$ for the controller. By the law of total probability, equation (14.48), the exact probability averaged over the counts is its present value for every candidate. A record settles the order when the probability after it is at least 0.95, the controller's share the larger, or at most 0.05, a CRAC unit's.

The band of the top event. A further check of a group is taken to be of the kind already recorded, a Bernoulli trial on the group's entry, with its information evaluated at $\hat{\mathbf{z}}$: this gives the value of a record in closed form, without an average over its outcomes, and it is least reliable for the controller, whose logit posterior is the least normal (Section 14.6.3). One check of group j adds $q_j(1-q_j)$ to the diagonal entry Λ_{jj} of $\mathbf{J}^{\mathsf{T}}\mathbf{J}$, Section 14.1.7, so after n_a checks

$$\Sigma_{jj}(n_a) = \frac{1}{1/\Sigma_{jj} + n_a q_j(1-q_j)}, \qquad \mathrm{Var}(\ln U_{\mathrm{sys}} \mid \mathbf{y}, a) \approx \sum_{j'} \Big(\frac{g_{j'}}{U_{\mathrm{sys}}}\Big)^2 \Sigma_{j'j'}(n_a\,\delta_{jj'}), \quad (26.24)$$

the other entries unchanged, with $\delta_{jj'}$ one when $j'=j$. The budget term $V_j = (g_j/U_{\mathrm{sys}})^2\Sigma_{jj}$ of equation (26.23) is the uncertainty importance of the group's record, the part of the band it leaves. Its rate of fall per added check, at $n_a = 0$, with $\Sigma_{jj} = 1/(n_j q_j(1-q_j))$ at $\hat{\mathbf{z}}$, equation (14.7), is

$$-\frac{\partial}{\partial n_a}\,\mathrm{Var}(\ln U_{\mathrm{sys}} \mid \mathbf{y}, a)\Big|_{n_a=0} = \Big(\frac{g_j}{U_{\mathrm{sys}}}\Big)^2 \Sigma_{jj}^2\, q_j(1-q_j) = \frac{V_j}{n_j}, \quad (26.25)$$

the budget divided by the number of checks already in the record: one more check removes the fraction $1/n_j$ of what that record leaves. The Fussell–Vesely importance ranks the units by what their failures carry; the budget $V_j = (1-q_j)^2(\sum_{i\in j} I_{\mathrm{FV},i})^2\Sigma_{jj}$ ranks the groups' records by what their units carry times how poorly the record knows it, and the rate divides that by the record's length. A unit can carry much of the top event and its record little of the band, if the record is long and holds many failures.

Table 26.9 gives the budget of equation (26.23) by entry. Each row is an entry of the state; the columns give its units, the sum of their shares, the logarithmic gradient $(1-q_j)\sum_{i\in j} I_{\mathrm{FV},i}$, the diagonal entry of $\mathbf{\Sigma}$, the budget term V_j and its share of the variance of $\ln U_{\mathrm{sys}}$, the length of the record and the rate (26.25) at which one more check removes variance.

Table 26.9: The budget of the variance of $\ln U_{\text{sys}}$ by entry of the state, equation (26.23), at $\hat{\mathbf{z}}$, and the variance one more check of each group removes, equation (26.25).

Entry	Units	$\sum_{i\in j} I_{\text{FV},i}$	$(1-q_j)\sum I_{\text{FV},i}$	Σ_{jj}	Budget V_j	Share	n_j	Per check V_j/n_j
ξ_C	C1–C4	1.616	1.568	0.3436	0.845	97.1 %	100	8.4×10^{-3}
ξ_P	P1, P2	0.0316	0.0313	0.2525	2.5×10^{-4}	0.03 %	400	6.2×10^{-7}
ξ_K	K	0.158	0.158	1.0010	0.0251	2.9 %	1000	2.5×10^{-5}
total					0.870	100 %		

Figure 26.5 sets the two rankings side by side. The horizontal axis lists the three groups; for each, the left bar is the group's share of the variance of $\ln U_{\text{sys}}$ and the right bar the Fussell–Vesely importance of one of its units, both on the same vertical scale from zero to one, with the value printed above each bar.

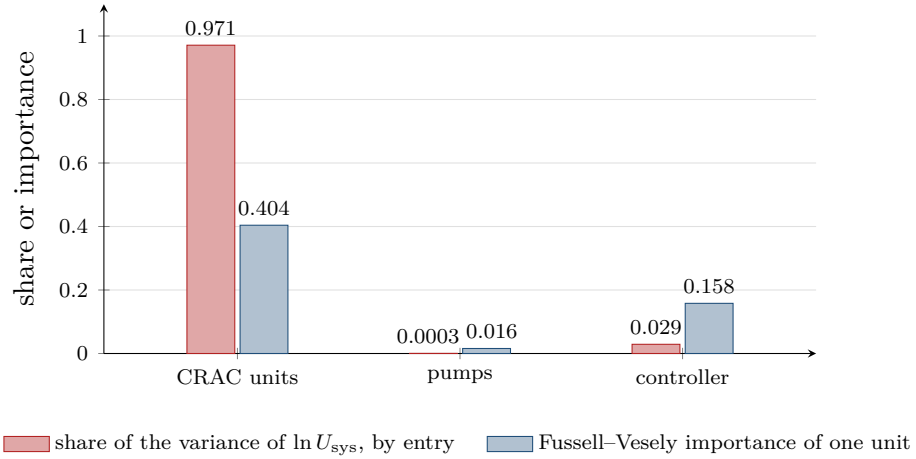


Figure 26.5: The uncertainty importance of each group's record, its share of the variance of $\ln U_{\text{sys}}$ (Table 26.9), against the Fussell–Vesely importance of one of its units (Table 26.2), at $\hat{\mathbf{z}}$.

The budget and the shares order the groups alike, CRAC units, controller, pumps, with different sizes: a CRAC unit carries two and a half times the controller's share, the CRAC record thirty-four times the controller's record's part of the band, and per check 337 times as much. The Birnbaum ranking swaps the first two.

26.6.2 The preposterior covariance

Under the Laplace posterior the spread of d after a record has a closed form. A record of n_a checks of group j is one more count row on entry j with Jacobian row $\mathbf{J}_a = \sqrt{n_a q_j (1-q_j)}\,\mathbf{e}_j^{\mathsf{T}}$ at $\hat{\mathbf{z}}$, and the preposterior covariance of equation (2.17) and the spread of the ranking's logarithm are

$$\boldsymbol{\Sigma}_a = \left(\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a\right)^{-1}, \qquad \operatorname{sd}(d \mid \mathbf{y}, a) \approx \sqrt{\mathbf{g}_d^{\mathsf{T}}\boldsymbol{\Sigma}_a\mathbf{g}_d}, \tag{26.26}$$

with $\mathbf{g}_d$ of equation (26.20). A hundred CRAC checks take Σ_{CC} from 0.344 to 0.172 and four hundred to 0.069; a hundred controller checks take Σ_{KK} from 1.001 to 0.910 and a thousand to 0.500. Through $\mathbf{g}_d$, whose two entries weigh the CRAC logit at 1.89 and the controller's at one, the standard deviation of d falls from 1.49 to 1.27 and 1.12 for the CRAC records and to 1.46 and 1.31 for the controller records.

Table 26.10 gives the result. Each row is a candidate record; the columns give the entry it feeds, the information it adds, the Laplace standard deviation of d after it, the probability that

it settles the order, with the share of that probability in which a CRAC unit is confirmed first, and the expected absolute change of the probability of the order. The exact columns sum the quadrature of equation (26.21) over the Beta–binomial predictive of the counts.

Table 26.10: The preposterior of $\Pr(I_{\mathrm{FV,K}} > I_{\mathrm{FV,C}} \mid \mathbf{y})$ for four candidate records. A CRAC check is one check of one unit. Added information and Laplace standard deviation from equation (26.26); the other columns from the state refitted after each count, weighted by the Laplace predictive, and in parentheses exact, by the quadrature over each count. Settled: the probability after the record is at least 0.95 or at most 0.05.

Record	Entry fed	Added information	Laplace sd of d	Settled	Of it, CRAC unit first	Expected change of the probability
present	–	–	1.49	–	–	–
100 more CRAC checks	ξ_C	2.91	1.27	0.13 (0.14)	all (all)	0.15 (0.15)
400 more CRAC checks	ξ_C	11.64	1.12	0.19 (0.27)	all (all)	0.19 (0.19)
100 more controller checks	ξ_K	0.10	1.46	0.00 (0.00)	–	0.04 (0.03)
1000 more controller checks	ξ_K	1.00	1.31	0.00 (0.00)	–	0.12 (0.11)

A hundred more CRAC checks settle the order with probability 0.13, each time in favor of the CRAC unit: if six or more of them find a unit down, the probability that the controller's share is the larger falls below 0.08, and below 0.05 from seven; if none does, it rises only to 0.61 (exactly 0.14; below 0.07 from five and below 0.05 from six; 0.56). Four hundred settle it with probability 0.19 (0.27 exactly), again only toward the CRAC unit; even none found down in four hundred leaves the probability at 0.93 (0.88 exactly). Controller checks settle nothing. A thousand of them find no unit down with probability 0.38, after which the probability is 0.14, and they move the probability by 0.12 on average without carrying it to either end (exactly 0.50, 0.13 and 0.11); they pass 0.95 with probability 3×10^{-4}, the Laplace predictive's long upper tail, against the exact 2×10^{-7} of twenty-two or more units found down. The reason is the one Chapters 14 and 25 found for their own questions, now with the weights of $\mathbf{g}_d$: a CRAC check adds 0.0291 to Λ_{CC} and a controller check 0.000999 to Λ_{KK}, twenty-nine times less, and the ranking's gradient weighs the CRAC logit at 1.89 against one.

26.6.3 The recommendation

The reported answer whose band the record is to narrow is the order of the controller and a CRAC unit by their shares. Four hundred further checks of the CRAC units, kept as the present record was, are the record that can settle it, with probability 0.19 (0.27 exactly), and that lower the Laplace spread of d from 1.49 to 1.12; a hundred settle it with probability 0.13 (0.14 exactly). Both routes rank the records alike. The record enters as one more count row on ξ_C, and the method returns to step 3 with it. The same record serves the band of the top event, whose budget it carries at 97 % and 337 times the controller's record per check, and the questions of Chapters 14 and 25, which recommended the same record for their own answers. Controller checks remain the record of the controller's own cut set and of its achievement worth, whose level they set; they do not settle the order. Until the record is in, the watch covers both the CRAC units and the controller. The price of the record against the decisions it informs is the subject of a later chapter on the value of information.

26.6.4 What the analysis showed

The first reading took the controller's slope, twelve times a CRAC unit's, as the hall's ranking for every purpose; the analysis keeps it for one purpose and overturns it for the others. The controller is the unit the hall may never let fail: its outage makes the loss of cooling certain, and that order holds whatever the records leave open. But the losses of cooling in this hall run mostly through its CRAC units, each carrying about two fifths of them against the controller's sixth, because a CRAC unit is down thirty times as often and only needs one other to be down with it; and that order is held with about three chances in four, not with certainty, because the controller's record, a thousand checks long, holds a single unit found down. The record that narrows what is open is the CRAC record, which carries nearly all of the band of the loss of cooling and can settle the order of the shares; longer controller records cannot. The engineer carries back to the site three things: keep the controller in service and never schedule its maintenance with the hall loaded; watch the CRAC units and the controller both until the order of their shares is settled; and order the CRAC checks, which serve this question and the band at once.

26.6.5 Limits of the model

The measures are exact for the drawing and the posterior; the drawing and the posterior rest on assumptions the records do not test.

Independence given the unavailabilities. The pivotal decomposition and with it every identity of the chapter assume that the units fail independently given their unavailabilities. A common cause that takes both pumps down, the subject of Chapter 15, couples the pinned values: pinning one pump down then raises the probability that the other is down, the share is no longer the criticality importance, and the pumps' measures rise with the common-cause share. The measures of this chapter are those of the drawing without common causes.

One unavailability per group. The four CRAC units share one entry, and each unit's measure is computed at the shared value. If one unit fails more often than the others, its share and its achievement worth are larger and the others' smaller, and the group totals of the records cannot reveal it; a record kept per unit would.

An outage changes nothing else. The achievement worth takes a unit certainly down with the others' unavailabilities unchanged. A CRAC unit out of service leaves the other three running harder, and if their unavailability rises with the load, the achievement worth of a CRAC unit is understated; the controller's is not, since its outage fails the hall whatever the others do.

Measures of one top event. Every measure is of the loss of cooling. The same units ranked for the expected loss in reais of Chapter 25 keep the ratio measures of the loss-state term, which is proportional to the top event, and change through the degraded terms, which weigh the CRAC units and the pumps by their single outages; a unit's importance belongs to a top event, not to the unit.

Importance is not investment. A measure ranks the units by an effect on the top event, not by the value of changing them. A unit first by every measure may be the dearest to improve, and a design change, such as a fifth CRAC unit, is valued by no unit's measure; the cost of a change set against the reduction of equation (26.17) is the subject of a later chapter on redundancy and reliability investment.

The next chapter turns from ranking units to accepting them: a rule that passes or fails a unit or a plant on a test, with the probability that the rule decides correctly.

26.7 What to retain

Definition 26.1 (Importance measures from three evaluations)**.** For a coherent tree of units independent given their unavailabilities $\mathbf{q}$, with top event $U_{\text{sys}}(\mathbf{q})$ and the pinned values $U_i^+ = U_{\text{sys}}(\mathbf{q} \mid q_i = 1)$ and $U_i^- = U_{\text{sys}}(\mathbf{q} \mid q_i = 0)$, the pivotal decomposition is $U_{\text{sys}} = q_i U_i^+ + (1-q_i)U_i^-$ and the measures of unit i are the Birnbaum importance $I_{\text{B},i} = \partial U_{\text{sys}}/\partial q_i = U_i^+ - U_i^-$, the Fussell–Vesely importance $I_{\text{FV},i} = (U_{\text{sys}} - U_i^-)/U_{\text{sys}} = q_i I_{\text{B},i}/U_{\text{sys}}$, equal to the criticality importance, the achievement worth $\text{RAW}_i = U_i^+/U_{\text{sys}}$ and the reduction worth $\text{RRW}_i = U_{\text{sys}}/U_i^-$, with $\text{RAW}_i - 1 = (1 - q_i)I_{\text{B},i}/U_{\text{sys}}$ and $1 - 1/\text{RRW}_i = I_{\text{FV},i}$. With the unavailabilities carried as logits $\mathbf{z}$, the gradient of the top event is $\mathbf{g} = \mathbf{D}\mathbf{A}^{\mathsf{T}}\mathbf{I}_{\text{B}}$ and its logarithmic entries are $g_j/U_{\text{sys}} = (1 - q_j)\sum_{i \in j} I_{\text{FV},i}$, so that $\text{Var}(\ln U_{\text{sys}} \mid \mathbf{y}) \approx \sum_j (1 - q_j)^2 (\sum_{i \in j} I_{\text{FV},i})^2 \Sigma_{jj}$ for a diagonal $\mathbf{\Sigma}$. Each measure is a derived quantity of the state, with a band, and the order of two units under a measure is an event with a posterior probability.

> **Principle 26.1.** *Name the decision before choosing the measure: the achievement worth for a unit that must never be let fail, the share for where the losses come from and for a proportional improvement, the Birnbaum importance for a fixed decrement of unavailability, and the budget of the top event's band for the record to extend. Report each ranking with the posterior probability that it holds.*

Three conclusions carry beyond the hall. The importance measures are not four computations but one: the top event at the present unavailabilities and with each unit pinned down and up, after which every measure is a difference, a ratio or a share of the same numbers, and two identities tie them so that the slope and the top event determine the rest. The rankings they give differ because they answer different questions, and they differ in firmness: a slope that is the probability that the rest of the system holds is known closely from any record, while a share inherits the spread of the unit's own unavailability, which a record with few failures leaves wide whatever its length. And the band of the top event is built from the same objects: its gradient in the unavailabilities is the Birnbaum vector, its logarithmic gradient in the state the vector of shares scaled by one minus each unavailability, and its budget ranks the records by the part of the band each leaves and the rate at which a check removes it; for this hall that order is the shares' order with different sizes, and the Birnbaum order swaps its first two.

Table 26.11 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 26.11: The terms of reliability importance measures and where each is defined.

Term	Expression	Units	Meaning	Defined in
Pinned values	$U_i^+ = U(q_i = 1)$, $U_i^- = U(q_i = 0)$	–	unit certainly down, never down	26.3.1
Pivotal decomposition	$U = q_i U_i^+ + (1 - q_i)U_i^-$	–	straight line in q_i	26.3.1
Birnbaum importance	$I_{\mathrm{B},i} = U_i^+ - U_i^- = \partial U / \partial q_i$	–	probability unit i is critical	26.3.1
Gradient in the state	$\mathbf{g} = \mathbf{D}\mathbf{A}^{\mathsf{T}}\mathbf{I}_{\mathrm{B}}$	–	units sharing an entry summed	26.3.1
Fussell–Vesely importance	$I_{\mathrm{FV},i} = (U - U_i^-)/U$	–	share removed by a perfect unit	26.3.2
Criticality importance	$I_{\mathrm{CR},i} = q_i I_{\mathrm{B},i}/U = I_{\mathrm{FV},i}$	–	down and critical, given the top event	26.3.2
Cut-set forms of the share	$\Pr(\cup_{c \ni i} E_c)/U$; $\sum_{c \ni i} P_c / S_1$	–	exceed the exact share by the overlaps	26.3.2
Risk achievement worth	$\mathrm{RAW}_i = U_i^+/U$; $1/U$ for a cut set of order one	–	factor of an outage	26.3.3
Risk reduction worth	$\mathrm{RRW}_i = U/U_i^- = 1/(1 - I_{\mathrm{FV},i})$	–	factor of a perfect unit	26.3.3
Reduction of a change	$(q_i - q_i')I_{\mathrm{B},i} = (1 - q_i'/q_i)I_{\mathrm{FV},i}U$	–	absolute or relative improvement	26.3.3
Ranking of the shares	$d = \xi_K - 2\xi_C - \ln 3 - \ln(1 - q_C) + \ln(1 + 3q_C)$	–	$\Pr(d > 0 \mid \mathbf{y})$, one integral	26.3.4
Logarithmic gradient	$g_j/U = (1 - q_j)\sum_{i \in j} I_{\mathrm{FV},i}$	–	the band from the shares	26.3.5
Uncertainty importance	$V_j = (g_j/U)^2 \Sigma_{jj}$; per check V_j/n_j	–	the band a record leaves	26.6.1

26.8 Exercises

Exercise 26.1. *Formulation.* Two units have known unavailabilities $q_1 = 0.01$ and $q_2 = 0.05$, independent. (a) For the two units in series, compute the top event and the four measures of each unit. (b) The same for the two units in parallel. (c) Rank the units under each measure in each arrangement. ((a) $U = 1 - 0.99 \times 0.95 = 0.0595$; $I_{\mathrm{B}} = 1 - q_2 = 0.95$ and $1 - q_1 = 0.99$; $I_{\mathrm{FV}} = 0.0095/0.0595 = 0.160$ and $0.0495/0.0595 = 0.832$; $\mathrm{RAW} = 1/U = 16.8$ for both; $\mathrm{RRW} = 0.0595/0.05 = 1.19$ and $0.0595/0.01 = 5.95$. (b) $U = q_1 q_2 = 5 \times 10^{-4}$; $I_{\mathrm{B}} = q_2 = 0.05$ and $q_1 = 0.01$; $I_{\mathrm{FV}} = 1$ for both, since pinning either up removes the top event; $\mathrm{RAW} = 1/q_1 = 100$ and $1/q_2 = 20$; RRW infinite for both. (c) In series the slope, the share and the reduction worth rank the worse unit, unit 2, first, and the achievement worth cannot separate the two; in parallel the slope and the achievement worth rank the better unit, unit 1, first, since each unit's slope is the other's unavailability, and the share and the reduction worth cannot separate them.)

Exercise 26.2. *Formulation.* A controller A, $q_A = 0.001$, is in series with a pair of pumps B and C in parallel, $q_B = q_C = 0.05$, all independent. (a) Compute the top event and the four measures of A and of B. (b) A maintenance program can halve the unavailability of one unit; which unit should it take, and which measure says so? (c) Which unit should never be taken out of service? ((a) $U = q_A + q_B q_C - q_A q_B q_C = 3.4975 \times 10^{-3}$; $I_{\mathrm{B},A} = 1 - q_B q_C = 0.9975$, $I_{\mathrm{B},B} = (1 - q_A)q_C = 0.0500$; $I_{\mathrm{FV},A} = 0.285$, $I_{\mathrm{FV},B} = 0.714$; $\mathrm{RAW}_A = 1/U = 286$, $\mathrm{RAW}_B = 0.05095/U = 14.6$; $\mathrm{RRW}_A = 1.40$, $\mathrm{RRW}_B = U/q_A = 3.50$. (b) By equation (26.17) halving removes $\frac{1}{2} I_{\mathrm{FV}} U$: 1.25×10^{-3} for B against 5.0×10^{-4} for A, so B, by the share, although A's slope is twenty times B's. (c) A, by its achievement worth, 286 against 14.6.)

Exercise 26.3. *Derivation.* (a) Derive the pivotal decomposition of a top event in the unavailability of one unit from the law of total probability, and from it the Birnbaum importance and its independence of q_i. (b) Derive the identities $\mathrm{RAW}_i - 1 = (1-q_i)I_{\mathrm{B},i}/U$ and $1 - 1/\mathrm{RRW}_i = I_{\mathrm{FV},i}$. (c) Show that under the rare-event approximation the share $(U - U_i^-)/U$ equals the share of the rare-event sum in the cut sets that contain i. ((a) $U = \Pr(\text{top} \mid i \text{ down})q_i + \Pr(\text{top} \mid i \text{ up})(1-q_i) = q_i U_i^+ + (1-q_i)U_i^-$, and the pinned values do not involve q_i, so $\partial U/\partial q_i = U_i^+ - U_i^-$. (b) $U_i^+ - U = (1-q_i)(U_i^+ - U_i^-)$, divide by U; $1 - U_i^-/U = (U - U_i^-)/U$. (c) With $U \approx S_1 = \sum_c P_c$ at every $\mathbf{q}$, pinning i up sets to zero exactly the terms with $i \in c$, so $U - U_i^- \approx \sum_{c \ni i} P_c$ and the share is $\sum_{c \ni i} P_c / S_1$; the exact gap is equation (26.13), of the order of the joint terms.)

Exercise 26.4. *Derivation.* For the hall, (a) derive the closed forms of the Birnbaum importances (26.8) from the pinned forms of equation (26.2); (b) show that the event that the controller's share exceeds a CRAC unit's does not involve the pumps and write it in the logits; (c) show that a CRAC unit's share never exceeds one half. ((a) C1 down leaves the CRAC factor $(1-q_C)^3$, C1 up $(1-q_C)^3 + 3q_C(1-q_C)^2$, a difference of $3q_C(1-q_C)^2$ times $(1-q_P^2)(1-q_K)$; a pump down leaves $1-q_P$, up one, a difference of q_P times $R_{\mathrm{C}}(1-q_K)$; the controller down gives one, up $1 - R_{\mathrm{C}}(1-q_P^2)$. (b) Both shares have the denominator U, and both $q_K I_{\mathrm{B,K}}$ and $q_C I_{\mathrm{B,C}}$ carry $1 - q_P^2$; equation (26.19). (c) $q_C I_{\mathrm{B,C}} = 3q_C^2(1-q_C)^2(1-q_P^2)(1-q_K)$, while $U \geq (6q_C^2 - 8q_C^3 + 3q_C^4)(1-q_P^2)(1-q_K)$, and $3q_C^2(1-q_C)^2 \leq \frac{1}{2}(6q_C^2 - 8q_C^3 + 3q_C^4)$ reduces to $-2q_C^3 + 1.5q_C^4 \leq 0$, true for $q_C \leq 4/3$.)

Exercise 26.5. *Application.* Three pumps of one design form a 2-out-of-3 group, failing when two or more are down, with the shared posterior $q \sim \mathrm{Beta}(2, 98)$. (a) Write the top event and the four measures of one pump as functions of q. (b) Evaluate them at the posterior mean $q = 0.02$. (c) Give their central 95 % bands from the Beta band [0.00246, 0.0550], and say why the share is nearly certain while the other measures are not. ((a) $U = 3q^2 - 2q^3$; $U^+ = 2q - q^2$, $U^- = q^2$; $I_{\mathrm{B}} = 2q(1-q)$, $I_{\mathrm{FV}} = 2(1-q)/(3-2q)$, $\mathrm{RAW} = (2-q)/(q(3-2q))$, $\mathrm{RRW} = 3 - 2q$. (b) $I_{\mathrm{B}} = 0.0392$, $I_{\mathrm{FV}} = 0.662$, RAW = 33.4, RRW = 2.96. (c) Each is monotone in q, so the bands are the images of the Beta band: I_{B} [0.0049, 0.104], I_{FV} [0.654, 0.666], RAW [12.2, 272], RRW [2.89, 2.995]. With all units on one entry the share is a ratio of two polynomials whose leading terms both grow as q^2, so it tends to the structural value $\frac{2}{3}$ and the record's uncertainty nearly cancels; the slope and the achievement worth keep the full spread of q.)

Exercise 26.6. *Application.* In the hall, the owner can either pay for a program that halves the controller's unavailability or one that halves the CRAC units' unavailability, and must also decide which record to extend first. (a) Using equation (26.17) and Table 26.2, compare the plug-in reductions of the two programs. (b) Using Table 26.9, compare the variance of $\ln U_{\mathrm{sys}}$ removed by a hundred further CRAC checks and by a thousand further controller checks, to first order. (c) Name the decision each of the Birnbaum, the share and the budget rankings is right for. ((a) Halving q_K removes $\frac{1}{2} \times 0.158 \times 6.28 \times 10^{-3} = 5.0 \times 10^{-4}$; halving q_C for all four units is a change of the shared entry, $U(0.03) - U(0.015) = 6.28 \times 10^{-3} - 2.42 \times 10^{-3} = 3.86 \times 10^{-3}$, nearly eight times as much, where the first-order sum of the four shares would give $\frac{1}{2} \times 1.616 \times 6.28 \times 10^{-3} = 5.1 \times 10^{-3}$ and overstate it, the top event being quadratic in q_C. (b) Exactly by equation (26.24), a hundred CRAC checks halve Σ_{CC} and remove $0.845/2 = 0.42$; a thousand controller checks halve Σ_{KK} and remove $0.0251/2 = 0.013$; the first-order rates, 0.84 and 0.025, overstate both by the factor two that the curvature in n_a takes back. (c) The Birnbaum importance for a fixed decrement of one unit's unavailability and for the unit to keep in service with the achievement worth; the share for where the losses come from and for a proportional improvement; the budget for the record to extend.)

Chapter 27

Reliability Demonstration and Acceptance Tests

27.1 Accepting a capacity read through a model

The design basis of the two-hall plant of Chapter 17 states the N+1 condition of hall 1 as a capacity: with one unit of the cooling plant out of service, hall 1's heat-removal path must carry the design load of 60 kW at a rise of no more than 12 K above the 12°C supply, that is, its conductance must be at least $G^{*} = 60/12 = 5.0\,\mathrm{kW/K}$. The commissioning engineer must sign that the condition holds, and the record on the desk is the steady afternoon of Chapter 18: five readings at a hall-1 load of about 31 kW, fitted on the reduced state to $\hat{G}_1 = 5.49\,\mathrm{kW/K}$. The first reading of the fit is simple: 5.49 is ten percent above 5.0, so the path has the capacity and N+1 holds.

The plant is being handed over by the contractor who built the cooling, the producer in the language of acceptance sampling, to the data-center operator who bought it, the customer; they are the two parties of the contract. The customer's acceptance clause asks something else: that the condition hold with a posterior probability of at least 0.95, the assurance of Chapter 5, and that a test run before handover be able to show it. A conductance inferred from one temperature sensor at half the design load has a spread, and a test can be too short, at too low a load, or judged on the wrong statistic to show anything. The question of this chapter is

whether a physical capacity read through a model meets a requirement at a stated assurance, how strongly the evidence favours the requirement over its complement, and how an acceptance test of that capacity, its number of readings and its load, is designed so that a good path passes and a bad one fails with stated probabilities.

What changes from the demonstrations of Volume I (Section 27.1.1) is the evidence: readings of a temperature through the physics row of the hall, in place of a count.

Jargon: reliability demonstration test, acceptance test, acceptance sampling by attributes, acceptance sampling by variables, assurance, operating characteristic, power, producer's risk, consumer's risk, held load, load bank, point-estimate rule, shrinkage, two-stage test

A *reliability demonstration test* is a test planned in advance to show that a requirement is met; an *acceptance test* is the same test written into a contract, with the rule that decides it fixed before the test is run. *Acceptance sampling by attributes* judges a lot by a count of defects; *acceptance sampling by variables* judges it by measured values of a characteristic, and this chapter's test is of the second kind. The *assurance* κ is the posterior probability the requirement must reach; the literature of Bayesian test design also uses the word for the probability, before a test, that the test will end accepted, which this book writes $U(a)$ and does not call assurance. The *operating characteristic* (OC) of a test is the probability that it accepts, as a function of the true value of the characteristic; its *power* at a value that meets the requirement is the OC there. The *producer's risk* is the probability of rejecting a good item, one minus the OC at a stated good value; the *consumer's risk* is the probability of accepting a bad one, the OC at a stated bad value. A *held load* is a heat load fixed during a test, here by a *load bank*, a set of electric heaters of known power. A *point-estimate rule* accepts when an estimate clears the requirement, without its spread. *Shrinkage* is the pull of a prior on an estimate toward the prior's center. A *two-stage test* decides after a first set of readings to accept, to reject, or to take a second set.

Key idea

An acceptance rule is a tail of the posterior, $\Pr(G_1 \geq G^* \mid \mathbf{y}) \geq \kappa$, and for a capacity read through a model it is a tail of one entry of the state: the log-conductance g_1 above $\ln G^*$. The same tail, as posterior odds over prior odds, is the Bayes factor of the requirement against its complement, and the rule is a threshold on it. A test is a record not yet taken: its operating characteristic is the rule averaged over the outcomes the test would produce from a stated true state, a preposterior conditioned on the truth, and its design follows from the posterior spread the test's rows would give. For a conductance read through a hall balance that spread has a floor set by the supply sensor's error divided by the temperature rise.

27.1.1 Demonstration of a requirement on a physical quantity

A demonstration test asks the record to establish a requirement, and the companion volume *Reliability Engineering* states the standards that govern such tests in its chapter on reliability standards. Two kinds of demonstration appeared in Volume I. The valve of Chapter 5 was demonstrated by a count, a campaign of demands fixed to pass only with zero failures, equation (5.56), and the generator of Chapter 9 by an exposure, a run of hours without a failure, equation (9.55). In both the evidence was a count, the information grew with the number of trials or the hours, and the only lever of the plan was its length.

A capacity is different. No test counts the conductance of a coil; a test reads temperatures, and the conductance is an entry of the state seen through the physics row of the hall. The row's Jacobian entry on the log-conductance is the temperature rise across the path, so the load of a test enters its Jacobian beside the number of readings. The rule is the one of Chapter 5, equation (5.41), a tail of the posterior against an assurance, and the vocabulary of acceptance sampling, the operating characteristic with its producer's and consumer's risks, applies to it unchanged; what the plant adds is the model through which the characteristic is read.

27.1.2 The plant graph (step 0)

The plant is the two-hall plant of Chapter 17, on its reduced state $\mathbf{z} = (g_1, g_2, Q_1, Q_2, T_s)^{\mathsf{T}}$, equation (17.7), with the hall temperatures and the chiller's heat carried by the balances. The capacity question concerns hall 1's heat-removal path alone. Figure 27.1 draws the graph as the acceptance test sees it. The gray box at the top left is the test load Q_{test}, a pinned source: during the test a load bank of known power holds hall 1's load. The green box below it is hall 1, a free node whose temperature $T_1 = T_s + Q_{\text{test}}\mathrm{e}^{-g_1}/G_{\text{ref}}$ follows from its balance, equation (17.4); the green box at the bottom is the supply header, a free node carrying T_s; and the solid arrow between them is the path with conductance $G_1 = G_{\text{ref}}\mathrm{e}^{g_1}$, the entry the requirement is about. The dashed boxes are the readings of the test: n readings of the hall-1 sensor and one reading of the supply sensor. The box drawn in gray on the right is hall 2, which the test does not read. The dotted box above the path is the question's extension of the graph: the event that the path's conductance is at least G^*, the threshold conductance of Chapter 1, Section 1.3.2, here $G^* = Q_{\text{d}}/(T_{\max} - T_s^0)$ with the design load $Q_{\text{d}} = 60\,\text{kW}$, the limit $T_{\max} = 24°\text{C}$ and the set point $T_s^0 = 12°\text{C}$. G^* is a requirement level, the f^* of Chapter 2, and not a scenario value of the kind the notation table stars from Part V on.

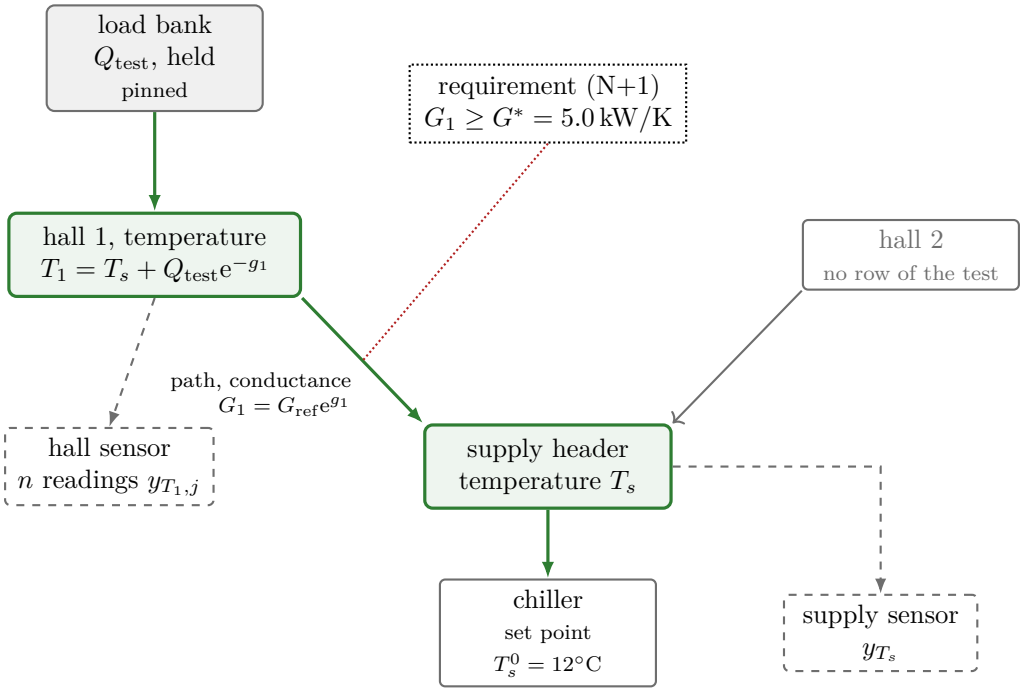


Figure 27.1: The plant graph of hall 1's acceptance test, on the reduced state of Chapter 17. Pinned: the test load Q_{test}, held by a load bank, and $G_{\text{ref}} = 1\,\text{kW/K}$. Free: the log-conductance g_1 of the path and the supply temperature T_s; the hall temperature T_1 is carried by its balance. Observed (dashed): n readings of the hall-1 sensor and one of the supply sensor. Hall 2 (gray) carries no row of the test. Dotted: the requirement, the question's extension of the graph. The priors on g_1 and T_s are rows of the evidence and are not drawn.

27.1.3 The two records

Two records bear on the capacity. The first is the record of Chapter 18, the five readings of Table 17.1 with the three priors of Table 17.2, taken at the plant's present loads; its fit is that chapter's, on the five-entry reduced state, with the MAP state of Table 18.4 and the covariance of equation (18.20), and Chapter 20 gave the full posterior of G_1 by reweighted draws. It is the first case of the acceptance rule.

The second record is the acceptance test, a new record on the reduced state, planned before it is taken. Hall 1's load is held at Q_{test} by the load bank, the plant is let reach steady state, and the hall-1 sensor is read n times, each reading with the error $\sigma_T = 0.5\,\text{K}$ of Table 17.1, independent of the others; the supply sensor is read once, with $\sigma_s = 0.3\,\text{K}$. The handbook prior on g_1 and the set-point prior on T_s of Table 17.2 are the test's priors: the contract judges the test by its own record and the handbook, not by the operating record, whose loads were not held. Table 27.1 lists the rows: each row is one piece of evidence, with its residual, the standard deviation that normalizes it and the entries it acts on. The handbook prior is centered on $\ln 5 = \ln G^*$: before the test, the requirement and its complement are equally probable.

Table 27.1: The rows of the acceptance test on the test state $\mathbf{z} = (g_1, T_s)^\mathsf{T}$, each divided by its standard deviation. $\bar{y}_T$ is the mean of the n hall-1 readings; Q_{test} is pinned; $g^0 = \ln 5$, $T_s^0 = 12°\text{C}$.

Row	Residual	Standard deviation	Acts on
n hall-1 readings, averaged	$T_s + Q_{\text{test}}\text{e}^{-g_1}/G_{\text{ref}} - \bar{y}_T$	$\sigma_T/\sqrt{n} = 0.5/\sqrt{n}\,\text{K}$	g_1, T_s
supply reading	$T_s - y_{T_s}$	$\sigma_s = 0.3\,\text{K}$	T_s
handbook prior	$g_1 - g^0$	$\sigma_g = 0.3$	g_1
set-point prior	$T_s - T_s^0$	$\sigma_0 = 1\,\text{K}$	T_s

The test is planned, so its outcome is not data yet, and its design questions are asked of a stated truth. The truth of the test is a state $\mathbf{z}_\text{t} = (g_\text{t}, T_{s,\text{t}})$ with $G_\text{t} = G_{\text{ref}}\text{e}^{g_\text{t}}$ and the supply at its set point, $T_{s,\text{t}} = 12°\text{C}$. Two values of G_t are named by the parties: 5.5 kW/K, the good path, which the contractor, the producer, wants accepted; and 4.75 kW/K, five percent below the requirement, which the operator, the consumer, wants rejected. The loads considered run from the present 30 kW to the design load 60 kW; a test above the design load would put a path at the requirement above its 24°C limit.

27.1.4 The questions (step 1)

Table 27.2 lists the questions. The first column is the question in engineering language; the second the general form of its category from Chapter 2; the third its mathematical form for the plant, on the record of Chapter 18 for the first two and on the test record for the others; the fourth its category; and the last where it is answered. The first is the acceptance rule; the second compares the two halves of the prior, the requirement met and not met; the third asks how well a test fixes the conductance and why more readings stop helping; the fourth and the fifth are the test's operating characteristic and its plan; the sixth is the common mistake of accepting on a point estimate. The atlas of Appendix A lists these questions.

27.1.5 The evidence as rows (step 2)

The record of Chapter 18 enters as its eight rows on the five-entry reduced state, equation (17.15), with the Jacobian $\mathbf{J} = \mathbf{J}_\text{e}\mathbf{N}$ of equation (17.18): the projection of Chapter 17 of the reading and prior rows onto the null directions of the physics rows, whose covariance is the free-entry marginal of the full state's in the limit $\sigma_p \to 0$, equation (17.19). The acceptance rule reads the first entry of that state.

The test is a new record on the same reduced state, and the same projection applies: its hall-1 rows are the reading row of equation (17.16) with the load pinned at Q_{test}. The test's rows touch two entries only, g_1 and T_s; the columns of g_2, Q_2 and of the pinned load carry none

Table 27.2: The questions asked of the capacity of hall 1's path. $\mathbf{y}$ is the record of Chapter 18; $\tilde{\mathbf{y}}_a = (\bar{y}_T, y_{T_s})$ the outcome of a test $a = (n, Q_{\text{test}})$; $\kappa = 0.95$; OC the operating characteristic; $\mathbf{z}_{\text{t}}$ a stated true state; M_1, M_0 the hypotheses that the requirement is met and not met.

Engineering question	General form	Mathematical form	Category	Answered in
Does N+1 hold?	threshold, eq. (2.12)	$\Pr(G_1 \geq G^* \mid \mathbf{y}) = \Pr(g_1 \geq \ln G^* \mid \mathbf{y}) \geq \kappa$	prediction under a threshold (the acceptance rule)	27.3.1
Which hypothesis does the evidence favour, N+1 holds or not?	Bayes factor, eq. (2.14)	$B_{10} = \dfrac{\Pr(M_1 \mid \mathbf{y})/\Pr(M_0 \mid \mathbf{y})}{\Pr(M_1)/\Pr(M_0)}$	hypothesis comparison	27.3.2
How well does a test fix the conductance, and why do more readings stop helping?	$\mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y})$, eq. (2.10)	$\mathrm{sd}(g_1 \mid \tilde{\mathbf{y}}_a) = \sqrt{\Sigma_{11}}$ against n and Q_{test}; its limit $n \to \infty$	estimation (the budget of the spread)	27.3.3
If the path really has 5.5 kW/K, how likely is the test to accept it? And if 4.75?	design, eq. (2.16), conditioned on a truth	$\mathrm{OC}(G_{\text{t}}; a) = \Pr(\text{accept} \mid \mathbf{z}_{\text{t}}, a)$	design (the operating characteristic)	27.3.4
How long a test, at what load?	design, eq. (2.16)	smallest n at each Q_{test} with $\mathrm{OC}(5.5) \geq 0.9$ and $\mathrm{OC}(4.75) \leq 0.01$	design (test planning)	27.3.5
What does accepting on a point estimate get wrong?	threshold at $\kappa = 0.5$	OC of $\hat{G}_1 \geq G^*$ against the rule on the tail	prediction under a threshold (the common mistake)	27.3.6

of them, so the posterior of the test separates and the test's state is the two-entry vector

$$\mathbf{z} = \begin{bmatrix} g_1 \\ T_s \end{bmatrix}, \tag{27.1}$$

with g_1 dimensionless and T_s in °C. Which row acts on which entry is a property of the graph, and the Jacobian of the four rows of Table 27.1 can be nonzero only where marked $*$:

$$\mathbf{J} = \begin{bmatrix} * & * \\ 0 & * \\ * & 0 \\ 0 & * \end{bmatrix} \begin{matrix} \text{hall-1 readings} \\ \text{supply reading} \\ \text{handbook prior} \\ \text{set-point prior} \end{matrix}. \tag{27.2}$$

The conductance has no reading of its own: it is reached by the hall-1 row alone, which marks both columns, so the readings see g_1 only together with the supply temperature, and the supply column, with three marks, is what separates the two. The design questions add nothing to this pattern: a test's design changes the values in the first row, through n and Q_{test}, not where they lie. All four rows are Gaussian.

The modeling choices. Three choices make the rows of the two records, and each buys something and can fail. The conductance of the record, taken in normal operation, is taken to be the conductance of the path in the N+1 configuration the clause names: this lets the record be the first case of the rule, and it fails if taking a unit out changes the path's air or water flow, which only a test in that configuration measures (Section 27.6.4). The test load is pinned, held by a calibrated load bank: this keeps the load out of the state, and it fails if the load is only read by the 1.5 kW meter, whose error then enters the rise (Section 27.4.3). The n readings are

independent given the state: this lets them average into one row with $\sigma_T/\sqrt{n}$, and it fails for an error that drifts slowly over the test, which averaging does not remove (Section 27.6.4).

The rows of the two records. The record of Chapter 18 has the eight rows of equation (17.15), each a Gaussian term $\mathcal{L}_k = \frac{1}{2}r_k^2$ of the objective, with r_1 the hall-1 reading row of equation (17.16). The test has the four rows of Table 27.1. Its n hall-1 readings $y_{T_1,j}$, each a row $r_j = (T_s + Q_{\text{test}}\mathrm{e}^{-g_1}/G_{\text{ref}} - y_{T_1,j})/\sigma_T$, put the terms

$$\sum_{j=1}^{n} \tfrac{1}{2} r_j^2 = \frac{n}{2\sigma_T^2}\Big(T_s + \frac{Q_{\text{test}}\mathrm{e}^{-g_1}}{G_{\text{ref}}} - \bar{y}_T\Big)^2 + \frac{1}{2\sigma_T^2}\sum_{j=1}^{n}\big(y_{T_1,j} - \bar{y}_T\big)^2$$

in the objective, with $\bar{y}_T = \frac{1}{n}\sum_j y_{T_1,j}$; the identity follows by writing $y_{T_1,j} = \bar{y}_T + (y_{T_1,j} - \bar{y}_T)$ inside each square, the cross terms summing to zero. The second sum does not depend on the state, so the n rows are one row on the mean with the standard deviation $\sigma_T/\sqrt{n}$, the first row of Table 27.1, as Chapter 23 averaged a record of days. The residual vector of the test is

$$\mathbf{r}(\mathbf{z}) = \begin{bmatrix} (T_s + Q_{\text{test}}\mathrm{e}^{-g_1}/G_{\text{ref}} - \bar{y}_T)\sqrt{n}/\sigma_T \\ (T_s - y_{T_s})/\sigma_s \\ (g_1 - g^0)/\sigma_g \\ (T_s - T_s^0)/\sigma_0 \end{bmatrix}, \qquad \mathcal{L}(\mathbf{z}) = \tfrac{1}{2}\,\mathbf{r}(\mathbf{z})^{\mathsf{T}}\mathbf{r}(\mathbf{z}), \tag{27.3}$$

the reduced-state rows of Chapter 17 for a test in which hall 1's load is pinned and hall 2 is not read; the physics row of hall 1 is carried by the projection, and Section 27.4.3 checks the limit $\sigma_p \to 0$ on the test's full state.

The Jacobian of the test. The rows of Table 27.1 at the prior point $\mathbf{z}^\circ = (g^0, T_s^0)$ have the Jacobian of the pattern (27.2), with the derivative of $Q_{\text{test}}\mathrm{e}^{-g_1}/G_{\text{ref}}$ in g_1 equal to minus the rise at the requirement,

$$\mathbf{J}(\mathbf{z}^\circ) = \begin{bmatrix} -\Delta T^*\sqrt{n}/\sigma_T & \sqrt{n}/\sigma_T \\ 0 & 1/\sigma_s \\ 1/\sigma_g & 0 \\ 0 & 1/\sigma_0 \end{bmatrix}, \qquad \Delta T^* = \frac{Q_{\text{test}}}{G^*}, \tag{27.4}$$

a 4×2 matrix, with ΔT^* in kelvin: the temperature rise the path would have at the test load if it sat exactly at the requirement level G^*, whose star it carries; it is not a scenario value of the kind starred from Part V on.

27.2 The posterior of the state (step 3)

Every question of Table 27.2 is read from a posterior of the log-conductance g_1: the posterior that the record of Chapter 18 gives it, or the posterior a test would give it, before the test is run. This section derives those posteriors from the rows of step 2, in their state-vector form, and computes them for the record and for a test; Section 27.3 reads every answer from them.

27.2.1 The posterior of the two records

The record of Chapter 18. Its MAP state $\hat{\mathbf{z}}$, its information $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ and its covariance $\boldsymbol{\Sigma}$ are those of equation (18.20), from the Gauss–Newton iterations of that chapter; the questions use the first entry, $\hat{g}_1$ with $\sqrt{\Sigma_{11}}$. Its full posterior is the one Chapter 20 obtained by reweighting $N = 200\,000$ draws of $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ with the weights of equation (20.10), from the PCG64 generator seeded with 17, numpy's `default_rng(17)`, the same draws and the same weights used here.

The test. The state is the two-entry vector (27.1) and the rows are equation (27.3). The Jacobian at any state $\mathbf{z}$ is the matrix (27.4) with the rise at the requirement ΔT^* replaced by the rise of the state, $Q_{\text{test}}\mathrm{e}^{-g_1}/G_{\text{ref}}$; only its first row changes with the state. The fit solves the normal equations of Chapter 3,

$$(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^{\mathsf{T}}\mathbf{r}, \qquad \mathbf{z}_{j+1} = \mathbf{z}_j + \Delta\mathbf{z}, \tag{27.5}$$

starting at the prior point $\mathbf{z}^\circ = (g^0, T_s^0)$, with the stopping rule of Chapter 18, and the Laplace covariance is $\mathbf{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$ at $\hat{\mathbf{z}}$. The hall-1 row is the one nonlinear row, through e^{-g_1}; its curvature term, $r_1\,\partial^2 r_1/\partial\mathbf{z}\,\partial\mathbf{z}^{\mathsf{T}}$, has the single entry $r_1\,(Q_{\text{test}}\mathrm{e}^{-g_1}/G_{\text{ref}})\sqrt{n}/\sigma_T$ in the $(1,1)$ position, which Gauss–Newton drops and which is small when the fitted reading row is small.

The exact tail of a test. The test admits its exact posterior of g_1 in one integral. Every row of equation (27.3) is linear in T_s, and the derivatives on T_s do not depend on the state, so for each g_1 the objective is a quadratic in T_s with the constant curvature $\Lambda_{22} = n/\sigma_T^2 + 1/\sigma_c^2$. Its minimum over T_s is at

$$\hat{T}_s(g_1) = \frac{n\,T_{s,\mathrm{h}}(g_1)/\sigma_T^2 + y_{T_s}/\sigma_s^2 + T_s^0/\sigma_0^2}{n/\sigma_T^2 + 1/\sigma_c^2}, \qquad T_{s,\mathrm{h}}(g_1) = \bar{y}_T - \frac{Q_{\text{test}}\mathrm{e}^{-g_1}}{G_{\text{ref}}},$$

the precision-weighted mean of the three values of the supply that the hall reading, the supply reading and the set point give, the first of them, $T_{s,\mathrm{h}}(g_1)$, the supply the hall reading implies at g_1. Integrating the Gaussian in T_s gives $\int \mathrm{e}^{-\mathcal{L}(g_1,T_s)}\,\mathrm{d}T_s = \mathrm{e}^{-\mathcal{L}(g_1,\hat{T}_s(g_1))}\sqrt{2\pi/\Lambda_{22}}$, and the square root is a constant, so the marginal of g_1 is $\mathrm{e}^{-\mathcal{L}(g_1,\hat{T}_s(g_1))}$ normalized, and

$$\Pr(G_1 \geq G^* \mid \tilde{\mathbf{y}}_a) = \frac{\int_{g^*}^{\infty} \mathrm{e}^{-\mathcal{L}(g,\hat{T}_s(g))}\,\mathrm{d}g}{\int_{-\infty}^{\infty} \mathrm{e}^{-\mathcal{L}(g,\hat{T}_s(g))}\,\mathrm{d}g}, \tag{27.6}$$

a ratio of one-dimensional integrals, evaluated by quadrature on $g^* \pm 0.7$. It is the full posterior of a test; the tail from the posterior of the state of the same record is the normal tail of its first entry, equation (27.8) with the test's $\hat{g}_1$ and Σ_{11}, and the exact tail gives the value beside it.

27.2.2 The posterior of the record and of a test

The inputs are the fit of Chapter 18 with the reweighted draws of Chapter 20, and the spreads of the test's rows. The record's fit is Chapter 18's, at the unrounded hall-1 reading 17.654545°C: $\hat{g}_1 = 1.70297$, $\hat{G}_1 = 5.490\,\mathrm{kW/K}$, $\Sigma_{11} = 0.011382$ and $\sqrt{\Sigma_{11}} = 0.10669$, from equation (18.20). The 200 000 draws of Chapter 20 carry log weights from -5.9 to 3.4 and the effective sample size 182 933.

For a test, the Jacobian (27.4) at the requirement, for $n = 16$ readings at $Q_{\text{test}} = 60\,\mathrm{kW}$, has the rise $\Delta T^* = 60/5 = 12\,\mathrm{K}$ and $\sqrt{n}/\sigma_T = 8\,\mathrm{K}^{-1}$:

$$\mathbf{J}(\mathbf{z}^\circ) = \begin{bmatrix} -96 & 8 \\ 0 & 3.3333 \\ 3.3333 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \mathbf{\Lambda} = \begin{bmatrix} 9227.1 & -768.0 \\ -768.0 & 76.111 \end{bmatrix},$$

$$\mathbf{\Sigma} = \begin{bmatrix} 6.7677 \times 10^{-4} & 6.8290 \times 10^{-3} \\ 6.8290 \times 10^{-3} & 0.082047 \end{bmatrix},$$

so $\sqrt{\Sigma_{11}} = 0.02601$ in g_1, $\sqrt{\Sigma_{22}} = 0.2864\,\mathrm{K}$ in T_s, and $\mathrm{corr}(g_1, T_s) = 0.916$. The supply's combined spread is $\sigma_c = (1/0.3^2 + 1/1^2)^{-1/2} = 0.2873\,\mathrm{K}$.

Every answer of step 4 is computed from these posteriors of the state: the record's from its first entry, $\hat{g}_1$ and Σ_{11}, and a test's from the test's $\mathbf{\Sigma}$ at the requirement. The full posterior of the record, by the reweighted draws, and the exact tail (27.6) of a test give the values set beside them.

27.3 The answers (step 4)

The questions of Table 27.2 fall into four categories: a prediction under a threshold (the first and the sixth), a hypothesis comparison (the second), an estimation (the third) and two design questions (the fourth and the fifth). Each is answered from the posteriors of Section 27.2: its engineering question, the general form of its category, the form it takes for the path, the quantities of the posterior it uses, and the path's numbers. The questions name the requirement $G^* = 5.0\,\mathrm{kW/K}$, $g^* = \ln 5 = 1.60944$, and the assurance $\kappa = 0.95$ with $z_\kappa = 1.64485$. The chapter reports the record's tail and Bayes factor from the posterior of the state, the Laplace normal of g_1, and a test's characteristic by the planning route, the same normal on the test's posterior; the full posterior of the record and simulated tests judged by the exact tail give the values set beside them in the same answer, with the gap.

27.3.1 The acceptance rule on the record

Engineering question. Does N+1 hold? The design basis asks for $G_1 \geq G^* = 5.0\,\mathrm{kW/K}$, and the acceptance clause asks that this hold with posterior probability at least $\kappa = 0.95$.

General form. The threshold form of Chapter 2, equation (2.12), the acceptance rule of Chapter 5, equation (5.41):

$$\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) \geq f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \qquad \text{accept if } \Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y}) \geq \kappa.$$

Specialized form. The question is equation (2.12) with $f(\mathbf{z}) = G_1 = G_{\mathrm{ref}}\mathrm{e}^{g_1}$, $f^* = G^* = 5.0\,\mathrm{kW/K}$, $\kappa = 0.95$, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of the record of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ whose first entry is $g_1 \mid \mathbf{y} \approx \mathcal{N}(\hat{g}_1, \Sigma_{11})$. One choice enters here. The N+1 condition is written as a requirement on the conductance alone, the rise at the design load against a fixed 12 K: this makes the event an event on one entry, and it fails if the supply temperature rises with the load, which would put T_s into the event as in Chapter 20 (Section 27.6.4). The tail is reported from the posterior of the state, the Laplace normal; the full posterior by reweighted draws gives the value beside it.

The exponential is increasing, so the event maps exactly onto the first entry of the state, as in Chapter 1,

$$G_1 \geq G^* \iff g_1 \geq g^* = \ln \frac{G^*}{G_{\mathrm{ref}}} = \ln 5, \qquad \Pr(G_1 \geq G^* \mid \mathbf{y}) = \int \mathbf{1}\big[g_1 \geq g^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \tag{27.7}$$

and the other four entries integrate out. Under the Laplace posterior of the reduced state, $\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$, the marginal of the first entry is $\mathcal{N}(\hat{g}_1, \Sigma_{11})$, so the tail is a normal tail of an entry read off with its spread,

$$\Pr(G_1 \geq G^* \mid \mathbf{y}) \approx \Phi\!\left(\frac{\hat{g}_1 - g^*}{\sqrt{\Sigma_{11}}}\right), \tag{27.8}$$

equation (4.54) applied to one entry; only the normal is an approximation, since the threshold maps exactly. Under the full posterior the tail is the self-normalized reweighted average of the indicator over draws of the Laplace normal, equation (20.11) with $\psi(\mathbf{z}) = \mathbf{1}[g_1 \geq g^*]$,

$$\Pr(G_1 \geq G^* \mid \mathbf{y}) \approx \sum_{j=1}^{N} \bar{\zeta}_j\, \mathbf{1}\big[g_1^{(j)} \geq g^*\big], \tag{27.9}$$

with the normalized weights $\bar{\zeta}_j$ of equation (20.10) and the standard error of equation (20.12). The decision can also be read as a one-sided bound, as in Chapter 5, equation (5.45): with $G_{1-\kappa}$

the $(1-\kappa)$ quantile of G_1, $\Pr(G_1 \le G_{1-\kappa} \mid \mathbf{y}) = 1-\kappa$, the rule holds exactly when $G^* \le G_{1-\kappa}$. Under the Laplace normal the quantile is the MAP entry moved down by $z_\kappa = \Phi^{-1}(\kappa)$ standard deviations,

$$G_{1-\kappa} \approx G_{\text{ref}} \exp\bigl(\hat{g}_1 - z_\kappa \sqrt{\Sigma_{11}}\bigr), \qquad \text{accept} \iff \hat{g}_1 - z_\kappa \sqrt{\Sigma_{11}} \ge g^*, \tag{27.10}$$

the form of the rule every test below is judged by; under the full posterior it is the weighted quantile of the draws.

Building blocks. The MAP entry $\hat{g}_1$ and Σ_{11} of the record (Section 27.2.1); the threshold $g^* = \ln 5$; Φ and z_κ. For the values beside them, the 200 000 reweighted draws of Chapter 20 with their weights.

The acceptance decision on the record. By the reported route, the Laplace normal of the record's first entry, equation (27.8) has the standardized distance and the tail

$$\frac{\hat{g}_1 - g^*}{\sqrt{\Sigma_{11}}} = \frac{1.70297 - 1.60944}{0.10669} = 0.8767, \qquad \Pr(G_1 \ge 5.0\,\text{kW/K} \mid \mathbf{y}) \approx \Phi(0.8767) = 0.810,$$

and by equation (27.10) the 0.05 quantile of G_1 is $G_{0.05} = \exp(1.70297 - 1.64485 \times 0.10669) = 4.607\,\text{kW/K}$. The tail is below $\kappa = 0.95$ and the bound below G^*: N+1 is not accepted on the record. The value beside it comes from the full posterior: equation (27.9) over the reweighted draws gives

$$\Pr(G_1 \ge 5.0\,\text{kW/K} \mid \mathbf{y}) = 0.8395 \pm 0.0008,$$

with the standard error of equation (20.12), and the weighted 0.05 quantile $G_{0.05} = 4.683\,\text{kW/K}$. The gap is 0.030 in probability, thirty-eight standard errors of the reweighted tail, and $0.076\,\text{kW/K}$ in the quantile, both with the state route lower. It errs low for the reason Chapter 20 traced to the hall-1 reading row: the full posterior of g_1 is skewed toward large conductances, and the symmetric normal of the log conductance puts more of its mass below the requirement. The decision is the same by both: neither tail reaches 0.95.

Figure 27.2 draws the posterior. The horizontal axis is the conductance of hall 1's path and the vertical axis its posterior density; the dashed blue curve is the Laplace normal of g_1 carried to G_1, a log-normal density, and the green curve the full posterior, a kernel estimate from the reweighted draws. The red vertical line is the requirement; the green area to its right is the full posterior's probability 0.840; the two short black ticks on the axis are the 0.05 quantiles of the two curves. The full posterior is skewed toward large conductances; the Laplace normal puts 0.19 of its mass below the requirement, the full posterior a sixth.

27.3.2 The Bayes factor of the requirement

Engineering question. Which hypothesis does the evidence favour, N+1 holds or not? The engineer asks how far the record has moved the case for the requirement, apart from where the case started.

General form. The Bayes factor of Chapter 2, equation (2.14), the ratio of the evidences of two models fitted to the same evidence:

$$B_{10} = \frac{p(\mathbf{y} \mid M_1)}{p(\mathbf{y} \mid M_0)}.$$

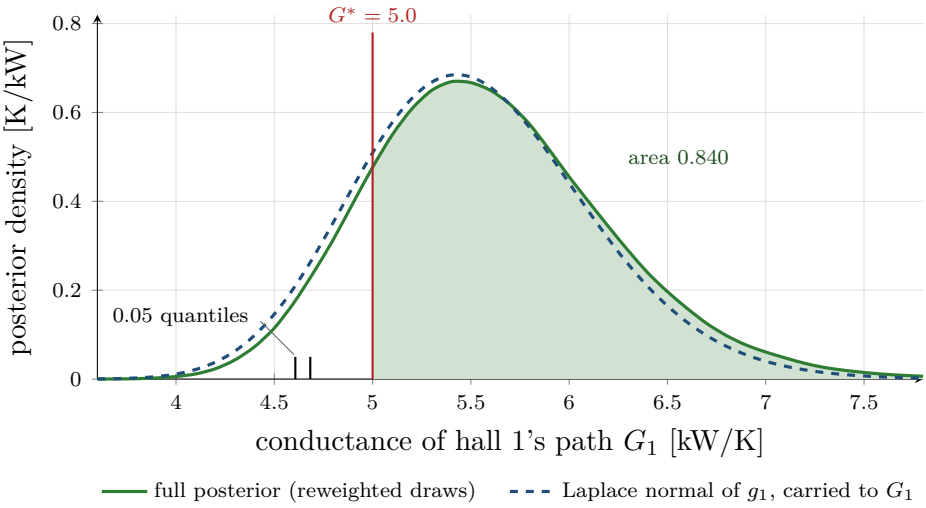


Figure 27.2: Posterior of the conductance of hall 1's path on the record of Chapter 18. Green: the full posterior, a Gaussian kernel estimate (bandwidth 0.08 kW/K) of the 200 000 draws of Chapter 20 reweighted by equation (20.11); shaded, $\Pr(G_1 \geq 5.0 \mid \mathbf{y}) = 0.840$. Blue dashed: the log-normal density of $G_1 = \mathrm{e}^{g_1}$ with $g_1 \sim \mathcal{N}(1.70297,\ 0.10669^2)$, the posterior of the state, whose mass above 5.0 is the reported 0.810. Ticks: the 0.05 quantiles, 4.683 (full) and 4.607 kW/K (Laplace).

Specialized form. The question is equation (2.14) with M_1 the plant with $g_1 \geq g^*$ and M_0 the plant with $g_1 < g^*$, $\mathbf{y}$ the record, and each evidence reduced below to a tail of $p(\mathbf{z} \mid \mathbf{y})$, the posterior of step 3, the Laplace normal of g_1. One choice enters here: the two models are hypotheses, the two halves of one drawing cut at the requirement, with the prior of each the handbook prior restricted to its half. This makes the comparison a statement about the event and nothing else, and it means the factor depends on where the cut lies against the prior; a prior not centered on G^* weighs the halves unequally and changes the factor (Section 27.4.3).

The hypothesis M_1 is the plant with $g_1 \geq g^*$ and M_0 the plant with $g_1 < g^*$; both have the rows and the priors of the record, and the prior of M_1 is the prior of the state renormalized on its half,

$$p(\mathbf{z} \mid M_1) = \frac{p(\mathbf{z})\,\mathbf{1}[g_1 \geq g^*]}{\Pr(g_1 \geq g^*)}.$$

The evidence of M_1 is the probability of the record with the state integrated out over that prior, equation (24.6):

$$p(\mathbf{y} \mid M_1) = \int p(\mathbf{y} \mid \mathbf{z})\, \frac{p(\mathbf{z})\,\mathbf{1}[g_1 \geq g^*]}{\Pr(g_1 \geq g^*)}\, \mathrm{d}\mathbf{z}.$$

By Bayes' theorem $p(\mathbf{y} \mid \mathbf{z})\, p(\mathbf{z}) = p(\mathbf{z} \mid \mathbf{y})\, p(\mathbf{y})$, with $p(\mathbf{y})$ the evidence of the whole model, so the integral is a tail of the posterior:

$$p(\mathbf{y} \mid M_1) = \frac{p(\mathbf{y})}{\Pr(g_1 \geq g^*)} \int \mathbf{1}[g_1 \geq g^*]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z} = p(\mathbf{y})\, \frac{\Pr(g_1 \geq g^* \mid \mathbf{y})}{\Pr(g_1 \geq g^*)}.$$

The same steps on the other half give $p(\mathbf{y} \mid M_0) = p(\mathbf{y})\ \Pr(g_1 < g^* \mid \mathbf{y}) / \Pr(g_1 < g^*)$. Dividing, $p(\mathbf{y})$ cancels,

$$B_{10} = \frac{\Pr(G_1 \geq G^* \mid \mathbf{y}) / \Pr(G_1 < G^* \mid \mathbf{y})}{\Pr(G_1 \geq G^*) / \Pr(G_1 < G^*)}, \qquad \ln B_{10} = \ln \frac{P}{1-P} - \ln \frac{P_0}{1-P_0}, \tag{27.11}$$

with $P = \Pr(G_1 \geq G^* \mid \mathbf{y})$ the tail of the previous question and $P_0 = \Pr(G_1 \geq G^*)$ its prior value: the Bayes factor of the requirement is its posterior odds over its prior odds, equation (24.8)

read backward. The handbook prior is the normal $\mathcal{N}(\ln 5, 0.3^2)$ centered on g^*, so $P_0 = \frac{1}{2}$, the prior odds are one, and B_{10} equals the posterior odds. The rule of the previous question is a threshold on the factor: $P \geq \kappa$ is $P/(1-P) \geq \kappa/(1-\kappa)$, so

$$\text{accept} \iff B_{10} \geq \frac{\kappa/(1-\kappa)}{P_0/(1-P_0)}, \tag{27.12}$$

a factor read on the scale of Chapter 24.

Two statements are thus available, and they answer different questions. The Bayes factor says how much the record moved the odds of the requirement from the prior's; it is not free of the prior, since the cut renormalizes the prior on each half and the likelihood is averaged over those halves, so the same record moves the odds by different factors from different priors. The tail at κ says whether the odds after the record are high enough to sign. With the prior centered on G^* the two are monotone in each other; with a prior leaning against the requirement, a record could favour N+1 by a large factor and still leave its probability low. The comparison also differs from those of Chapter 24. There the drawings differed in their rows and entries, a hall on the shared loop or on a branch of its own, each evidence integrated its own state over its own prior, and an Occam factor penalized the prior range a drawing did not use. Here the two hypotheses share one drawing and one prior, cut at g^*; their evidences are two pieces of one evidence, their ratio needs no evidence integral, only the posterior tail already computed, and the only Occam term is the prior mass of each half, which cancels when the halves are equal.

Building blocks. The tail P of the previous question by the reported route; the prior tail $P_0 = \frac{1}{2}$; the assurance κ.

The Bayes factor on the record. Table 27.3 sets out equation (27.11). Each row is a state of knowledge: the handbook prior, the record, and the level the rule (27.12) requires; the columns give the probability of the requirement, its odds, and the Bayes factor against the prior with its logarithm. With the reported tail of the previous question, the record's posterior odds are $0.8097/0.1903 = 4.25$, and the prior odds are one, so

$$B_{10} = 4.25, \qquad \ln B_{10} = 1.448 \text{ nats},$$

positive on the scale of Chapter 24: the record favours the hypothesis that N+1 holds by a factor of four. At even prior odds and $\kappa = 0.95$ the rule (27.12) needs $B_{10} \geq 19$, $\ln B_{10} \geq 2.944$ nats, at the edge between positive and strong on that scale; the record supplies 49 % of that logarithm. The value beside it comes from the full posterior: odds $0.8395/0.1605 = 5.23$, so $B_{10} = 5.23$ and $\ln B_{10} = 1.655$ nats, 56 % of the required logarithm. The gap is 0.207 nats, a factor of 1.23, from the same skew as the tail; it leaves the factor positive and short of the 19 the rule needs by both routes.

Table 27.3: The Bayes factor of the requirement $G_1 \geq G^*$ against its complement, equation (27.11), on the record of Chapter 18: from the posterior of the state (reported) and, beside it, from the full posterior. The last row is the threshold of equation (27.12) at $\kappa = 0.95$.

State of knowledge	$\Pr(G_1 \geq G^*)$	Odds	B_{10}	$\ln B_{10}$ [nat]
handbook prior $\mathcal{N}(\ln 5,\ 0.3^2)$	0.5000	1.00	–	–
record of Chapter 18, Laplace (reported)	0.8097	4.25	4.25	1.448
record of Chapter 18, full posterior	0.8395	5.23	5.23	1.655
required for acceptance	0.9500	19.0	19.0	2.944

27.3.3 The spread of a test and its floor

Engineering question. How well does a test of n readings at the load Q_{test} fix the conductance, and why do more readings stop helping?

General form. The variance of a quantity under the posterior, equation (2.10), here of an entry of the state, read off under the Laplace posterior as Σ_{11}:

$$\operatorname{Var}(f(\mathbf{z}) \mid \mathbf{y}) = \int \big(f(\mathbf{z}) - \mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]\big)^2 \, p(\mathbf{z} \mid \mathbf{y}) \, \mathrm{d}\mathbf{z}, \qquad f(\mathbf{z}) = g_1.$$

Specialized form. The question is the variance of equation (2.10) with $f(\mathbf{z}) = g_1$, the evidence $\tilde{\mathbf{y}}_a$ of a test of n readings at the load Q_{test}, and $p(\mathbf{z} \mid \tilde{\mathbf{y}}_a)$ the test's posterior of step 3, the Laplace normal whose covariance is $\boldsymbol{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$, so that the variance is its entry Σ_{11}. One choice enters here: the spread of a plan is read at the requirement, the linearization point $\mathbf{z}^{\circ} = (g^*, T_s^0)$: this is where the decision is close and where the plan must work, and it understates the spread of a path far above G^*, whose smaller rise reads it less well (Section 27.4.2).

At $\mathbf{z}^{\circ}$ the information matrix of the Jacobian (27.4) collects the supply sensor and the set-point prior into one precision $1/\sigma_c^2 = 1/\sigma_s^2 + 1/\sigma_0^2$,

$$\boldsymbol{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J} = \begin{bmatrix} n\,\Delta T^{*2}/\sigma_T^2 + 1/\sigma_g^2 & -n\,\Delta T^*/\sigma_T^2 \\ -n\,\Delta T^*/\sigma_T^2 & n/\sigma_T^2 + 1/\sigma_c^2 \end{bmatrix}; \tag{27.13}$$

The first entry of $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}^{-1}$ is $\Lambda_{22}/\det\boldsymbol{\Lambda}$. The determinant is

$$\det\boldsymbol{\Lambda} = \Big(\frac{n\Delta T^{*2}}{\sigma_T^2} + \frac{1}{\sigma_g^2}\Big)\Big(\frac{n}{\sigma_T^2} + \frac{1}{\sigma_c^2}\Big) - \frac{n^2\Delta T^{*2}}{\sigma_T^4} = \frac{n\,\Delta T^{*2}}{\sigma_T^2\,\sigma_c^2} + \frac{1}{\sigma_g^2}\Big(\frac{n}{\sigma_T^2} + \frac{1}{\sigma_c^2}\Big),$$

and dividing it by $\Lambda_{22} = n/\sigma_T^2 + 1/\sigma_c^2$,

$$\frac{1}{\Sigma_{11}} = \frac{1}{\sigma_g^2} + \frac{\Delta T^{*2}}{\sigma_T^2/n + \sigma_c^2}, \qquad \operatorname{sd}(g_1 \mid \tilde{\mathbf{y}}_a) = \sqrt{\Sigma_{11}}. \tag{27.14}$$

The precision of g_1 is the handbook's plus what the test adds, and the test adds the precision of one difference: the hall reading less the supply, $\bar{y}_T - T_s$, read with the variance $\sigma_T^2/n + \sigma_c^2$, the rise of the path. A rise read to δ kelvin fixes the logarithm of the conductance to $\delta/\Delta T^*$, since a relative error of the rise is the same relative error of G_1. As the number of readings grows the first term of the variance goes to zero and the second stays:

$$\lim_{n\to\infty} \Sigma_{11} = \Big(\frac{1}{\sigma_g^2} + \frac{\Delta T^{*2}}{\sigma_c^2}\Big)^{-1} \approx \Big(\frac{\sigma_c}{\Delta T^*}\Big)^2, \tag{27.15}$$

the floor of the test: the supply's error over the rise, with the handbook's small share beside it. The readings' share of the variance of the rise falls below the supply's once n exceeds σ_T^2/σ_c^2, whatever the load; further readings still lower the spread, but toward the floor and never past it. The floor moves only with the two quantities in it: the load, through ΔT^*, and the supply's precision, σ_c. The correlation of the two entries, $\Sigma_{12}/\sqrt{\Sigma_{11}\Sigma_{22}}$, is positive: a warmer supply explains the same reading with a larger conductance.

Building blocks. The Jacobian (27.4) and the information (27.13) at each (n, Q_{test}); the spreads σ_T, σ_s, σ_0, σ_g; G^*.

The spread of a test. Table 27.4 evaluates equation (27.14). Each row is a test load, with its rise ΔT^* at the requirement; the next four columns give $\mathrm{sd}(g_1 \mid \tilde{\mathbf{y}}_a)$ for $n = 1$, 4, 16 and 64 readings; the last two the floor (27.15) and the supply sensor's error over the rise, $\sigma_s/\Delta T^*$. One row of the arithmetic, 60 kW and $n = 16$: $\sigma_T^2/n + \sigma_c^2 = 0.015625 + 0.082569 = 0.098194\,\mathrm{K}^2$, $1/\Sigma_{11} = 11.111 + 144/0.098194 = 1477.6$, and $\sqrt{\Sigma_{11}} = 0.0260$. A spread of 0.026 in g_1 is 2.6 % in G_1.

Table 27.4: The posterior spread $\mathrm{sd}(g_1 \mid \tilde{\mathbf{y}}_a)$ of a test of n readings at the load Q_{test}, at the requirement, equation (27.14); the floor $n \to \infty$, equation (27.15); and the supply sensor's share $\sigma_s/\Delta T^*$.

Q_{test} [kW]	ΔT^* [K]	$n = 1$	$n = 4$	$n = 16$	$n = 64$	floor	$\sigma_s/\Delta T^*$
30	6	0.0915	0.0621	0.0515	0.0484	0.0473	0.0500
45	9	0.0627	0.0419	0.0346	0.0325	0.0317	0.0333
60	12	0.0475	0.0316	0.0260	0.0244	0.0239	0.0250

Figure 27.3 draws the same function. The horizontal axis is the number of readings on a logarithmic scale and the vertical axis the spread of g_1; the three solid curves are the three loads and the dashed lines their floors. Each curve falls steeply over the first few readings and flattens onto its floor: the readings' variance $0.25/n\,\mathrm{K}^2$ equals the supply's $0.0826\,\mathrm{K}^2$ at $n = 3.03$, and from about fifteen readings on every curve is within 10 % of its floor. Sixty-four readings at 30 kW give 0.048, about what a single reading gives at 60 kW, 0.0475, and four readings at 60 kW, 0.032, already lie below the 30 kW floor of 0.047; the floor at 60 kW, 0.024, is half the floor at 30 kW.

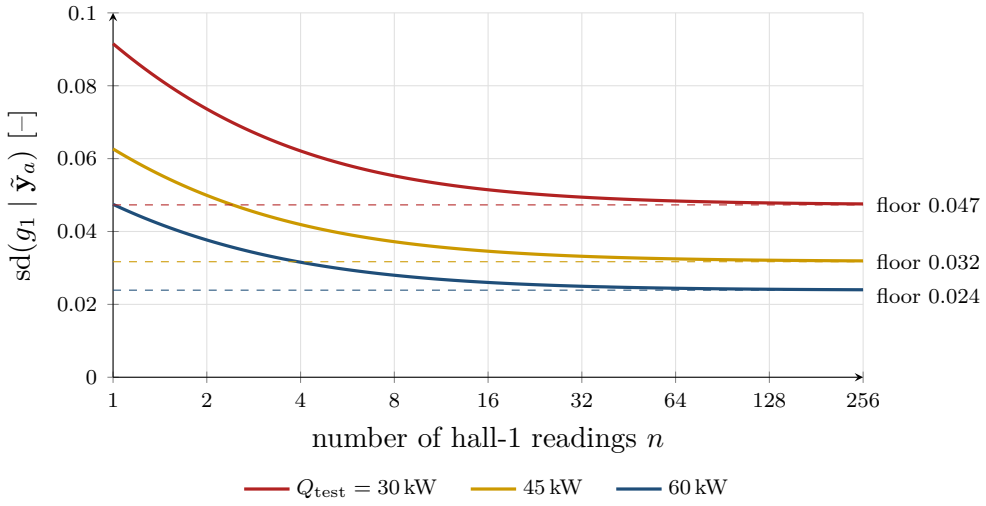


Figure 27.3: Posterior spread of the log-conductance from a test of n readings at three loads, equation (27.14): $\left(1/0.3^2 + \Delta T^{*2}/(0.5^2/n + 0.2873^2)\right)^{-1/2}$ with $\Delta T^* = 6$, 9 and 12 K. Dashed: the floors of equation (27.15).

27.3.4 The operating characteristic

Engineering question. If the path really has 5.5 kW/K, how likely is the test to accept it? And if it has 4.75? The producer asks the first, the consumer the second.

General form. The design form of Chapter 2, equation (2.16), with the indicator that the acceptance rule is met as the quantity u of the posterior the test would produce, and the

outcomes $\tilde{\mathbf{y}}_a$ of the test a weighted by their distribution under a stated true state $\mathbf{z}_\mathrm{t}$ rather than by their predictive:

$$\mathrm{OC}(\mathbf{z}_\mathrm{t}; a) = \int \mathbf{1}\big[\Pr(f(\mathbf{z}) \geq f^* \mid \tilde{\mathbf{y}}_a) \geq \kappa\big]\, p(\tilde{\mathbf{y}}_a \mid \mathbf{z}_\mathrm{t})\, \mathrm{d}\tilde{\mathbf{y}}_a, \qquad U(a) = \int \mathrm{OC}(\mathbf{z}; a)\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}. \tag{27.16}$$

The operating characteristic is the preposterior conditioned on the truth. Averaged over the posterior of the record $\mathbf{y}$, it becomes the design form with u the acceptance indicator: substituting the characteristic and exchanging the order of integration,

$$U(a) = \int \mathbf{1}\big[\Pr(f(\mathbf{z}) \geq f^* \mid \tilde{\mathbf{y}}_a) \geq \kappa\big]\Big(\int p(\tilde{\mathbf{y}}_a \mid \mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}\Big)\mathrm{d}\tilde{\mathbf{y}}_a = \int \mathbf{1}\big[\cdot\big]\, p(\tilde{\mathbf{y}}_a \mid \mathbf{y})\, \mathrm{d}\tilde{\mathbf{y}}_a,$$

the indicator averaged over the predictive of the outcome, equation (2.16): the probability that the test ends with the requirement accepted. Two priors are at work and they differ. The rule inside the indicator judges the test by the analysis prior, the handbook rows the contract names; the outer average weighs the truths by the design prior, here the posterior of the record, what the engineer believes before the test.

Specialized form. The question is equation (27.16) with the test a of n readings at the load Q_test, $f(\mathbf{z}) = G_1$, $f^* = G^*$, $\kappa = 0.95$, the truth $\mathbf{z}_\mathrm{t} = (g_\mathrm{t}, T_{s,\mathrm{t}})$, and $p(\mathbf{z} \mid \tilde{\mathbf{y}}_a)$ inside the indicator the test's posterior of step 3, the Laplace normal with the spread of the plan, so that the indicator is the rule (27.10). Three choices enter here. The truth's supply temperature sits at the set point, the center of its prior: this is the case the plan is written for, and an offset of the true supply moves the operating characteristic (Section 27.4.3). The handbook prior is centered on the requirement, $g^0 = g^*$: this makes the requirement point the prior point, and a prior centered elsewhere adds a pull of its own (Section 27.4.3). The planning route is a closed form, which uses the Laplace rule (27.10) with the spread of the plan, Σ at $\mathbf{z}^\circ$, and the estimate linearized about $\mathbf{z}^\circ$: this gives the characteristic as one normal tail, at the cost of the linearization and of the plan's spread in place of each test's own. Simulated tests, each refitted and judged by the exact tail (27.6) of its own posterior, give the values beside it.

Under the truth the outcomes are the mean of n readings and the supply reading,

$$\bar{y}_T \mid \mathbf{z}_\mathrm{t} \sim \mathcal{N}\Big(T_{s,\mathrm{t}} + \frac{Q_\mathrm{test}}{G_\mathrm{t}},\ \frac{\sigma_T^2}{n}\Big), \qquad y_{T_s} \mid \mathbf{z}_\mathrm{t} \sim \mathcal{N}\big(T_{s,\mathrm{t}},\ \sigma_s^2\big),$$

independent. The linearization point $\mathbf{z}^\circ = (g^*, T_s^0)$ is the prior point because $g^0 = g^*$: both prior rows vanish there, so one Gauss–Newton step from it, the normal equations of Section 27.2.1, gives the estimate as a linear function of the outcomes:

$$\hat{\mathbf{z}} - \mathbf{z}^\circ = -\boldsymbol{\Sigma}\,\mathbf{J}^\mathsf{T}\mathbf{r}(\mathbf{z}^\circ) = \boldsymbol{\Sigma}\begin{bmatrix} -\Delta T^* n/\sigma_T^2 & 0 \\ n/\sigma_T^2 & 1/\sigma_s^2 \end{bmatrix}\begin{bmatrix} \bar{y}_T - T_s^0 - \Delta T^* \\ y_{T_s} - T_s^0 \end{bmatrix},$$

whose first line is

$$\hat{g}_1 - g^* = k_T\,(\bar{y}_T - T_s^0 - \Delta T^*) + k_s\,(y_{T_s} - T_s^0), \qquad k_T = \frac{n}{\sigma_T^2}\big(\Sigma_{12} - \Delta T^* \Sigma_{11}\big), \quad k_s = \frac{\Sigma_{12}}{\sigma_s^2}. \tag{27.17}$$

The gain k_T is negative, a warmer hall reading a smaller conductance. Under the truth, $\hat{g}_1 - g^*$ is normal with

$$m(\mathbf{z}_\mathrm{t}) = k_T\Big(T_{s,\mathrm{t}} - T_s^0 + \frac{Q_\mathrm{test}}{G_\mathrm{t}} - \Delta T^*\Big) + k_s\big(T_{s,\mathrm{t}} - T_s^0\big), \qquad v = k_T^2\,\frac{\sigma_T^2}{n} + k_s^2\,\sigma_s^2, \tag{27.18}$$

the mean and variance of the estimate under the truth (these m and v are local to this chapter, not the conditional moments $m(x)$, $v(x)$ of Chapter 20 nor the Weibull log-shape v), and the rule (27.10) accepts when $\hat{g}_1 - g^* \geq z_\kappa\sqrt{\Sigma_{11}}$. Standardizing,

$$\mathrm{OC}(\mathbf{z}_\mathrm{t}; n, Q_\mathrm{test}) = \Phi\left(\frac{m(\mathbf{z}_\mathrm{t}) - z_\kappa\sqrt{\Sigma_{11}}}{\sqrt{v}}\right). \tag{27.19}$$

The shrinkage and the value at the requirement. The mean and the variance have a reading in the state-vector objects. Split the information into the test's readings and the priors, $\mathbf{\Lambda} = \mathbf{\Lambda}_{\mathrm{d}} + \mathbf{\Lambda}_{\mathrm{prior}}$ with $\mathbf{\Lambda}_{\mathrm{prior}} = \mathbf{J}_{\pi}^{\mathsf{T}}\mathbf{J}_{\pi} = \mathrm{diag}(1/\sigma_g^2,\ 1/\sigma_0^2)$. To first order in the departure of the truth, $Q_{\mathrm{test}}/G_{\mathrm{t}} - \Delta T^{*} \approx -\Delta T^{*}(g_{\mathrm{t}} - g^{*})$, and the expected estimate is $\mathbb{E}[\hat{\mathbf{z}} - \mathbf{z}^{\circ} \mid \mathbf{z}_{\mathrm{t}}] = \mathbf{\Sigma}\mathbf{\Lambda}_{\mathrm{d}}(\mathbf{z}_{\mathrm{t}} - \mathbf{z}^{\circ})$. Since $\mathbf{\Sigma}\mathbf{\Lambda} = \mathbf{I}$, $\mathbf{\Sigma}\mathbf{\Lambda}_{\mathrm{d}} = \mathbf{I} - \mathbf{\Sigma}\mathbf{\Lambda}_{\mathrm{prior}}$, and with the truth's supply at the set point the first entry is

$$m(\mathbf{z}_{\mathrm{t}}) \approx (1 - \pi_g)\,(g_{\mathrm{t}} - g^{*}), \qquad \pi_g = \frac{\Sigma_{11}}{\sigma_g^2}, \tag{27.20}$$

the departure of the truth pulled toward the prior's center by the prior's share π_g of the information of g_1, the analog of the share π_0 of Chapter 5. The variance of the estimate under the truth is the data's part of the posterior covariance, $v = (\mathbf{\Sigma}\mathbf{\Lambda}_{\mathrm{d}}\mathbf{\Sigma})_{11} = (\mathbf{\Sigma} - \mathbf{\Sigma}\mathbf{\Lambda}_{\mathrm{prior}}\mathbf{\Sigma})_{11}$, that is

$$v = \Sigma_{11} - \frac{\Sigma_{11}^2}{\sigma_g^2} - \frac{\Sigma_{12}^2}{\sigma_0^2} \le \Sigma_{11}. \tag{27.21}$$

At the requirement, $\mathbf{z}_{\mathrm{t}} = \mathbf{z}^{\circ}$, which needs both conditions, the prior centered on the requirement and the truth's supply at its set point, the mean is zero and

$$\mathrm{OC}(G^{*}) = \Phi\!\left(-z_{\kappa}\sqrt{\Sigma_{11}/v}\right) \le \Phi(-z_{\kappa}) = 1 - \kappa \qquad (g^{0} = g^{*},\ T_{s,\mathrm{t}} = T_s^{0}), \tag{27.22}$$

with equality when the prior rows carry no information. Under the two conditions the consumer's risk at the requirement itself is at most the complement of the assurance, and below it by as much as the priors shrink the estimate; without them the mean is not zero, and a prior centered above the requirement, or a true supply below the set point, pushes the characteristic at G^{*} above $1 - \kappa$ (Section 27.4.3). Averaged over the analysis prior the rule gives a second guarantee, in the Bayesian sense, which needs neither condition: when the truth is drawn from the prior the rule itself uses, the probability of accepting a path below the requirement is the expectation, over the outcomes, of the indicator of acceptance times the posterior probability of $G_1 < G^{*}$, and on every accepted outcome that probability is at most $1 - \kappa$, so

$$\Pr(G_1 < G^{*},\ \text{accept}) \le (1 - \kappa)\ \Pr(\text{accept}), \qquad \Pr(G_1 < G^{*} \mid \text{accept}) \le 1 - \kappa. \tag{27.23}$$

The producer's risk at a good value G_{g} is $1 - \mathrm{OC}(G_{\mathrm{g}})$, the consumer's risk at a bad value G_{b} is $\mathrm{OC}(G_{\mathrm{b}})$, and the plan of the next question sets both.

Building blocks. $\mathbf{\Sigma}$ at $\mathbf{z}^{\circ}$ from the information (27.13); the gains (27.17); the truths $G_{\mathrm{t}} = 4.5$ to $6.0\,\mathrm{kW/K}$ with $T_{s,\mathrm{t}} = 12^{\circ}\mathrm{C}$; z_{κ} and Φ. For the values beside them, 2000 simulated tests per truth (numpy's `default_rng(27)`), each refitted by the normal equations (27.5) and judged by the exact tail (27.6).

The operating characteristic of the planned test. For 16 readings at $60\,\mathrm{kW}$ the gains (27.17) are

$$k_T = 64\,(0.0068290 - 12 \times 0.00067677) = -0.082707\,\mathrm{K}^{-1},$$
$$k_s = 0.0068290/0.09 = 0.075878\,\mathrm{K}^{-1};$$

the prior's share is $\pi_g = 0.00067677/0.09 = 0.00752$, so $-k_T\Delta T^{*} = 0.99248 = 1 - \pi_g$, as equation (27.20) states; and $\sqrt{v} = 0.025001$ against $\sqrt{\Sigma_{11}} = 0.026015$. One row of the arithmetic, the truth $G_{\mathrm{t}} = 5.5\,\mathrm{kW/K}$: $m = -0.082707 \times (60/5.5 - 12) = 0.090225$, the standardized distance is $(0.090225 - 1.64485 \times 0.026015)/0.025001 = 1.897$, and $\mathrm{OC} = \Phi(1.897) = 0.971$. At the

requirement, equation (27.22) gives $\Phi(-1.64485 \times 1.04055) = \Phi(-1.712) = 0.0435$, below $1 - \kappa = 0.05$.

Table 27.5 gives the characteristic of three tests at six true conductances: each row is a test, each column a truth. Figure 27.4 draws the three as curves. The horizontal axis is the true conductance and the vertical axis the probability of acceptance; the blue curve is 16 readings at 60 kW, the red curve 16 readings at 30 kW and the dashed blue curve 64 readings at 60 kW; the dotted vertical lines mark the consumer's value 4.75, the requirement and the producer's value 5.5, and the gray dot at the requirement marks $1 - \kappa = 0.05$. The two 60 kW curves are steep and nearly coincide: quadrupling the readings moves the acceptance of a 5.5 kW/K path from 0.971 to 0.984. The 30 kW curve is shallow: it accepts a 5.5 kW/K path about half the time and passes 0.9 only near 5.9 kW/K. All three pass close under the gray dot, at 0.042 to 0.044. The values beside them come from the simulated tests judged by the exact tail: for sixteen readings at 60 kW, 0.973 ± 0.004 at 5.5 kW/K and 0.0450 ± 0.0046 at the requirement, against the reported 0.971 and 0.0435; for sixty-four, 0.988 ± 0.002 against 0.984; for sixteen at 30 kW, 0.566 ± 0.011 against 0.529, a gap of 0.037, three to four standard errors. The planning route lies low throughout, by at most 0.010 at 50 and 60 kW, because the exact tail of a test record, skewed toward large conductances as the record's was, accepts more often than the normal tail.

Table 27.5: The operating characteristic, the probability that the test accepts N+1, against the true conductance G_t with the supply at its set point, equation (27.19), $\kappa = 0.95$.

Test	4.5	4.75	5.0	5.25	5.5	6.0
$n = 16$ at 60 kW	0.0000	0.0001	0.0435	0.571	0.971	1.000
$n = 16$ at 30 kW	0.0000	0.0028	0.0417	0.216	0.529	0.943
$n = 64$ at 60 kW	0.0000	0.0000	0.0428	0.620	0.984	1.000

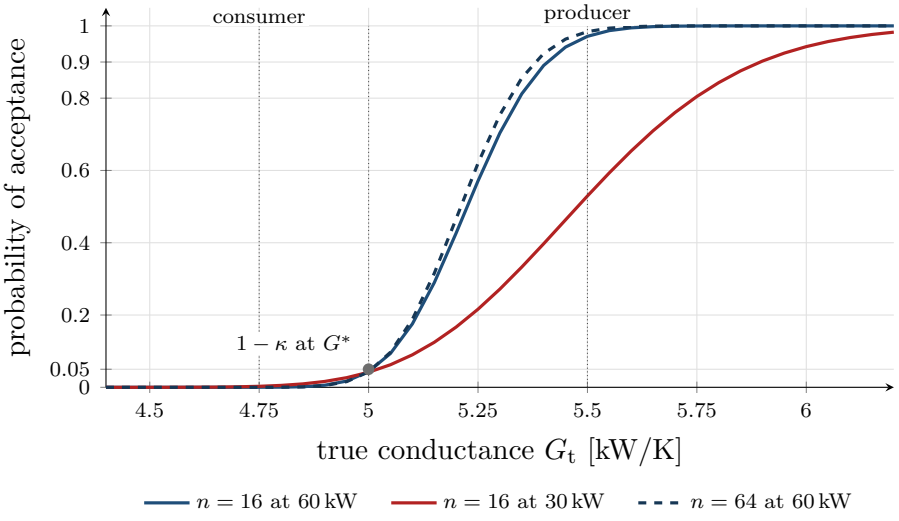


Figure 27.4: Operating characteristic of the acceptance test, equation (27.19), against the true conductance with the supply at 12°C: the probability that a test ends with $\Pr(G_1 \geq 5.0 \mid \tilde{\mathbf{y}}_a) \geq 0.95$. Dotted: the consumer's value 4.75, the requirement 5.0 and the producer's value 5.5 kW/K; the dot marks $1 - \kappa = 0.05$ at the requirement.

27.3.5 The test plan

Engineering question. How long a test, at what load?

General form. The design form, equation (2.16), with the plan chosen by its operating characteristic at the two values the parties name: the smallest test that meets both risks,

$$a^* = \min\{\, a : \mathrm{OC}(G_\mathrm{g}; a) \geq 1 - \epsilon_\mathrm{p}, \quad \mathrm{OC}(G_\mathrm{b}; a) \leq \epsilon_\mathrm{c} \,\},$$

with ϵ_p and ϵ_c the producer's and the consumer's risks the contract allows (written ϵ because α is the level entry of the trend form).

Specialized form. The question is equation (2.16) in the form written above, with the action $a = (n, Q_\mathrm{test})$, the characteristic OC of equation (27.19) from the test's posterior of step 3 at the requirement, $G_\mathrm{g} = 5.5$, $G_\mathrm{b} = 4.75\,\mathrm{kW/K}$, $\epsilon_\mathrm{p} = 0.1$ and $\epsilon_\mathrm{c} = 0.01$. One choice enters here: the two values and the two risks, $G_\mathrm{g} = 5.5\,\mathrm{kW/K}$ accepted with probability at least 0.9 and $G_\mathrm{b} = 4.75\,\mathrm{kW/K}$ accepted with probability at most 0.01, are the parties' and not the analysis's; another pair moves the plan along the same curves (Section 27.6.4). The plan is a single stage, its length fixed in advance; a test that may stop after its first reading is Exercise 27.6.

The plan has two dimensions, the number of readings n and the load Q_test, and they enter the operating characteristic (27.19) differently. The number of readings divides the reading's share of the variance of the rise; the load multiplies the rise itself and with it every term. By the floor (27.15), at a fixed load the characteristic approaches a limit as n grows, the characteristic of a test with the reading's error removed; if that limit is below 0.9 at G_g, no length of test at that load meets the producer's risk. The smallest load at which some length suffices solves $\mathrm{OC}(G_\mathrm{g}; \infty, Q_\mathrm{test}) = 0.9$. The load has an upper bound of its own: above the design load a path at the requirement would exceed its 24°C limit during the test. Against the counts of equations (5.62) and (9.57) (Section 27.1.1), the operating condition of the test enters through the Jacobian entry (27.4).

Building blocks. The operating characteristic (27.19) over $n = 1$ to 64 and its limit; the loads 30, 45, 50 and 60 kW.

The test plan for the path. Table 27.6 is the planning table. Each row is a test load; the columns give the operating characteristic at the producer's 5.5 kW/K for $n = 1$ to 64 readings and in the limit, then the consumer's risk at 4.75 kW/K for one and four readings, and the smallest n that meets both $\mathrm{OC}(5.5) \geq 0.9$ and $\mathrm{OC}(4.75) \leq 0.01$, all by the planning route. At 60 kW five readings clear the producer's risk, with 0.916, and four sit on its boundary, with 0.892; the consumer's risk is 0.0026 with one reading and 0.0003 with four, met at any length. At 50 kW sixteen readings are needed, 0.902. At 45 kW the limit is 0.894 and at 30 kW it is 0.596: by the planning route no length of test at either load meets the producer's risk, and at 30 kW a single reading also misses the consumer's, 0.0108. Solving $\mathrm{OC}(5.5; \infty, Q_\mathrm{test}) = 0.9$ gives the smallest load at which some length suffices, 45.5 kW; at sixteen readings the smallest load is 49.8 kW, at five 58.1 kW, and four readings would need 60.9 kW, above the design load. The values beside them come from the simulated tests judged by the exact tail: at 60 kW, 0.923 ± 0.006 for five readings and 0.902 ± 0.007 for four, against the reported 0.916 and 0.892; at 50 kW with sixteen, 0.903 ± 0.007 against 0.902. The gap for four readings, 0.010, about 1.4 standard errors, puts them on the boundary of the producer's risk by the exact tail where the reported route puts them just below it; five readings meet it by both, and the plan is five.

27.3.6 The rule on a point estimate

Engineering question. What does accepting on a point estimate get wrong? The habit the opening scene followed is to accept when the fitted conductance clears the requirement.

General form. The threshold form, equation (2.12), at the assurance $\kappa = \frac{1}{2}$.

Table 27.6: The test plan: the operating characteristic at the producer's value 5.5 kW/K against the number of readings and the load, the consumer's risk at 4.75 kW/K, and the smallest n meeting OC(5.5) ≥ 0.9 and OC(4.75) ≤ 0.01; equation (27.19).

	OC(5.5) for $n =$								OC(4.75)		
Q_{test} [kW]	1	2	4	5	8	16	64	∞	$n = 1$	$n = 4$	smallest n
30	0.215	0.306	0.399	0.427	0.477	0.529	0.578	0.596	0.0108	0.0051	none
45	0.395	0.552	0.691	0.727	0.786	0.839	0.880	0.894	0.0057	0.0015	none
50	0.460	0.632	0.772	0.806	0.858	0.902	0.933	0.943	0.0045	0.0009	16
60	0.590	0.773	0.892	0.916	0.949	0.971	0.984	0.987	0.0026	0.0003	5

Specialized form. The question is equation (2.12) with $f(\mathbf{z}) = G_1$, $f^* = G^*$ and $\kappa = \frac{1}{2}$, $p(\mathbf{z} \mid \tilde{\mathbf{y}}_a)$ the test's posterior of step 3, the Laplace normal of g_1, judged over the outcomes of the test through the characteristic of equation (27.19). No new choice enters. The point rule accepts when $\hat{G}_1 \geq G^*$, that is $\hat{g}_1 \geq g^*$: the rule (27.10) with $z_\kappa = 0$, the assurance at which the posterior median of G_1 under the Laplace normal is the bound. Its operating characteristic is equation (27.19) with $z_{0.5} = 0$, $\Phi(m(\mathbf{z}_t)/\sqrt{v})$. At the requirement it is $\frac{1}{2}$ whatever the test: the point rule accepts a path exactly at the limit on every other test. Below the requirement its consumer's risk is set by the test's information alone, with no level the contract chose, and it is largest for the weakest test. Above the requirement the order of the two rules reverses: the tail rule's bound lies z_κ spreads below the estimate, so its curve is shifted toward larger conductances, and it accepts a good path less often than the point rule; that shift is what buys its low consumer's risk.

Building blocks. Those of the operating characteristic, with $z_\kappa = 0$.

The point-estimate rule. Table 27.7 sets the point rule beside the rule on the tail for the weakest test, one reading at the present load of 30 kW: each row is a true conductance, and the columns give the probability of acceptance under each rule. At 4.75 kW/K the point rule accepts with probability 0.290 and the tail rule with 0.0108; at the requirement the point rule accepts half the time and the tail rule 0.040. At 5.5 kW/K the point rule accepts with 0.830, the tail rule with 0.215. For sixteen readings at 60 kW the point rule's consumer's risk at 4.75 falls to 0.018, against 0.0001. Beside the tail rule's values for the weak test, the simulated tests judged by the exact tail give 0.0140 ± 0.0026 at 4.75 and 0.251 ± 0.010 at 5.5 kW/K; the point rule's consumer's risk is more than twenty times either. Figure 27.5 draws the two characteristics of the weak test. The horizontal axis is the true conductance and the vertical axis the probability of acceptance; the red curve is the point rule and the blue curve the tail rule, and the dotted line marks the consumer's value. The red curve crosses $\frac{1}{2}$ at the requirement and keeps a long shoulder below it; the blue curve lies to its right throughout.

Table 27.7: The point-estimate rule $\hat{G}_1 \geq G^*$ ($\kappa = 0.5$) against the rule on the tail ($\kappa = 0.95$), for one reading at 30 kW, equation (27.19).

Rule, one reading at 30 kW	4.5	4.75	5.0	5.25	5.5	6.0
point estimate, $\hat{G}_1 \geq G^*$	0.121	0.290	0.500	0.692	0.830	0.960
tail, $\Pr(G_1 \geq G^* \mid \tilde{\mathbf{y}}_a) \geq 0.95$	0.0018	0.0108	0.0405	0.107	0.215	0.503

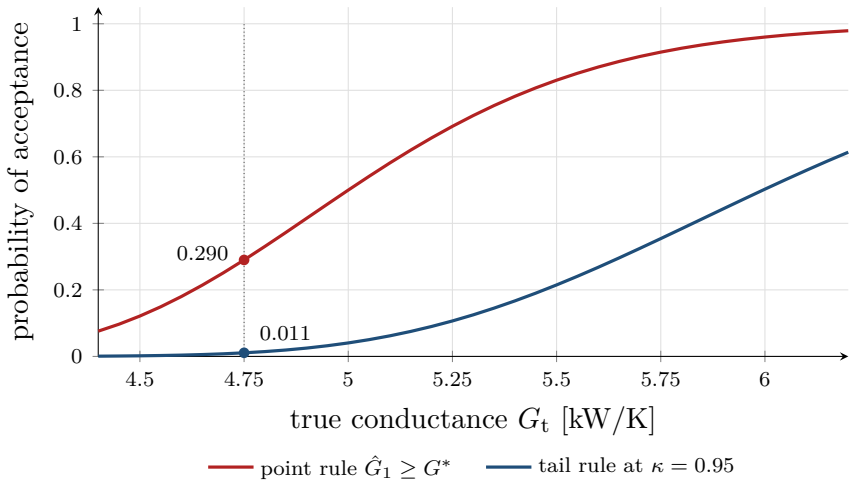


Figure 27.5: The consumer's risk of the point-estimate rule. Operating characteristics of one reading at 30 kW, equation (27.19), with $z_\kappa = 0$ (red, accept when $\hat{G}_1 \geq 5.0$) and $z_\kappa = 1.645$ (blue, $\kappa = 0.95$). Dotted: the consumer's value 4.75 kW/K.

Common mistake

Accepting a capacity because its fitted value clears the requirement. The point rule is the acceptance rule at an assurance of one half: it accepts a path exactly at the limit on every other test, and a path five percent below it with a probability the test's information alone sets, 0.29 for one reading at 30 kW. The rule is stated on the posterior tail, $\Pr(G_1 \geq G^* \mid \mathbf{y}) \geq \kappa$, which bounds the consumer's risk at the requirement by $1 - \kappa$.

27.4 The checks (step 5)

The answers rest on the Laplace normal of the record, on a linearized closed form for the test, on the priors and the truth's supply, on the pinned load and on the limit of exact balances. Each is set beside a calculation that does without it.

27.4.1 The Laplace tail against the full posterior on the record

Under the Laplace normal, equation (27.8) gives the standardized distance and the tail

$$\frac{1.70297 - 1.60944}{0.10669} = 0.8767, \qquad \Pr(G_1 \geq 5.0\,\text{kW/K} \mid \mathbf{y}) \approx \Phi(0.8767) = 0.810,$$

against the full posterior's 0.8395 ± 0.0008; the unweighted draws reproduce the Laplace value, 0.809 ± 0.0009. The difference, 0.030, is thirty-eight standard errors of the reweighted tail: a gap of the approximation, not of the sampling. The other summaries move with it: the Laplace 0.05 quantile, equation (27.10), is $\exp(1.70297 - 1.64485 \times 0.10669) = 4.607$ kW/K against the full 4.683, and the Laplace Bayes factor is $0.8097/0.1903 = 4.25$, $\ln B_{10} = 1.448$ nats, against the full 5.23 and 1.655. The cause is the one Chapter 20 traced to the hall-1 reading row: the full posterior of g_1 is skewed toward large conductances, with mean 1.716 above the MAP value 1.703, and the symmetric normal puts too much of its mass below the requirement. The decision does not change, both routes leaving the tail below 0.95; the Bayes factor changes by a quarter.

27.4.2 The closed form against simulated tests

The operating characteristic (27.19) rests on a linearized estimate, on the plan's spread $\boldsymbol{\Sigma}(\mathbf{z}^\circ)$ in place of each fit's own, and on the Laplace form of the rule. Each test was therefore simulated

2000 times at each truth, with numpy's `default_rng(27)`, drawn in the order of the designs and truths below, the means of the readings before the supply readings: the mean $\bar{y}_T$ from $\mathcal{N}(12 + Q_{\text{test}}/G_{\text{t}},\ 0.25/n)$ and the supply reading from $\mathcal{N}(12,\ 0.09)$; each record was fitted by the normal equations (27.5) from the prior point, and judged twice, by the exact tail (27.6) on a grid of 2801 points and by the Laplace rule (27.10) with the fit's own Σ_{11}. Table 27.8 gives the results: each row is a test and a truth, and the columns are the closed form, the fraction of simulated tests accepted under the exact tail with its standard error $\sqrt{\hat{P}(1-\hat{P})/2000}$, the fraction under the Laplace tail, and the number of the 2000 tests on which the two tails decide differently.

Table 27.8: The operating characteristic, closed form (27.19) against 2000 simulated tests per truth (numpy `default_rng(27)`), each refitted and judged by the exact tail (27.6) and by the Laplace tail with its own Σ_{11}.

Test	G_{t} [kW/K]	Closed form	Simulated, exact tail	Simulated, Laplace tail	Decisions differ
$n = 16$ at 60 kW	4.75	0.0001	0.0000 ± 0.0000	0.0000	0
	5.0	0.0435	0.0450 ± 0.0046	0.0355	19
	5.25	0.571	0.579 ± 0.011	0.551	55
	5.5	0.971	0.973 ± 0.004	0.968	10
$n = 16$ at 30 kW	4.75	0.0028	0.0025 ± 0.0011	0.0015	2
	5.0	0.0417	0.0450 ± 0.0046	0.0320	26
	5.5	0.529	0.566 ± 0.011	0.513	107
	6.0	0.943	0.951 ± 0.005	0.934	34
$n = 64$ at 60 kW	5.0	0.0428	0.0455 ± 0.0047	0.0405	10
	5.5	0.984	0.988 ± 0.002	0.985	7
$n = 4$ at 60 kW	5.5	0.892	0.902 ± 0.007	0.886	32
$n = 5$ at 60 kW	5.0	0.0448	0.0500 ± 0.0049	0.0430	14
	5.5	0.916	0.923 ± 0.006	0.910	27
$n = 16$ at 50 kW	5.5	0.902	0.903 ± 0.007	0.888	30
$n = 1$ at 30 kW	4.75	0.0108	0.0140 ± 0.0026	0.0065	15
	5.5	0.215	0.251 ± 0.010	0.167	168

At 50 and 60 kW the closed form, the planning route, agrees with the simulated tests under the exact tail at every entry of the table, lying slightly low; the largest difference is 0.010, for four readings at 5.5 kW/K, about 1.4 standard errors. The plan stands with one qualification: five readings at 60 kW accept a 5.5 kW/K path with 0.923 ± 0.006 and clear the producer's risk on both routes, while four sit on its boundary, 0.902 ± 0.007 simulated against 0.892 by the closed form, so the boundary between four and five readings is not resolved by either route and the plan takes five. At 30 kW the closed form is low, by 0.037 for sixteen readings at 5.5 kW/K and by the same for one reading, three to four standard errors; the conclusion of the plan is unchanged, since 0.566 is far from 0.9. The limits of long tests move the same way: 4000 simulated tests of 10 000 readings each (numpy `default_rng(99)`) accept a 5.5 kW/K path with 0.628 at 30 kW and 0.907 at 45 kW, against the closed form's 0.596 and 0.894, so the smallest load at which some length suffices is about 45 kW rather than 45.5, and at 45 kW only a test of thousands of readings reaches the producer's risk. At the requirement the simulated tests accept with 0.045 to 0.050, within their standard errors of the bound $1 - \kappa$, as equation (27.22) states for a prior centered on the requirement and a truth's supply at the set point.

The tail used to judge a single test matters more than the linearization. On every design the Laplace tail of a test record accepts less often than the exact tail, because the exact posterior of g_1 is skewed toward large conductances, as the record's was: the largest difference between the two tails on one simulated record is 0.014 at sixteen readings and 60 kW, 0.022 at sixteen readings and 30 kW, and 0.035 for one reading at 30 kW, and a difference of that size moves the

decision of a test whose tail lies near 0.95. The weaker the test, the more decisions differ: 10 of 2000 at 60 kW against 168 for one reading at 30 kW. For the plan, five readings at 60 kW accept a 5.5 kW/K path with 0.910 under the Laplace tail of each test and 0.923 under the exact tail, and four with 0.886 and 0.902: five meet the producer's risk whichever tail judges the test, and the tail is a term the acceptance clause states.

The guarantee over the prior. The bound (27.23) was checked by drawing 400 000 truths from the handbook priors, $g_{\mathrm{t}} \sim \mathcal{N}(\ln 5,\ 0.3^2)$ and $T_{s,\mathrm{t}} \sim \mathcal{N}(12,\ 1)$, with numpy's `default_rng(5)`, and averaging the closed form over them. For sixteen readings at 60 kW the test accepts with probability 0.442, and of the accepted paths a fraction 0.0016 lies below the requirement; for one reading at 30 kW the figures are 0.281 and 0.0075. Both lie well under $1-\kappa = 0.05$, because most accepted paths are far above G^*. The point rule breaks the bound on the weak test: of the paths it accepts after one reading at 30 kW, 0.095 lie below the requirement, and 0.027 after sixteen readings at 60 kW.

27.4.3 Sensitivity to the priors, the supply and the load

Table 27.9 varies the inputs of the plan for sixteen readings at 60 kW. Each row is one change; the columns give the spread of g_1 at the requirement and the operating characteristic at the consumer's value, at the requirement and at the producer's value. The handbook spread barely matters: a flat prior on g_1 or a spread halved to 0.15 moves no entry by more than 0.002, since the test carries more than 99 % of the information. The truth's supply does matter at the requirement: a true supply 0.5 K below the set point, which the set-point prior pulls the estimate toward, raises the acceptance of a path at G^* to 0.058, above $1-\kappa$; the bound (27.22) holds for a truth at the prior's center, and a better supply reading is the remedy. The prior's center is the second condition of the bound. Centered at 5.5 kW/K instead of on the requirement, the handbook prior adds the pull $\pi_g \ln(5.5/5)$ to the mean (27.18), and the acceptance at G^* rises to 0.0462 for sixteen readings at 60 kW and to 0.0503, above $1-\kappa$, for one reading at 30 kW, where the prior's share is 0.093; centered at 4.5, it falls to 0.0406 and 0.0315. The largest change is the load. If the load is not held by a calibrated bank but read by the 1.5 kW meter of Table 17.1, as a free entry with its reading row, its relative error, 2.5 %, enters the rise as the supply's does: the spread of g_1 grows from 0.026 to 0.036, its floor from 0.024 to 0.034, and the acceptance of a 5.5 kW/K path falls to 0.809, with a limit of 0.839 as n grows; no length of test meets the producer's risk.

Table 27.9: Sensitivity of the plan of sixteen readings at 60 kW: $\mathrm{sd}(g_1 \mid \tilde{\mathbf{y}}_a)$ at the requirement and the operating characteristic at 4.75, 5.0 and 5.5 kW/K, equation (27.19).

Change	$\mathrm{sd}(g_1)$	OC(4.75)	OC(5.0)	OC(5.5)
none (handbook $\sigma_g = 0.3$, truth's supply at 12°C)	0.0260	0.0001	0.0435	0.971
flat prior on g_1	0.0261	0.0001	0.0441	0.972
handbook spread $\sigma_g = 0.15$	0.0257	0.0001	0.0417	0.970
handbook centered at 5.5 kW/K	0.0260	0.0001	0.0462	0.973
handbook centered at 4.5 kW/K	0.0260	0.0001	0.0406	0.969
truth's supply 12.5°C	0.0260	0.0000	0.0323	0.961
truth's supply 11.5°C	0.0260	0.0001	0.0576	0.979
load read by the 1.5 kW meter, not held	0.0359	0.0008	0.0457	0.809

The prior of the record. The record's answers depend on the handbook's center as well. Centered at 4.5 instead of 5 kW/K, the prior gives the requirement the probability 0.363 before

the record, and the refitted record, by the Laplace normal of its g_1, 0.778 after it (0.810 by reweighted draws), a Bayes factor of $3.508/0.569 = 6.16$ ($4.263/0.569 = 7.49$); centered at 5.5, 0.625 before and 0.836 after (0.864), a factor of $5.104/1.664 = 3.07$ ($6.371/1.664 = 3.83$). The probability moves by about 0.03 either way and stays below 0.95 by both routes; the Bayes factor moves the other way, from 3.07 to 6.16 (3.83 to 7.49), larger when the prior leaned against the requirement: it measures the shift the record made from that prior, and it depends on the prior through the halves the cut makes of it. The decision on the record does not depend on the handbook's center within this range.

The exact balance as a limit. The test runs on the reduced state, in the limit $\sigma_p \to 0$ of equation (17.19). On the test's full state (g_1, T_s, T_1), with hall 1's physics row $(T_1 - T_s - Q_{\text{test}}\mathrm{e}^{-g_1}/G_{\text{ref}})/\sigma_p$ beside the reading row on T_1, the spread of g_1 at the requirement for sixteen readings at 60 kW is 0.0838 at $\sigma_p = 1$ K, 0.02730 at 0.1 K, 0.026028 at 0.01 K and 0.026015 at 0.001 K, against the reduced state's 0.026015: the difference falls as σ_p^2.

27.5 The report (step 6)

The record does not demonstrate N+1, and no test at the present load can. Hall 1's path is fitted at 5.49 kW/K, ten percent above the 5.0 the design basis asks, but the probability that it is at least 5.0 is 0.81 from the posterior of the state (0.84 by the full posterior, which the skew of the log conductance raises by 0.03), short of the assurance of 0.95 by both. A test at the present 30 kW would accept a path that truly has 5.5 kW/K with probability near one half after sixteen readings and below two in three however long it ran, because the supply sensor's error over a 6 K rise sets a floor that more readings do not lower. At the design load of 60 kW, five readings accept such a path with probability 0.92 (0.916 by the planning route, 0.923 simulated) and a path five percent below the requirement with 0.0003; four readings sit on the producer's boundary of 0.9, 0.89 by the planning route and 0.90 by the simulated tests, and five clear it on both. The answers below are in the units in which the questions were asked.

27.5.1 The answers and their bands

Table 27.10 gives the answers in the order of Table 27.2: the answer from the posterior of the state in the second column, the exact value beside it in the third, and in the fourth the band and what the checks of Section 27.4 found.

The report line. On the record of Chapter 18, hall 1's path has a conductance of 5.5 kW/K (median, 95 % band 4.5 to 6.8 from the posterior of the state, 4.5 to 6.9 by the full posterior), and the probability that it meets the N+1 requirement of 5.0 kW/K is 0.81 (0.84 by the full posterior), so the requirement is not demonstrated at the assurance of 0.95; the record favours it by a factor of four (five by the full posterior). An acceptance test of five readings of the hall-1 sensor and one of the supply sensor, at the design load of 60 kW held by a calibrated load bank in the N+1 configuration, judged by the posterior tail at 0.95, accepts a 5.5 kW/K path nine times in ten and a 4.75 kW/K path three times in ten thousand. A test at the present load cannot settle the question at any length.

27.6 The next measurement (step 7)

The report leaves the requirement undemonstrated at 0.81 (0.84 by the full posterior), and the measurement that would settle it is the acceptance test itself: the test is the next measurement for the first question. The candidates are tests of the plan, rows added to the plant before handover.

Table 27.10: The answers for the capacity of hall 1's path, from the posterior of the state: the Laplace normal of the record's g_1, and for a test the closed form, the planning route. Beside them the exact values: the record's full posterior by reweighted draws, and 2000 simulated tests per truth judged by the exact tail.

Question	Answer (state route)	Exact	Band and check
Does N+1 hold?	not accepted: $\Pr(G_1 \geq 5.0 \mid \mathbf{y}) = 0.810$, below 0.95; $G_1 \geq 4.61$ kW/K with probability 0.95	0.840 ± 0.001 and 4.68, higher by 0.030	not accepted by both; Laplace 0.778 to 0.836 over handbook centers 4.5 to 5.5 (full 0.810 to 0.864)
Which hypothesis does the evidence favour, N+1 holds or not?	N+1 holds, by $B_{10} = 4.25$, $\ln B_{10} = 1.45$ nats, positive	$B_{10} = 5.23$, $\ln B_{10} = 1.66$ nats	19 (2.94 nats) needed to accept at even odds; 3.1 to 6.2 over the handbook centers (full 3.8 to 7.5)
How well does a test fix the conductance, and why do more readings stop helping?	$\mathrm{sd}(g_1) = 0.026$ for sixteen readings at 60 kW, 0.051 at 30 kW; floors 0.024 and 0.047	–	the floor is the supply's 0.287 K over the rise; past $n \approx 3$ the supply's error dominates, and the 60 kW spread still falls 18 % from four readings to sixteen, toward the floor; a metered load raises the 60 kW floor to 0.034
If the path really has 5.5 kW/K, how likely is the test to accept it? And if 4.75?	sixteen readings at 60 kW: 0.971 and 0.0001; at 30 kW: 0.53 and 0.003	0.973 ± 0.004 and 0.0000; 0.566 ± 0.011 at 30 kW	at G^*, 0.045 to 0.050 simulated, at most $1 - \kappa$ for a prior centered on G^* and a supply at the set point
How long a test, at what load?	five readings at 60 kW: 0.916 at 5.5, 0.0003 at 4.75	0.923 ± 0.006	four readings sit on the boundary, 0.892 planned and 0.902 ± 0.007 simulated, and five clear it; no practical length suffices below about 45 kW (45.5 by the closed form); sixteen at 50 kW
What does accepting on a point estimate get wrong?	its consumer's risk at 4.75: 0.290 for one reading at 30 kW, against 0.011 for the tail rule	tail rule 0.0140 ± 0.0026	at G^* it accepts half the time on any test; of its accepted paths 0.095 lie below G^* over the handbook prior

27.6.1 The candidate tests

Two numbers rank a candidate test a against the record. The first is the width it removes from the record's spread of g_1: appended to the record of Chapter 18, the test adds the rows of Table 27.1 on g_1 and on a new entry T_s', the supply temperature on the day of the test, with its own set-point prior, and the preposterior covariance is equation (2.17),

$$\boldsymbol{\Sigma}_a = \left(\boldsymbol{\Sigma}^{-1} \oplus 0 + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a\right)^{-1},$$

with $\boldsymbol{\Sigma}^{-1} \oplus 0$ the record's information bordered by a zero row and column for T_s', and $\mathbf{J}_a$ the test's three rows on (g_1, T_s') at $\hat{g}_1$. The second is the probability that the test, judged by itself, ends with N+1 accepted: $U(a)$ of equation (27.16), the operating characteristic averaged over the design prior $p(\mathbf{z} \mid \mathbf{y})$, the record's posterior, while the test is still judged by its analysis prior, the handbook. On the posterior of the state, the Laplace normal of g_1, it is a one-dimensional expectation, evaluated by the Gauss–Hermite rule of Chapter 5, equation (5.33), with 40 nodes,

$$U(a) \approx \frac{1}{\sqrt{\pi}} \sum_{q=1}^{40} w_q \, \mathrm{OC}\!\left(G_{\mathrm{ref}} \exp(\hat{g}_1 + \sqrt{2}\,\sqrt{\Sigma_{11}}\, t_q); a\right);$$

the value beside it averages over the 200 000 reweighted draws of G_1, the full posterior,

$$U(a) = \sum_j \bar{\zeta}_j \, \mathrm{OC}\!\left(G_1^{(j)}; a\right).$$

Table 27.11 lists six candidates. Each row is a test; the columns give $\sqrt{\Sigma_{a,11}}$ on the record with the test, the width removed from 0.1067, the spread of the test alone, and $U(a)$ from the posterior of the state with the full-posterior value beside it.

Table 27.11: Candidate acceptance tests against the record of Chapter 18: the spread of g_1 with the test appended (equation (2.17)), the width removed from 0.1067, the spread of the test alone at the requirement, and the probability $U(a)$ that the test accepts N+1, over the Laplace normal of the record's g_1 (reported) and over its full posterior.

Test	$\sqrt{\Sigma_{a,11}}$	Width removed	Test alone	$U(a)$, Laplace	$U(a)$, full
$n = 1$ at 60 kW	0.0473	0.0594 (56 %)	0.0475	0.530	0.566
$n = 4$ at 60 kW	0.0331	0.0736 (69 %)	0.0316	0.637	0.672
$n = 5$ at 60 kW	0.0318	0.0749 (70 %)	0.0302	0.646	0.681
$n = 16$ at 60 kW	0.0277	0.0790 (74 %)	0.0260	0.673	0.707
$n = 16$ at 45 kW	0.0360	0.0707 (66 %)	0.0346	0.617	0.653
$n = 16$ at 30 kW	0.0505	0.0562 (53 %)	0.0515	0.504	0.540

Every test at 60 kW removes more than half the record's spread with its first reading, and the spread with the test appended is close to the test's own: the test, not the record, will decide. The probability of acceptance rises from 0.530 for one reading to 0.646 for five and only to 0.673 for sixteen (0.566, 0.681 and 0.707 over the full posterior, each 0.034 to 0.036 higher, the record's skew carried into the average); it cannot approach one, because the record itself gives 0.19 of its probability to paths below the requirement (a sixth by the full posterior), which a good test rejects. The order of the candidates is the same by both. Sixteen readings at 30 kW do worse than one at 60 kW on both counts.

27.6.2 The recommendation

The next measurement is the acceptance test of the plan: hall 1 held at the design load of 60 kW by a calibrated load bank, in the N+1 configuration the clause names, five readings of

the hall-1 sensor in steady state and one of the supply sensor, judged by the posterior tail at $\kappa = 0.95$. Given the record, it accepts N+1 with probability 0.65 from the posterior of the state (0.68 over the full posterior); a path that truly has 5.5 kW/K passes with 0.92 (0.916 planned, 0.923 simulated under the exact tail, 0.910 under the Laplace tail of each test), and a path five percent short fails with probability 0.9997. Five readings rather than four because four sit on the producer's boundary and five clear it by the planning route, by the simulated tests and under either tail. Further readings add little: eleven more raise the probability of acceptance by 0.027 (0.026 over the full posterior). A better supply sensor would lower the floor that sets the plan, and a load that is not held would destroy it. What the test costs in hours of a hall at its design load, and what its verdict is worth against that cost, is the value-of-information question that Chapter 31, in Volume III, prices.

27.6.3 What the analysis showed

The first reading of the fit, 5.49 kW/K against 5.0, said that N+1 holds; the analysis does not support it. The record puts the probability of the requirement at 0.81 and favours it by a factor of four (0.84 and five by the full posterior, whose skew toward large conductances the symmetric normal of the log conductance does not carry), which is evidence for the requirement but not a demonstration at the customer's assurance by either. The engineer's instinct for a demonstration, to run longer, is the wrong lever for this path: the conductance is seen through the rise of the hall above the supply, the supply sensor's error is a fixed fraction of a small rise, and at the present load no number of readings brings the acceptance of a good path above two in three. What settles the question is the operating condition of the test. At the design load the same sensor reads a rise twice as large, five readings suffice, and the test should be run in the configuration the clause names, with a load bank whose power is known, judged by the tail of the posterior and never by the fitted value, which would accept a path at the limit on every other test.

27.6.4 Limits of the model

The configuration of the record. The record was taken in normal operation. If taking a unit out for N+1 changes the path's air or water flow, the record's conductance is not the conductance the clause is about, and only a test in the N+1 configuration measures it; the record's answer is then a prior for the test, not a case of the rule.

The supply at the design load. The requirement was written with the supply at its set point. A chiller near its capacity lets the supply rise with the load, and the rise of 12 K would then be measured from a warmer supply; the requirement becomes a joint event in g_1 and T_s, as the scenario temperature of Chapter 20 was, and the test must read the supply at the test load, which this plan does.

Independent readings in steady state. The readings were taken as independent given the state. A sensor error that drifts slowly, or a hall not yet in steady state when the readings start, gives errors that the mean of n readings does not average away; the spread of the reading row then stops falling as $1/\sqrt{n}$, and the floor is reached sooner. The plan of five readings is less exposed to this than a long test, but the readings should be spaced past the hall's time constant.

The conductance at the test load. The test reads the conductance at the load at which it is run. If the conductance depends on the load, through the chilled-water flow, a test at the present load would also measure the wrong quantity, a second reason to test at the design load.

The truths and the risks. The producer's and consumer's values, 5.5 and 4.75 kW/K, and their risks are the parties'. The plan follows them along the curves of Figure 27.4: the recommended five readings at 60 kW already reject a 4.9 kW/K path with probability 0.992 (OC = 0.0085), but a consumer value closer still to the requirement, or a producer value closer to it from above, needs a steeper curve, more readings or a better supply sensor.

The next chapter turns from a requirement on one capacity to operating margins: the probability that a thermal limit is violated at a future load, and whether a unit delivers the capacity its plate states.

27.7 What to retain

Definition 27.1 (Acceptance of a capacity read through a model)**.** A requirement on a capacity, $G_1 \geq G^*$, is an event on one entry of the state, $g_1 \geq \ln(G^*/G_{\mathrm{ref}})$, and the acceptance rule is its posterior tail against the assurance, $\Pr(G_1 \geq G^* \mid \mathbf{y}) \geq \kappa$, read from the posterior of the state as the normal tail of that entry, with the full posterior's tail set beside it where the two depart. The same tail as posterior odds over prior odds is the Bayes factor of the requirement against its complement, and the rule is $B_{10} \geq (\kappa/(1-\kappa))/(P_0/(1-P_0))$. A test of n readings of the hall temperature at a held load Q_{test} adds to the precision of g_1 the precision of the rise, $\Delta T^{*2}/(\sigma_T^2/n + \sigma_c^2)$, with $\Delta T^* = Q_{\mathrm{test}}/G^*$ the Jacobian entry of the reading row; its floor as n grows is $(\sigma_c/\Delta T^*)^2$, set by the supply's error and the load, not by the duration. The operating characteristic of the test is the rule averaged over the outcomes the test would produce from a stated truth, a preposterior conditioned on the truth, normal in the linearized estimate with mean $(1-\pi_g)(g_{\mathrm{t}} - g^*)$ and variance $(\boldsymbol{\Sigma} - \boldsymbol{\Sigma}\boldsymbol{\Lambda}_{\mathrm{prior}}\boldsymbol{\Sigma})_{11}$; at the requirement it is at most $1-\kappa$ when the prior is centered on the requirement and the truth's supply sits at its set point. Averaged over the design prior, the posterior of what is known before the test, it is the probability that the test ends accepted. The plan is the smallest test whose characteristic meets the producer's and the consumer's risks, planned on the closed form and checked by simulated tests judged by the exact tail; the point-estimate rule, $\kappa = \frac{1}{2}$, accepts a path at the limit half the time.

Table 27.12 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 27.12: The terms of acceptance testing of a capacity and where each is defined.

Term	Expression	Units	Meaning	Defined in
Requirement	$G_1 \geq G^* \iff g_1 \geq g^* = \ln(G^*/G_{\mathrm{ref}})$	kW/K	N+1 as a capacity	27.3.1
Acceptance rule	$\Pr(G_1 \geq G^* \mid \mathbf{y}) \geq \kappa$	–	a tail of one entry	27.3.1
One-sided bound	$G_{1-\kappa} \approx G_{\mathrm{ref}} \exp(\hat{g}_1 - z_\kappa \sqrt{\Sigma_{11}})$	kW/K	accept iff $G^* \leq G_{1-\kappa}$	27.3.1
Bayes factor of the requirement	posterior odds over prior odds of $G_1 \geq G^*$	–	weight of the record	27.3.2
Rule as a Bayes factor	$B_{10} \geq (\kappa/(1-\kappa))/(P_0/(1-P_0))$	–	19 at even odds	27.3.2
Rise at the requirement	$\Delta T^* = Q_{\mathrm{test}}/G^*$	K	Jacobian entry on g_1	27.3.3
Spread of a test	$1/\Sigma_{11} = 1/\sigma_g^2 + \Delta T^{*2}/(\sigma_T^2/n + \sigma_c^2)$	–	reading plus supply over the rise	27.3.3
Floor	$(1/\sigma_g^2 + \Delta T^{*2}/\sigma_c^2)^{-1} \approx (\sigma_c/\Delta T^*)^2$	–	limit $n \to \infty$	27.3.3
Supply precision	$1/\sigma_c^2 = 1/\sigma_s^2 + 1/\sigma_0^2$	K^{-2}	sensor and set point	27.3.3
Operating characteristic	$\Phi\big((m - z_\kappa \sqrt{\Sigma_{11}})/\sqrt{v}\big)$	–	acceptance at a truth	27.3.4
Gains	$k_T = (n/\sigma_T^2)(\Sigma_{12} - \Delta T^* \Sigma_{11})$, $k_s = \Sigma_{12}/\sigma_s^2$	K^{-1}	estimate as a linear function	27.3.4
Shrinkage	$m \approx (1 - \pi_g)(g_{\mathrm{t}} - g^*)$, $\pi_g = \Sigma_{11}/\sigma_g^2$	–	pull toward the prior	27.3.4
Value at the requirement	$\Phi(-z_\kappa \sqrt{\Sigma_{11}/v}) \leq 1 - \kappa$	–	consumer's risk at G^*, for $g^0 = g^*$ and $T_{s,\mathrm{t}} = T_s^0$	27.3.4
Guarantee over the prior	$\Pr(G_1 < G^* \mid \mathrm{accept}) \leq 1 - \kappa$	–	Bayesian consumer's risk	27.3.4
Producer's, consumer's risk	$1 - \mathrm{OC}(G_{\mathrm{g}})$, $\mathrm{OC}(G_{\mathrm{b}})$	–	the two errors of the test	27.3.4
Preposterior acceptance	$U(a) = \int \mathrm{OC}(\mathbf{z}; a)\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}$	–	the test ends accepted	27.3.4
Point-estimate rule	$\hat{G}_1 \geq G^*$, the rule at $\kappa = \frac{1}{2}$	–	accepts half at G^*	27.3.6
Averaged readings	one row on $\bar{y}_T$ with $\sigma_T/\sqrt{n}$	K	n rows summed	27.1.5
Exact tail of a test	ratio of integrals of $\mathrm{e}^{-\mathcal{L}(g, \hat{T}_s(g))}$	–	T_s integrated in closed form	27.2.1

27.8 Exercises

Exercise 27.1. *Formulation.* (a) A fit gives $\hat{g}_1 = 1.68$ with $\sqrt{\Sigma_{11}} = 0.04$ for a path whose requirement is $G^* = 5.0\,\mathrm{kW/K}$. Write the acceptance rule at $\kappa = 0.95$ on the state, decide it by the tail and by the one-sided bound, and give the Bayes factor at even prior odds. (b) A test's linearized estimate is normal about $g^* + 0.95 \ln(G_{\mathrm{t}}/G^*)$ with standard deviation 0.025, and its rule uses $\sqrt{\Sigma_{11}} = 0.026$. Give its operating characteristic at $G_{\mathrm{t}} = 5.4\,\mathrm{kW/K}$. ((a) $\Pr(g_1 \geq \ln 5 \mid \mathbf{y}) = \Phi((1.68 - 1.6094)/0.04) = \Phi(1.764) = 0.961 \geq 0.95$, accepted; the bound $\exp(1.68 - 1.6449 \times 0.04) = 5.024 \geq 5.0$ agrees; $B_{10} = 0.961/0.039 = 24.7$, above 19. (b) $m = 0.95 \ln 1.08 = 0.0731$; $(0.0731 - 1.6449 \times 0.026)/0.025 = 1.214$, $\mathrm{OC} = \Phi(1.214) = 0.888$.)

Exercise 27.2. *Formulation.* The supply sensor of the test is replaced by one with $\sigma_s = 0.1\,\mathrm{K}$, the set-point prior kept at 1 K. Give the combined supply spread and the floor of $\mathrm{sd}(g_1)$ for a test at 60 kW, and state what the floor was with the present sensor and why the duration does not appear. ($\sigma_c = (1/0.01 + 1)^{-1/2} = 0.0995\,\mathrm{K}$; the floor is $(1/0.09 + 144/0.0099)^{-1/2} = 0.0083$, about $\sigma_c/12$, against 0.0239 with the 0.3 K sensor; the floor is the limit $n \to \infty$ of equation (27.14), in which the reading's variance σ_T^2/n has vanished, leaving the supply's error over the rise.)

Exercise 27.3. *Derivation.* For the test state $\mathbf{z} = (g_1, T_s)^{\mathsf{T}}$ with the four rows of Table 27.1, derive the Jacobian at $\mathbf{z}^\circ$, the information matrix, and the first entry of $\mathbf{\Sigma}$; show that it

equals $(1/\sigma_g^2 + \Delta T^{*2}/(\sigma_T^2/n + \sigma_c^2))^{-1}$, and take the limit $n \to \infty$. ($\mathbf{J}$ as in equation (27.4); $\mathbf{\Lambda}$ as in equation (27.13); $\Sigma_{11} = \Lambda_{22}/\det\mathbf{\Lambda}$ with $\det\mathbf{\Lambda} = n\Delta T^{*2}/(\sigma_T^2\sigma_c^2) + \Lambda_{22}/\sigma_g^2$, so $1/\Sigma_{11} = 1/\sigma_g^2 + \Delta T^{*2}(n/\sigma_T^2)(1/\sigma_c^2)/\Lambda_{22}$, which is the stated form; as $n \to \infty$, $\sigma_T^2/n \to 0$ and $\Sigma_{11} \to (1/\sigma_g^2 + \Delta T^{*2}/\sigma_c^2)^{-1}$.)

Exercise 27.4. *Derivation.* Derive the operating characteristic of the test with the prior's shrinkage: write the estimate after one Gauss–Newton step from the prior point, taken centered on the requirement ($g^0 = g^*$), as a linear function of the outcomes, give its mean and variance under a truth with the supply at the set point, show that the mean is $(1 - \pi_g)(g_\mathrm{t} - g^*)$ to first order, and show that at the requirement the characteristic is at most $1 - \kappa$, with equality when the priors carry no information; say what changes when $g^0 \neq g^*$. (With $g^0 = g^*$ the prior rows vanish at $\mathbf{z}^\circ$, so $\hat{\mathbf{z}} - \mathbf{z}^\circ = \mathbf{\Sigma}\mathbf{J}_\mathrm{d}^\mathsf{T}\tilde{\mathbf{e}}$ with $\tilde{\mathbf{e}}$ the normalized departures of the outcomes; under the truth $\mathbb{E}[\hat{\mathbf{z}} - \mathbf{z}^\circ] = \mathbf{\Sigma}\mathbf{\Lambda}_\mathrm{d}(\mathbf{z}_\mathrm{t} - \mathbf{z}^\circ) = (\mathbf{I} - \mathbf{\Sigma}\mathbf{\Lambda}_\mathrm{prior})(\mathbf{z}_\mathrm{t} - \mathbf{z}^\circ)$, whose first entry is $(1 - \Sigma_{11}/\sigma_g^2)(g_\mathrm{t} - g^*)$; $\mathrm{Var} = \mathbf{\Sigma}\mathbf{\Lambda}_\mathrm{d}\mathbf{\Sigma} = \mathbf{\Sigma} - \mathbf{\Sigma}\mathbf{\Lambda}_\mathrm{prior}\mathbf{\Sigma}$, so $v \leq \Sigma_{11}$; at $\mathbf{z}^\circ$ the mean is zero and $\mathrm{OC} = \Phi(-z_\kappa\sqrt{\Sigma_{11}/v}) \leq \Phi(-z_\kappa) = 1 - \kappa$, equal when $\mathbf{\Lambda}_\mathrm{prior} \to \mathbf{0}$. With $g^0 \neq g^*$ the prior row adds $\pi_g(g^0 - g^*)$ to the mean, and a prior centered above the requirement can lift the value at G^* above $1 - \kappa$: 0.0503 for one reading at 30 kW with the prior at 5.5 kW/K.)

Exercise 27.5. *Application.* The pump pair of Chapter 16 has $\mathrm{MTTF}_\mathrm{pair} = 8.5\,\mathrm{yr}$ at the MAP state, equation (16.25), the gradient of $\ln\mathrm{MTTF}_\mathrm{pair}$ of equation (16.44) and the covariance of equation (16.7). On the delta route, a normal on $\ln\mathrm{MTTF}_\mathrm{pair}$, give $\Pr(\mathrm{MTTF}_\mathrm{pair} > 5\,\mathrm{yr} \mid \mathbf{y})$, and find how many more hours of the pair's log would demonstrate $\mathrm{MTTF}_\mathrm{pair} > 5\,\mathrm{yr}$ at an assurance of 0.9, with the MAP state held and each hour adding, at the MAP rates, 2λ expected pump failures to the failure-rate row and $(2 - \beta_\mathrm{cc})\lambda$ expected events to the share row, each an event of information $\beta_\mathrm{cc}(1-\beta_\mathrm{cc})$ in the logit. ($\mathbf{g}^\mathsf{T}\mathbf{\Sigma}\mathbf{g} = 1.023^2/11 + 0.023^2/12 + 0.874^2/1.8 = 0.520$, a spread of 0.721, so $\Pr = \Phi(\ln(8.51/5)/0.721) = \Phi(0.738) = 0.770$, against the 0.85 of the draws in Chapter 16; the assurance needs the spread $0.532/1.2816 = 0.415$; with T' further hours the variance is $1.023^2/(11 + 2\lambda T') + 0.023^2/12 + 0.874^2/(1.8 + 0.09(2 - 0.1)\lambda T')$, which reaches 0.1726 at $T' = 145\,000\,\mathrm{h}$, about 16.6 yr more and 36 further events: the share, which carries 82 % of the variance and gains only 0.09 of information per event, sets the length: a demonstration by record length, as in Chapter 9, for a derived quantity of three entries.)

Exercise 27.6. *Application.* A two-stage test at 60 kW takes one reading, accepts if the tail is at least 0.95, rejects if it is at most 0.05, and otherwise takes four more readings and decides on all five at 0.95. With the linearized estimates of the two stages jointly normal, give its operating characteristic at 5.5, 5.0 and 4.75 kW/K and its expected number of readings at 5.5, and compare it with the single test of five readings. (The estimates after one and after five readings are linear in the first reading, the next four and the supply reading, so they are jointly normal with the covariance of those linear forms; the characteristic is the probability of accepting at the first stage plus that of continuing and accepting at the second. At 5.5: 0.927, of which 0.590 at the first stage, with $1 + 4 \times 0.410 = 2.64$ readings expected; at 5.0: 0.077; at 4.75: 0.0028. The single test gives 0.916, 0.045 and 0.0003. The two-stage test saves readings and gains power, but its acceptance at the requirement exceeds $1 - \kappa$: two looks at the rule each spend part of the consumer's risk, so a sequential plan needs a stricter assurance at each stage.)

Chapter 28

Operating Margins and Capacity Risk

28.1 Margins on the path and on the chiller

The operator of the two data halls of Chapter 17 has Chapter 20's answer for hall 1 at 60 kW: a temperature of 23.0°C and a probability of 0.966 of staying under the 25°C rating, with the supply water held at its set point. The questions of the daily shift are wider. How much margin is there, and how much more load can the hall take? The record gives a line of arithmetic for that too: the conductance of hall 1's path, 5.49 kW/K, times the 12.89 K between the supply and the limit is 70.8 kW, the load at which the hall reaches 25°C, so the move to 60 kW leaves 10.8 kW to spare. In the same week the maintenance engineer brings a ten-day load-bank test of the chiller: on each day a temporary load bank pushed the demand above what the chiller could remove, and the chiller's heat meter read the duty it delivered, which averaged 54.6 kW, against a plate of 60 kW of cooling at a 12°C supply. The plate is the number on the drawings, the two halls draw 52.9 kW between them, and the first reading of the chiller is that it has some 7 kW in hand.

The engineer who must sign for more load asks what the 10.8 kW and the 7 kW rest on. The first is a plug-in value: it says nothing of how sure the hall is to stay under its limit at 70 kW, or of the largest load the plant can take with the 95 % assurance the customer asks for. The second takes the plate at its word, while the ten-day record says something else, of a size that may or may not exceed what one meter and ten days can resolve; and a chiller that cannot deliver its demand lets the supply temperature rise, which Chapter 20 held fixed. The question of this chapter is the engineer's,

how much load the hall and its chiller can take with stated assurance, whether the chiller delivers what its plate says, and what a shortfall of the chiller does to the hall's margin.

The method is that of Chapter 20, the posterior of the reduced state pushed through functions of it. What changes is the kind of answer, a load read off by inverting a tail instead of a probability at a stated load, and the state, which gains the chiller's capacity factor, an entry with its own record joined to the hall through the supply temperature.

Jargon: operating margin, capacity, capacity margin, load–capacity interference, assurance load, plug-in crossing, derating, capacity factor, plate rating, shortfall, onset shortfall, partial expectation

An *operating margin* is the distance from an operating quantity to its limit, here the limit temperature less the hall's temperature at a load; it is a derived quantity of the state, with a posterior. The *capacity* of a path is the load at which the margin is zero, and a *capacity margin* the capacity less a demand. *Load–capacity interference* (stress–strength interference) is the event that a load exceeds the capacity that must carry it, and its probability is the standard measure of capacity risk. The *assurance load* is the largest load at which the probability of staying under the limit reaches a stated assurance level; the *plug-in crossing* is the load at which the plug-in temperature reaches the limit. *Derating* is the delivery by a unit of less than its *plate rating*, the capacity stated by its maker; the *capacity factor* is the delivered capacity divided by the plate rating. A *shortfall* is the part of a demand that a unit cannot deliver, and the *onset shortfall* model takes the supply temperature to rise by the shortfall over the loop's heat-capacity rate, the rise on the first pass of the water. The *partial expectation* of a normal quantity is its expectation over its positive part, the mean of $\max(0, X)$. A capacity margin is not an availability, and a derate is not a failure rate: the first is a distance in kilowatts, the second a degraded state whose posterior says nothing of how often the unit fails.

Key idea

A margin, a capacity and an assurance load are three readings of one event, the load exceeding what the path can carry: the probability that the hall stays under its limit at a load is the probability that the path's capacity exceeds that load, so the assurance load is a quantile of the capacity's posterior. An approximation that linearizes one of those functions need not agree with one that linearizes another, and the full posterior is the check. A unit's rating is an entry of the state, with the plate as its prior and a duty record as its row, and a derate reaches the hall through the supply temperature; the margin of a plant is the smaller of the margins of its path and of its source.

28.1.1 Margins, capacities and derates

Chapter 20 answered one operating question, whether hall 1 stays under its rating at a stated load, as a threshold probability of a scenario temperature. An operator reads the same plant in three further ways. The *margin* at a load is the limit less the temperature there, a derived quantity whose band shows how much room is left and how well it is known. The *capacity* of the path is the load that brings the hall to its limit, a derived quantity in kilowatts, the unit in which loads are planned. And the *assurance load* turns the threshold question around: instead of the probability at a given load, the load at a given probability, the answer to how much load can be taken. The three are functions of the same two entries of the state, the log-conductance g_1 and the supply temperature T_s, and they are tied together by one event, the load exceeding what the path can carry.

The fourth reading concerns the source. The hall rejects its heat to the supply header, and the header is held at its set point only while the chiller can deliver the whole demand. Chapter 20 named the held supply as a limit of its scenario, Section 20.6.6, and said that a row for the chiller's capacity would be needed to relax it. Such a row is a derating question in its own right: is the unit delivering what its plate says? The chiller's delivered capacity becomes an entry of the state, carried as a capacity factor with the plate as its prior, and a record of its duty at full demand is a reading row on it. The derate then reaches the hall through the one quantity they share, the supply temperature.

28.1.2 The plant, the chiller and its record (step 0)

The plant is the two-hall plant of Chapter 17 on its reduced state, Section 17.2.1, as Chapter 20 used it, equation (20.1): the log-conductances g_1, g_2 of the two heat-removal paths, the loads Q_1, Q_2 and the supply temperature T_s, with each hall's temperature $T_i = T_s + Q_i\,\mathrm{e}^{-g_i}$ a function of them and the chiller's heat the sum $Q_1 + Q_2$, the header balance of equation (17.5). The record is the five readings and three priors of Tables 17.1 and 17.2, fitted in Chapters 18 and 19; the hall-1 reading is 17.654545°C, printed 17.655, and every number of this chapter uses the unrounded value.

The chiller is the new node. Its plate rating is $P_{\mathrm{pl}} = 60\,\mathrm{kW}$ of cooling at a 12°C supply, and its delivered capacity is written $P_{\mathrm{pl}}\,c_{\mathrm{ch}}$, with c_{ch} the capacity factor, a sixth entry of the state, dimensionless, equal to one for a chiller that delivers its plate. (The subscript keeps it apart from the constants and costs written c elsewhere in the book.) Its record is the chiller's heat meter of Table 17.1, of standard deviation $\sigma_P = 2\,\mathrm{kW}$, read once a day for $N = 10$ days of a load-bank test. The halls draw about 53 kW, less than the chiller can remove, so on each test day a temporary load bank on the loop added heat until the chiller ran at *full demand*, the one sense of the term in this chapter: a demand at or above what the chiller can remove, so that the meter read the delivered capacity. The ten readings average $\bar{y}_P = 54.6\,\mathrm{kW}$, with the spread of the average $\sigma_P/\sqrt{N} = 0.632\,\mathrm{kW}$, the averaged reading of Chapter 23. The prior on the capacity factor is the plate with the maker's tolerance, $c_{\mathrm{ch}} \sim \mathcal{N}(1,\ 0.08^2)$. The demand at which an undamaged chiller would just hold the supply is its plate, 60 kW, the *plate demand*. Hall 1 at 60 kW puts the chiller at its plate demand only if hall 2's load is taken off this chiller, onto another source, which is a modeling choice of the seventh question (Section 28.3.7); with hall 2 kept on it, a free entry at its present 22.23 kW, the demand is larger. The loop that joins the chiller to the halls is that of Chapter 21, Section 21.1.3, with the heat-capacity rate $\dot{m}c_p = 10\,\mathrm{kW/K}$ pinned, equation (21.4).

Figure 28.1 draws the plant extended for the questions. The green boxes are free nodes and carry entries of the state: the two loads, the supply header with T_s, and now the chiller with c_{ch}. The blue boxes are the halls, balance nodes whose temperatures are functions of the free entries. The solid green arrows are the physical paths, labeled by the log-conductances where they carry one; the dashed gray arrows leave the graph for the readings. The dashed blue box on the left is the scenario node of Chapter 20, hall 1 at a pinned load Q_1^*, here scanned from 40 to 80 kW. The dashed red arrow from the chiller back to the header is the new coupling: the shortfall of the chiller at a demand P^* raises the supply temperature.

28.1.3 The questions (step 1)

The operator asks seven questions of the extended plant, and step 1 writes each as a probability or an expectation over the state, with the unknown on the left of the bar and the evidence on the right: the five readings, the three priors of Chapter 17, the ten-day duty record and the plate prior. Table 28.1 lists them. The first column is the question in engineering language; the second the general form of its category in the notation of Chapter 2, for a quantity $f(\mathbf{z})$ derived from the state, a level f^* of it and an assurance level κ; the third its mathematical form for the plant; the fourth its category; and the last the subsection where it is answered. The first question is Chapter 20's, stated at the loads an operator scans, and the third inverts its tail. The atlas of Appendix A lists the same forms.

28.1.4 The evidence as rows (step 2)

Step 2 extends the graph twice. The first extension is the scenario node of Chapter 20, Section 20.1.4, hall 1 at a pinned load Q_1^*, with the temperature

$$T_1^*(\mathbf{z}; Q_1^*) = T_s + Q_1^*\,\mathrm{e}^{-g_1}, \tag{28.1}$$

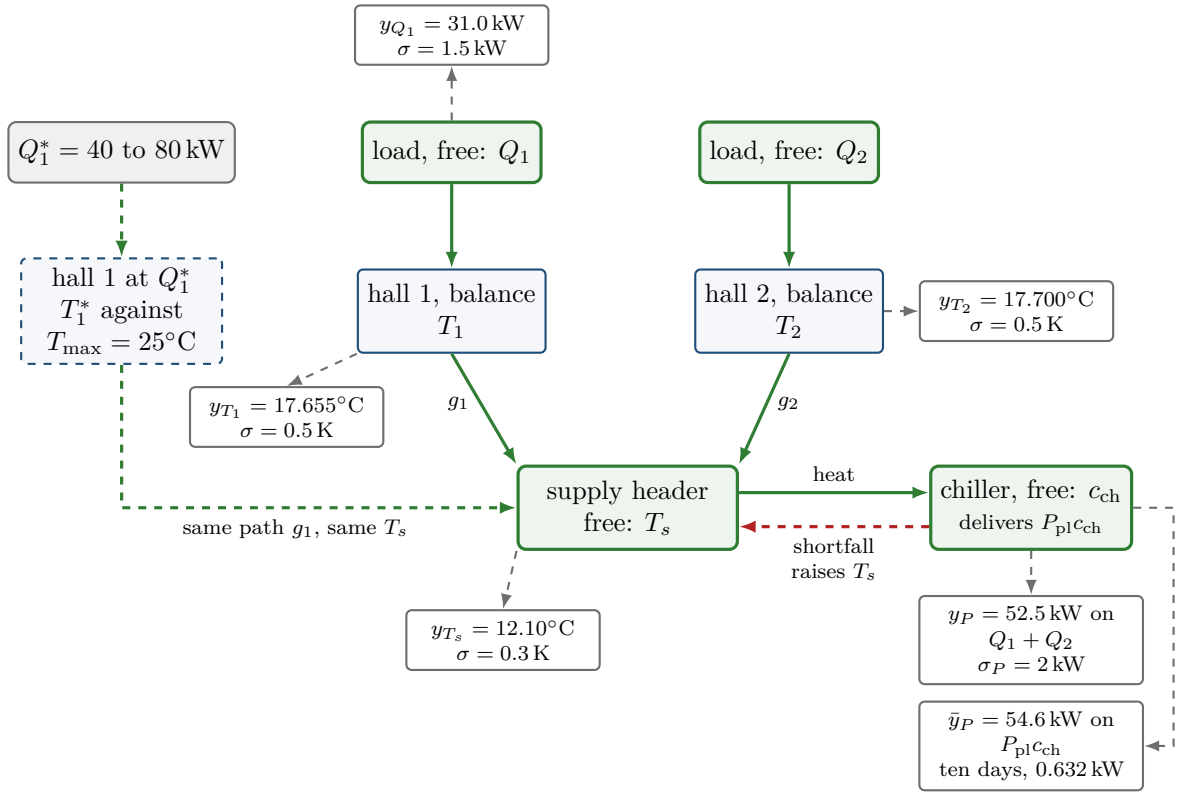


Figure 28.1: The two-hall plant of Chapter 17 extended for the questions of this chapter. Free (green): the loads Q_1, Q_2, the supply temperature T_s, the log-conductances g_1, g_2 that label the hall paths, and the chiller's capacity factor c_{ch}, with delivered capacity $P_{\text{pl}}c_{\text{ch}}$ and plate $P_{\text{pl}} = 60\,\text{kW}$. Balance nodes (blue): the halls. Pinned (gray): the scenario load Q_1^*. Observed (dashed gray arrows): the five readings of Chapter 17 and the ten-day duty record of the chiller's heat meter at full demand; the priors on g_1, g_2, T_s and c_{ch} are rows not drawn. Dashed blue: the scenario node, sharing g_1 and T_s. Dashed red: the onset shortfall $(P^* - P_{\text{pl}}c_{\text{ch}})^+/\dot{m}c_p$ by which a demand P^* above the delivered capacity raises the supply, with $\dot{m}c_p = 10\,\text{kW/K}$ pinned.

Table 28.1: The questions asked of the two-hall plant and its chiller, the general form of their category, their mathematical form for the plant, their category and where each is answered. $T_1^* = T_s + Q_1^* \mathrm{e}^{-g_1}$ is hall 1's temperature at a scenario load Q_1^* with the supply held; $m_1^* = T_{\max} - T_1^*$ its margin; $Q_{\mathrm{cap}} = (T_{\max} - T_s)\,\mathrm{e}^{g_1}$ the path's capacity; c_{ch} the chiller's capacity factor, M_{f} the drawing that frees it and M_{p} the one that pins it at the plate; ΔT_s the onset rise of the supply at a demand P^*; $T_{\max} = 25°\mathrm{C}$, $\kappa = 0.95$.

Engineering question	General form	Mathematical form	Category	Answered in
Will the hall stay under the limit at 60 kW?	$\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y})$	$\Pr(T_1^* < T_{\max} \mid \mathbf{y})$ at $Q_1^* = 40$ to 80 kW	prediction under a threshold	28.3.1
How much margin is there, and how well is it known?	$\Pr(f(\mathbf{z}) \leq f^* \mid \mathbf{y})$; location and interval	$p(m_1^* \mid \mathbf{y})$, $m_1^* = T_{\max} - T_1^*$	estimation of a derived quantity	28.3.2
How much load can the hall take with 95 % assurance?	the largest level with $\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y}) \geq \kappa$	$Q_\kappa = \max\{Q_1^* : \Pr(T_1^* < T_{\max} \mid \mathbf{y}) \geq \kappa\}$	threshold inverted (a design limit)	28.3.3
What is the path's capacity, and is it at least the design load?	location and interval; $\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y})$	$p(Q_{\mathrm{cap}} \mid \mathbf{y})$; $\Pr(Q_{\mathrm{cap}} \geq 60\,\mathrm{kW} \mid \mathbf{y})$	estimation; prediction under a threshold	28.3.4
Is the chiller delivering what its plate says?	location; $\Pr(f(\mathbf{z}) \leq f^* \mid \mathbf{y})$; B_{AB}	$p(c_{\mathrm{ch}} \mid \mathbf{y})$; $\Pr(c_{\mathrm{ch}} < 0.95 \mid \mathbf{y})$; $\ln B_{\mathrm{fp}} = \ln[p(\mathbf{y} \mid M_{\mathrm{f}})/p(\mathbf{y} \mid M_{\mathrm{p}})]$	estimation; threshold; hypothesis comparison	28.3.5
How much load does the chiller have in hand now?	location and interval	$p(P_{\mathrm{hr}} \mid \mathbf{y})$, $P_{\mathrm{hr}} = P_{\mathrm{pl}} c_{\mathrm{ch}} - Q_1 - Q_2$	estimation of a derived quantity	28.3.6
What does the derate do to the margin at full demand?	$\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y})$	$\Delta T_s = (P^* - P_{\mathrm{pl}} c_{\mathrm{ch}})^+ / \dot{m} c_p$; $\Pr(T_1^* + \Delta T_s < T_{\max} \mid \mathbf{y})$ at $Q_1^* = P^* = 60$ kW, hall 2 off this chiller; the assurance load there	prediction under a threshold, through the loop	28.3.7

equation (20.2) with the load left free to be scanned. It adds no row, and from it come three derived quantities of the state, each a function of g_1 and T_s only: the margin, the capacity and, through the tail, the assurance load.

The second extension is the chiller's entry and its two rows. The state becomes

$$\mathbf{z} = (g_1,\ g_2,\ Q_1,\ Q_2,\ T_s,\ c_{\mathrm{ch}})^{\mathsf{T}}, \tag{28.2}$$

the five entries of equation (20.1) and the capacity factor. The duty record is a reading row on the delivered capacity and the plate a prior row on the factor, each divided by its standard deviation:

$$r_9(\mathbf{z}) = \frac{P_{\mathrm{pl}}\, c_{\mathrm{ch}} - \bar{y}_P}{\sigma_P/\sqrt{N}}, \qquad r_{10}(\mathbf{z}) = \frac{c_{\mathrm{ch}} - 1}{\sigma_{\mathrm{pl}}}, \qquad \sigma_{\mathrm{pl}} = 0.08, \tag{28.3}$$

added to the eight rows of equation (17.15). The demand on the chiller at a scenario is the header balance at the scenario loads, $P^* = Q_1^* + Q_2^*$; it is pinned with the loads, and adds no row either. Which row acts on which entry fixes where the Jacobian $\mathbf{J}$ can be nonzero, marked $*$:

$$\mathbf{J} = \begin{bmatrix} * & 0 & * & 0 & * & 0 \\ 0 & * & 0 & * & * & 0 \\ 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & * & * & 0 & 0 \\ 0 & 0 & * & 0 & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & * \end{bmatrix} \begin{matrix} \text{hall-1 reading} \\ \text{hall-2 reading} \\ \text{supply reading} \\ \text{heat meter} \\ \text{load meter} \\ \text{prior on } g_1 \\ \text{prior on } g_2 \\ \text{prior on } T_s \\ \text{duty record} \\ \text{plate prior} \end{matrix} \qquad \begin{aligned} \mathbf{g}_T^{\mathsf{T}} &= \begin{bmatrix} * & 0 & 0 & 0 & * & 0 \end{bmatrix}, \\ \mathbf{g}_{\mathrm{cap}}^{\mathsf{T}} &= \begin{bmatrix} * & 0 & 0 & 0 & * & 0 \end{bmatrix}, \\ \mathbf{g}_{\mathrm{hr}}^{\mathsf{T}} &= \begin{bmatrix} 0 & 0 & * & * & 0 & * \end{bmatrix}, \\ \mathbf{g}_{\Delta}^{\mathsf{T}} &= \begin{bmatrix} * & 0 & 0 & 0 & * & * \end{bmatrix}, \end{aligned} \tag{28.4}$$

with the columns in the order of equation (28.2); the gradients are those of the scenario temperature (and of the margin), of the capacity, of the chiller's headroom and of the scenario temperature with the shortfall. The upper left block is Chapter 20's pattern (20.3). The capacity factor's column meets only its own two rows, and no row of the hall meets it: the five readings, taken with the demand below the capacity, say nothing of the capacity. The chiller and the hall meet only in the last two gradients, the headroom and the shortfall, functions asked about and not rows.

The modeling choices. Five choices make the scenario node and the chiller's rows, and each buys something and can fail. The scenario node keeps the two choices of Chapter 20, Section 20.3.1: the path keeps at the scenario load the conductance inferred at the present load, and the chiller holds the supply temperature. The first lets the node share g_1 and fails if the coil's conductance changes with the flow; it is a limit of the model (Section 28.6.5). The second lets the node share T_s and fails whenever the chiller cannot deliver the demand; it is the assumption the seventh question removes (Section 28.3.7), and every margin of the first four questions is the path's margin under it. The ten load-bank days are taken as comparable, each at full demand, so that every reading is of the delivered capacity: this is what makes the record one row on c_{ch}, and it fails if on some days the load bank left the demand below the capacity, when those readings read the demand and the average understates the capacity (Section 28.6.5). The meter's daily errors are taken as independent and unbiased, so that the average has the spread $\sigma_P/\sqrt{N}$: this buys the factor $\sqrt{10}$, and it fails for a calibration error, common to every day, which no average removes and which a derate mimics exactly; Section 28.4.4 frees it as Chapter 22 frees a bias. The plate prior $\mathcal{N}(1,\ 0.08^2)$ is the maker's tolerance: it gives the drawing M_{f} a proper prior, which the Savage–Dickey ratio needs, and Section 28.4.5 moves it.

28.2 The posterior of the state (step 3)

Every question of Table 28.1 asks for a function of the state (28.2), and the objects of the fit recur in the specialized forms: the MAP state $\hat{\mathbf{z}}$, the information matrix $\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$ of the normalized rows, the Laplace covariance $\mathbf{\Sigma} = \mathbf{\Lambda}^{-1}$, and the gradient $\mathbf{g}$ of each derived quantity at $\hat{\mathbf{z}}$, with the variance $\mathbf{g}^{\mathsf{T}}\mathbf{\Sigma}\mathbf{g}$. The delta method and the full posterior by reweighted draws are those of Chapter 20, Sections 20.3.1 and 20.2.4. Every answer is reported from the posterior of the state, by the delta method on $(\hat{\mathbf{z}}, \mathbf{\Sigma})$, and the full posterior by reweighted draws gives the value set beside it, with the gap. All of them are read from one posterior, that of the six-entry state the ten rows of step 2 give. This section derives it and computes it for the plant; Section 28.3 reads every answer from it.

28.2.1 The posterior from the rows

The rows are the ten of Section 28.1.4: the eight of the reduced state, equation (17.15), displayed in Chapter 20, Section 20.2.1, and the two of the chiller, equation (28.3). The objective is $\mathcal{L}(\mathbf{z}) = \frac{1}{2}\mathbf{r}^{\mathsf{T}}\mathbf{r}$, and because no row involves both the hall's entries and the capacity factor it splits,

$$\mathcal{L}(\mathbf{z}) = \mathcal{L}_{\mathrm{h}}(g_1, g_2, Q_1, Q_2, T_s) + \tfrac{1}{2}(r_9^2 + r_{10}^2)(c_{\mathrm{ch}}), \tag{28.5}$$

with $\mathcal{L}_{\mathrm{h}}$ the objective of Chapters 18 and 19. The posterior, proportional to $\mathrm{e}^{-\mathcal{L}}$, is therefore the product of the hall's posterior and the capacity factor's, exactly, not only under the Laplace approximation.

The capacity factor's two rows (28.3) act on its column alone, by the pattern (28.4), so its posterior is a one-entry problem in state-vector form: the state (c_{ch}), the residual vector $(r_9, r_{10})^{\mathsf{T}}$, the 2×1 Jacobian

$$\mathbf{J}_{\mathrm{c}} = \begin{bmatrix} P_{\mathrm{pl}}\sqrt{N}/\sigma_P \\ 1/\sigma_{\mathrm{pl}} \end{bmatrix}, \qquad \mathbf{\Lambda}_{\mathrm{c}} = \mathbf{J}_{\mathrm{c}}^{\mathsf{T}}\mathbf{J}_{\mathrm{c}} = \frac{NP_{\mathrm{pl}}^2}{\sigma_P^2} + \frac{1}{\sigma_{\mathrm{pl}}^2}, \tag{28.6}$$

constant, since both rows are linear. One Gauss–Newton step from any starting point lands on the minimum, and the Laplace covariance $\Sigma_{66} = \mathbf{\Lambda}_{\mathrm{c}}^{-1}$ is the exact posterior variance: for linear Gaussian rows and a Gaussian prior the posterior is normal, the precisions adding as in Chapter 3, equation (3.26). The same posterior is the rank-one update of Chapter 4, equation (4.63), applied to the plate prior with the record as the new reading: with $\mathbf{h} = P_{\mathrm{pl}}$ the reading's sensitivity to the entry, the predictive variance of the averaged reading is $V_{\mathrm{c}} = P_{\mathrm{pl}}^2\sigma_{\mathrm{pl}}^2 + \sigma_P^2/N$, the one-entry gain (the $\mathbf{K}$ of that equation) $\varphi_{\mathrm{c}} = \sigma_{\mathrm{pl}}^2 P_{\mathrm{pl}}/V_{\mathrm{c}}$, and

$$\hat{c}_{\mathrm{ch}} = 1 + \varphi_{\mathrm{c}}\,(\bar{y}_P - P_{\mathrm{pl}}), \qquad \Sigma_{66} = \sigma_{\mathrm{pl}}^2 - V_{\mathrm{c}}\varphi_{\mathrm{c}}^2, \tag{28.7}$$

the innovation $\bar{y}_P - P_{\mathrm{pl}}$ being the part of the record the plate did not predict. The delivered capacity $P_{\mathrm{pl}}c_{\mathrm{ch}}$ is the entry scaled, with spread $P_{\mathrm{pl}}\sqrt{\Sigma_{66}}$.

The 10×6 Jacobian $\mathbf{J} = \mathbf{J}_{\mathrm{e}}\mathbf{N}$ of the reduced state, equation (17.18), is block diagonal, an 8×5 block for the hall and the 2×1 block (28.6); the normal equations $(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^{\mathsf{T}}\mathbf{r}$ separate into the hall's, iterated in Chapter 18, and the capacity factor's single linear step; and

$$\mathbf{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1} = \begin{bmatrix} \mathbf{\Sigma}_{\mathrm{h}} & \mathbf{0} \\ \mathbf{0}^{\mathsf{T}} & \Sigma_{66} \end{bmatrix}, \qquad \Sigma_{66} = \mathbf{\Lambda}_{\mathrm{c}}^{-1}, \tag{28.8}$$

with $\mathbf{\Sigma}_{\mathrm{h}}$ the committed covariance of Chapter 19, Table 19.3, the free-entry marginal of the full state's covariance in the limit $\sigma_p \to 0$, equation (17.19). The two hall reading rows are nonlinear on the reduced state, and Chapter 18 kept the Gauss–Newton curvature at $\hat{\mathbf{z}}$; the chiller's rows are linear and carry no curvature term.

The full posterior is reached as in Chapter 20, Section 20.2.4: $N_{\mathrm{d}} = 200\,000$ draws of the hall's Laplace Gaussian, reweighted by the ratio (20.10) to the hall's exact posterior, equation (20.11), with the effective sample size (20.12). Because the posterior factorizes, the capacity factor needs no weight: each hall draw is joined with an independent draw of c_{ch} from its exact normal. A probability is the weighted fraction of draws in its event, with the standard error of equation (20.12); a quantile, the assurance load among them, is read where the weights accumulated in the order of the values reach the level, and its standard error is estimated by resampling the draws with their weights. Every answer is computed from $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ and reported; the draws give the value set beside it.

28.2.2 The posterior of the hall and of the chiller

The plant's inputs are the fit of Chapters 18 and 19, the ten-day record and the plate prior. The hall's block is the fit of Chapter 18, Table 18.4, with the covariance of Chapter 19, as Chapter 20, Section 20.2.2, quotes it: $\hat{g}_1 = 1.70297$ ($\hat{G}_1 = 5.4903\,\mathrm{kW/K}$), $\hat{g}_2 = 1.39993$, $\hat{Q}_1 = 30.688\,\mathrm{kW}$, $\hat{Q}_2 = 22.231\,\mathrm{kW}$, $\hat{T}_s = 12.1115^{\circ}\mathrm{C}$, and the block that every hall-1 question uses, equation (20.4), $\Sigma_{11} = 0.011382$, $\Sigma_{15} = 0.012974\,\mathrm{K}$, $\Sigma_{55} = 0.078433\,\mathrm{K}^2$, with $\mathrm{corr}(g_1, T_s) = 0.434$. The thermal resistance of hall 1's path is $\mathrm{e}^{-\hat{g}_1} = 0.18214\,\mathrm{K/kW}$, and the room between the supply and the limit is $a = 25 - 12.1115 = 12.8885\,\mathrm{K}$.

The capacity factor's fit starts at the plate, $c_{\mathrm{ch}} = 1$, where the residual vector (28.3) is $(r_9, r_{10}) = ((60 - 54.6)/0.6325,\ 0) = (8.5381,\ 0)$ and $\mathcal{L} = 36.45$. The Jacobian (28.6) is $(60/0.6325,\ 1/0.08)^{\mathsf{T}} = (94.868,\ 12.5)^{\mathsf{T}}$, and the information

$$\boldsymbol{\Lambda}_{\mathrm{c}} = 94.868^2 + 12.5^2 = 9000 + 156.25 = 9156.25.$$

The normal equation gives the step $\Delta c_{\mathrm{ch}} = -(94.868 \times 8.5381)/9156.25 = -0.088464$, and at $c_{\mathrm{ch}} = 0.911536$ the residuals are $(0.1457,\ -1.1058)$, $\mathcal{L} = 0.622$, and a second step is zero: one step, since the rows are linear. The covariance is $\Sigma_{66} = 1/9156.25 = 1.0922 \times 10^{-4}$, a spread $\sqrt{\Sigma_{66}} = 0.010451$. The record carries $9000/9156.25 = 98.3\,\%$ of the information and the plate $1.7\,\%$. The rank-one form (28.7) gives the same: $V_{\mathrm{c}} = 60^2 \times 0.08^2 + 0.4 = 23.04 + 0.40 = 23.44\,\mathrm{kW}^2$, $\varphi_{\mathrm{c}} = 0.0064 \times 60/23.44 = 0.016382\,\mathrm{kW}^{-1}$, the innovation $54.6 - 60 = -5.4\,\mathrm{kW}$, $\hat{c}_{\mathrm{ch}} = 1 - 0.016382 \times 5.4 = 0.911536$ and $\Sigma_{66} = 0.0064 - 23.44 \times 0.016382^2 = 1.0922 \times 10^{-4}$. The covariance (28.8) is the hall's block with this sixth diagonal entry and zeros between them.

28.3 The answers (step 4)

Each question of Table 28.1 is answered from the posterior of Section 28.2, in the table's order: its engineering question, the general form of its category, the form it takes for the plant, the quantities of the posterior it uses, and the plant's numbers. The questions name the limit $T_{\max} = 25^{\circ}\mathrm{C}$, the assurance level $\kappa = 0.95$ and the heat-capacity rate $\dot{m}c_p = 10\,\mathrm{kW/K}$. Every answer below is a formula evaluated on the posterior of the state, by the delta method on $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$, and it is the value reported; the full posterior by reweighted draws gives the value set beside it in the same answer, with the gap.

28.3.1 The threshold at a scenario load

Engineering question. Will the hall stay under the limit at 60 kW? And, since the operator plans loads, at the other loads from 40 to 80 kW that the hall might be asked to carry.

General form. The threshold form of Chapter 2, equation (2.12),

$$\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) \geq f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

with the upper limit written as $f = -T_1^*$, $f^* = -T_{\max}$.

Specialized form. The question is equation (2.12) with $f(\mathbf{z}) = -T_1^*(\mathbf{z}; Q_1^*)$, the scenario temperature at the load Q_1^*, $f^* = -T_{\max} = -25°\mathrm{C}$, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$. The scenario node carries the two choices of Section 28.1.4. The reported route is the normal tail of the linearized output on that posterior. It buys one evaluation of Φ per load, and it errs where the limit lies near the center of the band, as Chapter 20 found, Section 20.4.3; the full posterior gives the value beside it at every load.

At a load Q_1^* the event is $T_1^*(\mathbf{z}; Q_1^*) < T_{\max}$ and the probability is the integral (20.17) with the load left free. The gradient of the scenario temperature (28.1) is that of equation (20.14) with a sixth entry,

$$\mathbf{g}_T(Q_1^*) = \left(-Q_1^* \, \mathrm{e}^{-\hat{g}_1},\ 0,\ 0,\ 0,\ 1,\ 0\right)^{\mathsf{T}}, \tag{28.9}$$

and the delta tail is equation (20.18), $\Phi\big((T_{\max} - T_1^*(\hat{\mathbf{z}}))/\sqrt{\mathbf{g}_T^{\mathsf{T}} \boldsymbol{\Sigma} \mathbf{g}_T}\big)$. By the budget (20.20) the variance at a load is a quadratic in the load,

$$\mathbf{g}_T^{\mathsf{T}} \boldsymbol{\Sigma} \mathbf{g}_T = Q_1^{*2} \, \mathrm{e}^{-2\hat{g}_1} \, \Sigma_{11} - 2 \, Q_1^* \, \mathrm{e}^{-\hat{g}_1} \, \Sigma_{15} + \Sigma_{55}, \tag{28.10}$$

whose leading term, the conductance's, grows as the square of the temperature rise $Q_1^* \mathrm{e}^{-\hat{g}_1}$, the lever of Chapter 20, while the supply's term stays fixed; the cross term grows with the rise and stays negative.

Building blocks. The block of $\boldsymbol{\Sigma}$ in the rows and columns of g_1 and T_s, equation (20.4); the MAP values $\hat{g}_1$, $\hat{T}_s$; Φ. For the values beside them, the reweighted draws of Section 28.2.1.

28.3.2 The margin and its band

Engineering question. How much margin is there, and how well is it known? The operator asks how many kelvin stand between the hall and its rating at a load, and how sure that distance is.

General form. The estimation form, equations (2.9) to (2.11), for a derived quantity: its posterior, summarized by a location and the central 95 % interval.

Specialized form. The question is equations (2.9) to (2.11) with $f(\mathbf{z}) = m_1^*(\mathbf{z}; Q_1^*)$, the margin below, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$, under which the delta method gives its center $m_1^*(\hat{\mathbf{z}})$ and spread. One choice enters here: the hall has one limit, $T_{\max} = 25°\mathrm{C}$ on its temperature. It buys a scalar margin in kelvin; it fails for equipment with several limits, a temperature and a temperature rise, say, whose joint margin is the intersection of several events (Section 28.6.5 and Exercise 28.5).

The margin at a load is

$$m_1^*(\mathbf{z}; Q_1^*) = T_{\max} - T_1^*(\mathbf{z}; Q_1^*) = T_{\max} - T_s - Q_1^* \, \mathrm{e}^{-g_1}, \tag{28.11}$$

in kelvin; the symbol m_1^* is not Chapter 20's conditional mean $m(x)$. It is an affine function of the scenario temperature, with gradient $-\mathbf{g}_T$ and the same variance: its posterior is that of T_1^* carried by the map $t \mapsto T_{\max} - t$, its delta band is $T_{\max}$ less the band (20.16) with the ends exchanged, and its full-posterior quantiles are the reflected quantiles of T_1^*, by the rule of equation (2.11) for a decreasing function. The threshold question is the event $m_1^* > 0$, so the first two questions read one posterior, the first as its mass above zero and the second as its location and band. The margin's spread grows with the load as the square root of the quadratic (28.10); its center falls linearly, by $\mathrm{e}^{-\hat{g}_1}$ kelvin per kilowatt, the thermal resistance of the path.

Building blocks. Those of the threshold; for the values beside them, the reflected quantiles of the reweighted draws.

The margin by load. Table 28.2 gives the first two questions at five loads by the delta method. Each row is a load; the columns are the scenario temperature at the MAP state, the temperature rise $Q_1^* \mathrm{e}^{-\hat{g}_1}$ (the magnitude of the gradient's first entry), the spread $\mathrm{sd}(m_1^* \mid \mathbf{y})$ from equation (28.10), the margin $m_1^*(\hat{\mathbf{z}})$ with its central 95 % band, the standardized distance $m_1^*(\hat{\mathbf{z}})/\mathrm{sd}$ and the probability of staying under the limit. One row of the arithmetic, at 70 kW: the rise is $70 \times 0.18214 = 12.750$ K, the variance $12.750^2 \times 0.011382 + 0.078433 - 2 \times 12.750 \times 0.012974 = 1.8503 + 0.0784 - 0.3308 = 1.5979\,\mathrm{K}^2$, the spread 1.264 K, the margin $12.8885 - 12.750 = 0.139$ K, and $\Phi(0.139/1.264) = \Phi(0.110) = 0.544$.

Table 28.2: Hall 1's margin $m_1^* = T_{\max} - T_1^*$ by load, by the delta method, with the supply held: the scenario temperature at $\hat{\mathbf{z}}$, the temperature rise, the margin's spread, its value at $\hat{\mathbf{z}}$ and central 95 % band, the standardized distance and the probability of staying under $T_{\max} = 25$°C, with the full posterior's probability beside it. The row at 60 kW is Chapter 20's, Section 20.3.1.

Q_1^* [kW]	$T_1^*(\hat{\mathbf{z}})$ [°C]	rise [K]	sd [K]	$m_1^*(\hat{\mathbf{z}})$ [K]	95 % band [K]	m_1^*/sd	$\Pr(T_1^* < 25)$	full
40	19.397	7.286	0.703	5.603	[4.23, 6.98]	7.975	1.0000	1.0000
50	21.219	9.107	0.887	3.782	[2.04, 5.52]	4.265	0.99999	0.99996
60	23.040	10.928	1.074	1.960	[−0.15, 4.07]	1.824	0.966	0.966
70	24.861	12.750	1.264	0.139	[−2.34, 2.62]	0.110	0.544	0.577
80	26.683	14.571	1.455	−1.683	[−4.53, 1.17]	−1.157	0.124	0.135

The margin falls by 1.82 K for every 10 kW, and its spread grows from 0.70 K at 40 kW to 1.07 at 60 and 1.46 K at 80 kW. The values beside them come from the full posterior, 200 000 reweighted draws: the probability of staying under the limit is the same at 60 kW, 0.966, and higher by 0.034 at 70 kW and 0.011 at 80 kW; the full median margin is 2.05 K at 60 kW, band [−0.16, 4.03] K, and lies 0.05 to 0.13 K above the delta center at every load. The state route understates the probability near the plug-in crossing because the full posterior's center lies on the safe side of the delta center, and its temperature is skewed, +0.17, which works the other way in the far tail; the two effects cancel near 1.8 standard deviations, where the 60 kW question lies. No decision at 60 kW changes.

28.3.3 The assurance load

Engineering question. How much load can the hall take with 95 % assurance? The customer's assurance level is fixed, and the operator asks for the load.

General form. The threshold form (2.12) with its acceptance rule, $\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y}) \geq \kappa$, read the other way: the requirement depends on a design variable, here the load, and the question asks for the largest value of it that the rule accepts.

Specialized form. The question is equation (2.12) with its acceptance rule, with $f(\mathbf{z}) = -T_1^*(\mathbf{z}; Q_1^*)$, $f^* = -T_{\max}$, $\kappa = 0.95$, the load Q_1^* the design variable, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$. One choice enters here: the assurance level $\kappa = 0.95$ is taken as the customer's, kept apart from the limit as Chapter 2 keeps it. It buys a load in kilowatts; the load moves with κ, and Section 28.3.3 gives it at three levels so that the price of each is visible.

Write the probability of staying under the limit as a function of the load, $p_{\mathrm{s}}(Q_1^*) = \Pr(T_1^*(\mathbf{z}; Q_1^*) < T_{\max} \mid \mathbf{y})$. It is nonincreasing in the load: for $Q_1^{*\prime} > Q_1^*$ and any state, $\mathrm{e}^{-g_1} > 0$

gives $T_1^*(\mathbf{z}; Q_1^{*\prime}) > T_1^*(\mathbf{z}; Q_1^*)$, so the event at the larger load is contained in the event at the smaller,

$$\{T_s + Q_1^{*\prime}\,\mathrm{e}^{-g_1} < T_{\max}\} \subset \{T_s + Q_1^*\,\mathrm{e}^{-g_1} < T_{\max}\},$$

and the probability of a contained event is no larger. It is strictly decreasing and continuous wherever the posterior has a positive density along the boundary curve, which a normal or a reweighted normal has. The assurance load is therefore the unique root

$$Q_\kappa : \quad p_{\mathrm{s}}(Q_\kappa) = \kappa, \tag{28.12}$$

and every load below it meets the rule. Under the delta method, by equation (28.10),

$$p_{\mathrm{s}}(Q_1^*) \approx \Phi\!\left(\frac{a - x}{\sqrt{x^2\Sigma_{11} - 2x\Sigma_{15} + \Sigma_{55}}}\right),$$

with $a = T_{\max} - \hat{T}_s$ the room between the supply and the limit and $x = Q_1^*\mathrm{e}^{-\hat{g}_1}$ the temperature rise. Setting it equal to κ with $z_\kappa = \Phi^{-1}(\kappa)$, and squaring, which keeps the root when $a - x > 0$, as it is for $\kappa > \frac{1}{2}$,

$$(a - x)^2 = z_\kappa^2(x^2\Sigma_{11} - 2x\Sigma_{15} + \Sigma_{55}),$$

a quadratic in the rise,

$$\begin{gathered} Ax^2 - 2Bx + C = 0, \qquad x_\kappa = \frac{B - \sqrt{B^2 - AC}}{A}, \\ A = 1 - z_\kappa^2\Sigma_{11}, \quad B = a - z_\kappa^2\Sigma_{15}, \quad C = a^2 - z_\kappa^2\Sigma_{55}, \end{gathered} \tag{28.13}$$

the smaller root, and $Q_\kappa = x_\kappa\,\mathrm{e}^{\hat{g}_1}$. At $\kappa = \frac{1}{2}$, $z_\kappa = 0$ and the root is $x = a$: the load at which the plug-in temperature reaches the limit, the *plug-in crossing* $a\,\mathrm{e}^{\hat{g}_1}$. The difference between the plug-in crossing and Q_κ is the load the uncertainty costs the operator at assurance κ.

Building blocks. a, Σ_{11}, Σ_{15}, Σ_{55}, $\hat{g}_1$ and the quantile $z_{0.95} = 1.6449$; for the values beside them, the weighted quantile of the reweighted draws of the capacity (Section 28.3.4).

The assurance load of hall 1. With $z_{0.95} = \Phi^{-1}(0.95) = 1.6449$, $z_{0.95}^2 = 2.7055$, the coefficients of equation (28.13) are

$$\begin{aligned} A &= 1 - 2.7055 \times 0.011382 = 0.969204, \\ B &= 12.8885 - 2.7055 \times 0.012974 = 12.853449, \\ C &= 12.8885^2 - 2.7055 \times 0.078433 = 165.902514, \end{aligned}$$

so $B^2 - AC = 165.211162 - 160.793445 = 4.417717$ and

$$x_{0.95} = \frac{12.853449 - 2.101837}{0.969204} = 11.0932\,\mathrm{K}, \qquad Q_{0.95} = 11.0932 \times 5.49025 = 60.90\,\mathrm{kW}.$$

At that load the scenario temperature is 23.205°C with a spread of 1.0915 K, and $(25 - 23.205)/1.0915 = 1.645$, as it must be. The plug-in crossing is $a\,\hat{G}_1 = 12.8885 \times 5.49025 = 70.76$ kW. Table 28.3 gives the assurance load at four levels: each row is an assurance level, with its normal quantile, the delta load from equation (28.13) and the load the uncertainty costs against the plug-in crossing.

At $\kappa = 0.95$ the delta assurance load is 60.90 kW, 9.86 kW below the plug-in crossing; from 0.95 to 0.99 it falls by another 3.29 kW. The values beside them come from the full posterior, the weighted quantiles of the capacity over the reweighted draws: $Q_{0.95} = 60.99 \pm 0.02$ kW, 0.08 kW above the delta inversion, four standard errors; 57.32 kW at 0.99, 63.06 kW at 0.90 and 71.37 kW at 0.50, the capacity's median. The delta inversion is high at 0.99 and low at the others, crossing the full posterior near 1.8 standard deviations; at 0.95 both put the design load of 60 kW inside the assured range.

Table 28.3: The assurance load Q_κ of hall 1 by the delta inversion, equation (28.13), with the supply held, and its distance from the plug-in crossing 70.76 kW.

Assurance κ	$z_\kappa = \Phi^{-1}(\kappa)$	Q_κ [kW]	plug-in crossing less Q_κ [kW]
0.99	2.3263	57.61	13.15
0.95	1.6449	60.90	9.86
0.90	1.2816	62.83	7.94
0.50	0	70.76	0

28.3.4 The capacity of the path

Engineering question. What is the path's capacity, and is it at least the design load? The operator asks the question in kilowatts: the load at which hall 1 reaches its limit, and whether it reaches 60 kW.

General form. The estimation form for a derived quantity, equations (2.10) and (2.11), and the threshold form (2.12) with $f = Q_{\text{cap}}$ and $f^* = 60\,\text{kW}$.

Specialized form. The question is equations (2.10) and (2.11) with $f(\mathbf{z}) = Q_{\text{cap}}(\mathbf{z})$ below, and equation (2.12) with the same f and $f^* = 60\,\text{kW}$, $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$. One choice enters here: the capacity is that of the path, the load that brings the hall to its limit with the conductance and the supply as they are. It buys a function of two entries; it is not the plant's capacity, which is the smaller of the path's and the chiller's, and it fails as an answer to how much load the plant can take whenever the chiller binds first, which the seventh question decides.

Setting $T_1^* = T_{\max}$ in equation (28.1) and solving for the load,

$$Q_{\text{cap}}(\mathbf{z}) = (T_{\max} - T_s)\,\mathrm{e}^{g_1}, \qquad \mathbf{g}_{\text{cap}} = \big(\,Q_{\text{cap}}(\hat{\mathbf{z}}),\ 0,\ 0,\ 0,\ -\mathrm{e}^{\hat{g}_1},\ 0\big)^{\mathsf{T}}, \tag{28.14}$$

in kW, the product of two functions of two correlated entries, the room $T_{\max} - T_s$ and the conductance $G_1 = \mathrm{e}^{g_1}$. By equation (20.20) its delta variance is

$$\mathbf{g}_{\text{cap}}^{\mathsf{T}} \boldsymbol{\Sigma} \mathbf{g}_{\text{cap}} = Q_{\text{cap}}(\hat{\mathbf{z}})^2\,\Sigma_{11} + \hat{G}_1^2\,\Sigma_{55} - 2\,Q_{\text{cap}}(\hat{\mathbf{z}})\,\hat{G}_1\,\Sigma_{15}, \tag{28.15}$$

and its cross term is negative for the reason of Section 20.3.3: the two entries are positively correlated and enter with opposite signs.

One event, two functions. The capacity and the scenario temperature are two functions of the same state, and their tails describe one event. Since $\mathrm{e}^{g_1} > 0$, for every state and every load Q_1^*,

$$Q_{\text{cap}}(\mathbf{z}) \ge Q_1^* \iff (T_{\max} - T_s)\,\mathrm{e}^{g_1} \ge Q_1^* \iff T_s + Q_1^*\,\mathrm{e}^{-g_1} \le T_{\max} \iff T_1^*(\mathbf{z}; Q_1^*) \le T_{\max}, \tag{28.16}$$

so that under any posterior

$$\Pr(Q_{\text{cap}} \ge Q_1^* \mid \mathbf{y}) = p_{\text{s}}(Q_1^*), \qquad Q_\kappa = F_{\text{cap}}^{-1}(1 - \kappa),$$

with F_{cap} the posterior distribution function of the capacity: the probability of staying under the limit at a load is the survival function of the capacity at that load, and the assurance load is the $(1 - \kappa)$ quantile of the capacity. Load–capacity interference at the load Q_1^* is the event $Q_{\text{cap}} < Q_1^*$, and its probability is $1 - p_{\text{s}}(Q_1^*)$. The full posterior, which evaluates the event itself, returns the same number by either function; at 60 kW that number is Chapter 20's, and it is cited, not recomputed. The delta method does not: it linearizes T_1^* in one case and Q_{cap} in the other, two functions related by a monotone but nonlinear map, and the tangent plane of the

one is not the image of the tangent plane of the other. The quantiles of the full posterior map through the monotone function, Chapter 9's quantile map, Section 9.3.2; the delta bands do not. The two delta answers, $\Phi\big((T_{\max}-T_1^*(\hat{\mathbf{z}}))/\mathrm{sd}(T_1^* \mid \mathbf{y})\big)$ and $\Phi\big((Q_{\mathrm{cap}}(\hat{\mathbf{z}})-Q_1^*)/\mathrm{sd}(Q_{\mathrm{cap}} \mid \mathbf{y})\big)$, agree at one load only, the plug-in crossing, where both are $\frac{1}{2}$, and differ elsewhere; the delta assurance load from the capacity is $Q_{\mathrm{cap}}(\hat{\mathbf{z}}) - z_\kappa\,\mathrm{sd}(Q_{\mathrm{cap}} \mid \mathbf{y})$, which is not the root of equation (28.13). Which of the two is nearer the truth is a question for the full posterior.

Building blocks. The gradient (28.14), the block of $\boldsymbol{\Sigma}$, the second-order mean of equation (20.6), whose Hessian has the entries $\partial^2 Q_{\mathrm{cap}}/\partial g_1^2 = Q_{\mathrm{cap}}$ and $\partial^2 Q_{\mathrm{cap}}/\partial g_1 \partial T_s = -G_1$. For the values beside them, the reweighted draws.

The capacity of hall 1's path. At the MAP state $Q_{\mathrm{cap}}(\hat{\mathbf{z}}) = 12.8885 \times 5.49025 = 70.76\,\mathrm{kW}$, the plug-in crossing again. The gradient (28.14) is $(70.761\,\mathrm{kW},\ 0,\ 0,\ 0,\ -5.490\,\mathrm{kW/K},\ 0)^{\mathsf{T}}$, and the variance (28.15)

$$\begin{aligned}\mathbf{g}_{\mathrm{cap}}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}_{\mathrm{cap}} &= 70.761^2 \times 0.011382 + 5.490^2 \times 0.078433 - 2 \times 70.761 \times 5.490 \times 0.012974 \\ &= 56.994 + 2.364 - 10.080 = 49.278\,\mathrm{kW}^2,\end{aligned}$$

a spread of 7.02 kW and the delta band [57.00, 84.52] kW; the conductance's diagonal term is 116 % of the variance and the cross term −20 %. The second-order mean, equation (20.6) with the Hessian of Section 28.3.4, is $70.76 + \frac{1}{2}(0.011382 \times 70.761 - 2 \times 0.012974 \times 5.490) = 70.76 + 0.33 = 71.09\,\mathrm{kW}$. Table 28.4 gives the delta tails: each row is a design load, with its standardized distance $(Q_{\mathrm{cap}}(\hat{\mathbf{z}}) - Q_1^*)/7.020$ and the probability that the capacity reaches it, beside the probability of staying under the limit at the same load from Table 28.2, the other delta answer to the same event, equation (28.16).

Table 28.4: The capacity of hall 1's path by the delta method, $Q_{\mathrm{cap}} = 70.76 \pm 7.02\,\mathrm{kW}$: the probability that it reaches each design load, against the delta threshold probability of the same event written on the scenario temperature.

Design load Q_1^* [kW]	$(Q_{\mathrm{cap}}(\hat{\mathbf{z}}) - Q_1^*)/\mathrm{sd}$	$\Pr(Q_{\mathrm{cap}} \geq Q_1^*)$, delta on Q_{cap}	$\Pr(T_1^* < 25)$, delta on T_1^*
50	2.958	0.998	0.99999
60	1.533	0.937	0.966
65	0.821	0.794	0.814
70	0.108	0.543	0.544

At 60 kW the two delta answers to the one event are 0.937 through the capacity and 0.966 through the temperature. The delta assurance load from the capacity is $70.76 - 1.6449 \times 7.020 = 59.21\,\mathrm{kW}$.

The values beside them come from the full posterior, the reweighted draws pushed through equation (28.14): mean 71.93 kW, standard deviation 7.30 kW, median 71.37 kW, central 95 % band [59.28, 87.80] kW and skewness 0.47; $\Pr(Q_{\mathrm{cap}} \geq Q_1^* \mid \mathbf{y}) = 0.99996$, 0.966, 0.832 and 0.577 at 50, 60, 65 and 70 kW, the same numbers as through the temperature, since the event is the same. Against the reported delta on the capacity the gaps are 0.029 at 60 kW, 0.038 at 65 kW and 0.034 at 70 kW, and the band is shifted down by 2.3 kW at its lower end and 3.3 kW at its upper. At 60 kW the gap crosses the assurance level: the delta on the capacity puts the design load below $\kappa = 0.95$, at 0.937, with an assurance load of 59.21 kW, while the full posterior puts it above, at 0.966, as does the delta on the temperature. The state route errs here because the capacity is the room divided by the thermal resistance, a reciprocal of a nearly symmetric quantity, and so skewed to the right; a symmetric normal gives its lower tail too much mass.

28.3.5 The capacity factor of the chiller

Engineering question. Is the chiller delivering what its plate says? The maintenance engineer asks three things of the record: what the chiller delivers, how probable it is that it delivers less than 95 % of its plate, and whether the record contradicts the plate outright.

General form. Three forms of Chapter 2: the estimation form (2.10) for an entry; the threshold form (2.12) with $f = -c_{\mathrm{ch}}$, $f^* = -0.95$; and the hypothesis comparison (2.14), $B_{\mathrm{fp}} = p(\mathbf{y} \mid M_{\mathrm{f}})/p(\mathbf{y} \mid M_{\mathrm{p}})$, between the drawing M_{f} that frees the capacity factor and the drawing M_{p} that pins it at the plate, $c_{\mathrm{ch}} = 1$; its reciprocal is B_{pf}.

Specialized form. The three questions are equation (2.10) with $f(\mathbf{z}) = c_{\mathrm{ch}}$, the sixth entry; equation (2.12) with $f = -c_{\mathrm{ch}}$ and $f^* = -0.95$ (and -0.90); and equation (2.14) with M_{f} and M_{p}; $p(\mathbf{z} \mid \mathbf{y})$ is the posterior of step 3, whose sixth entry is $\mathcal{N}(\hat{c}_{\mathrm{ch}}, \Sigma_{66})$. The record and the plate enter through the choices of Section 28.1.4, and the capacity factor's posterior is the exact normal of equation (28.7): its state route and its exact value are one number.

The threshold is a normal tail of the entry, exact:

$$\Pr(c_{\mathrm{ch}} < c^\dagger \mid \mathbf{y}) = \Phi\left(\frac{c^\dagger - \hat{c}_{\mathrm{ch}}}{\sqrt{\Sigma_{66}}}\right), \qquad c^\dagger = 0.95 \text{ or } 0.90. \tag{28.17}$$

The comparison is the Savage–Dickey ratio of Chapter 24, equation (24.22). The drawing M_{p} is M_{f} with c_{ch} pinned at one and every other prior unchanged, so the condition (24.21) holds; the hall's entries share no row with the capacity factor and drop out of the ratio. The Bayes factor of the free capacity factor against the plate is then the prior density at the plate over the posterior density there,

$$\ln B_{\mathrm{fp}} = \ln \mathcal{N}(1;\ 1,\ \sigma_{\mathrm{pl}}^2) - \ln \mathcal{N}(1;\ \hat{c}_{\mathrm{ch}},\ \Sigma_{66}) = -\ln(\sigma_{\mathrm{pl}}\sqrt{2\pi}) + \ln\sqrt{2\pi\Sigma_{66}} + \frac{(1-\hat{c}_{\mathrm{ch}})^2}{2\Sigma_{66}}, \tag{28.18}$$

in nats, with $\mathcal{N}(x;\ \mu,\ v)$ the normal density of mean μ and variance v; the posterior is exactly normal, so the tail evaluation that Chapter 24 warns about costs nothing here. The ratio is also available in closed form, as the predictive density of $\bar{y}_P$ under M_{p}, $\mathcal{N}(P_{\mathrm{pl}},\ \sigma_P^2/N)$, over that under M_{f}, $\mathcal{N}(P_{\mathrm{pl}},\ V_{\mathrm{c}})$, which Section 28.4.3 sets against equation (28.18).

Building blocks. $\mathbf{\Lambda}_{\mathrm{c}}$, $\hat{c}_{\mathrm{ch}}$, Σ_{66} (Section 28.2.1); Φ; the two normal log-densities at $c_{\mathrm{ch}} = 1$.

The capacity factor from the duty record. The chiller delivers $P_{\mathrm{pl}}\hat{c}_{\mathrm{ch}} = 60 \times 0.911536 = 54.69\,\mathrm{kW}$, with a spread of $60 \times 0.010451 = 0.627\,\mathrm{kW}$, and its capacity factor has the central 95 % band $[0.891,\ 0.932]$, the delivered capacity $[53.46,\ 55.92]\,\mathrm{kW}$. By equation (28.17),

$$\Pr(c_{\mathrm{ch}} < 0.95 \mid \mathbf{y}) = \Phi\Big(\frac{0.95 - 0.911536}{0.010451}\Big) = \Phi(3.681) = 0.99988,$$
$$\Pr(c_{\mathrm{ch}} < 0.90 \mid \mathbf{y}) = \Phi\Big(\frac{0.90 - 0.911536}{0.010451}\Big) = \Phi(-1.104) = 0.135.$$

The plate lies $(1 - 0.911536)/0.010451 = 8.465$ spreads above the posterior mean. The Savage–Dickey ratio (24.22) reads the prior's log-density at the plate, $-\ln(0.08\sqrt{2\pi}) = 1.6068$, and the posterior's there, $-\ln(0.010451\sqrt{2\pi}) - 8.465^2/2 = 3.6423 - 35.8281 = -32.1858$, so

$$\ln B_{\mathrm{fp}} = 1.6068 + 32.1858 = 33.79 \text{ nats}.$$

Table 28.5 collects the answers to the fifth question: each row is a quantity, with its value and the formula that gives it.

Table 28.5: The chiller's capacity factor: its posterior from the ten-day duty record at full demand and the plate prior $\mathcal{N}(1,\ 0.08^2)$, the threshold probabilities and the evidence against the plate.

Quantity	Value	From
Capacity factor $\hat{c}_{\text{ch}}$, $\sqrt{\Sigma_{66}}$	0.9115, 0.0105	equations (28.6) and (28.7)
Central 95 % band of c_{ch}	$[0.891,\ 0.932]$	normal, exact
Delivered capacity	54.69 kW, band $[53.46,\ 55.92]$ kW	$P_{\text{pl}} c_{\text{ch}}$
$\Pr(c_{\text{ch}} < 0.95 \mid \mathbf{y})$	0.99988	equation (28.17)
$\Pr(c_{\text{ch}} < 0.90 \mid \mathbf{y})$	0.135	equation (28.17)
Plate in posterior spreads	8.465	$(1 - \hat{c}_{\text{ch}})/\sqrt{\Sigma_{66}}$
$\ln B_{\text{fp}}$, free against $c_{\text{ch}} = 1$	33.79 nats	equation (28.18)

Figure 28.2 draws the densities of the capacity factor. The horizontal axis is c_{ch}, the vertical axis its density; the gray curve is the plate prior, the solid blue curve the posterior from the record, and the dashed red curve the posterior of the drawing of Section 28.4.4, in which the heat meter carries a free bias. The vertical lines mark the plate, $c_{\text{ch}} = 1$, and the thresholds 0.95 and 0.90. The posterior is nearly eight times narrower than the prior, and at the plate the ratio of the blue to the gray curve is $B_{\text{pf}} = \text{e}^{-33.79}$.

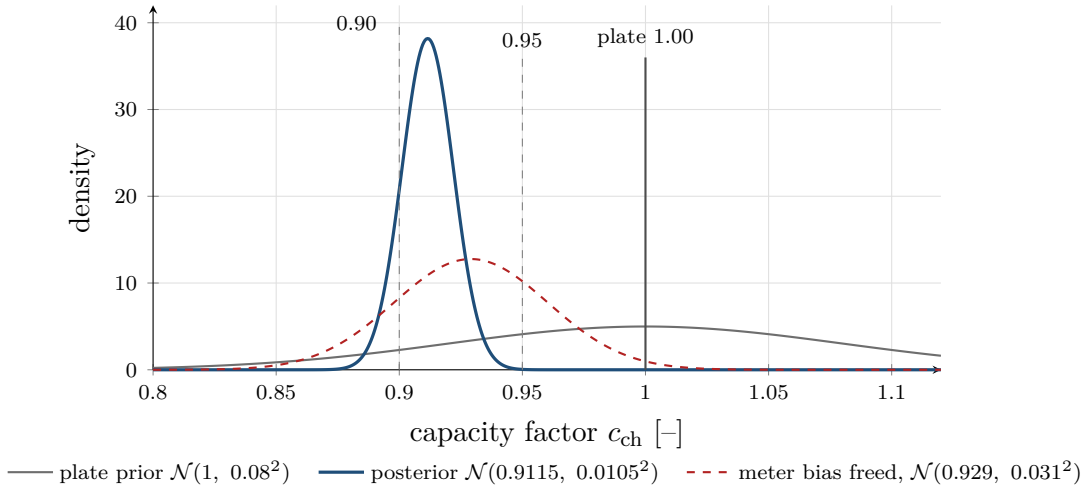


Figure 28.2: Densities of the chiller's capacity factor c_{ch}. Gray: the plate prior $\mathcal{N}(1,\ 0.08^2)$. Blue: the posterior from the ten-day duty record $\bar{y}_P = 54.6$ kW with spread 0.632 kW, $\mathcal{N}(0.9115,\ 0.01045^2)$, equation (28.7). Red dashed: the posterior when the heat meter carries a free bias with Chapter 22's prior spread of 2 kW, $\mathcal{N}(0.929,\ 0.0313^2)$ (Section 28.4.4). Vertical lines: the plate and the thresholds 0.95 and 0.90.

28.3.6 The chiller's headroom

Engineering question. How much load does the chiller have in hand now? The operator asks how many kilowatts can be added to the present loads before the chiller runs at full demand.

General form. The estimation form, equations (2.10) and (2.11), for a derived quantity.

Specialized form. The question is equations (2.10) and (2.11) with $f(\mathbf{z}) = P_{\text{hr}}(\mathbf{z})$ below and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$. One choice enters here: the

headroom is taken at the present loads, the chiller's delivered capacity less the present demand. It buys a quantity linear in the state; it says nothing of how the demand will change, which the scenario questions carry.

The headroom is

$$P_{\mathrm{hr}}(\mathbf{z}) = P_{\mathrm{pl}}c_{\mathrm{ch}} - Q_1 - Q_2, \qquad \mathbf{g}_{\mathrm{hr}} = \begin{pmatrix} 0, & 0, & -1, & -1, & 0, & P_{\mathrm{pl}} \end{pmatrix}^{\mathsf{T}},$$
$$\mathbf{g}_{\mathrm{hr}}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}_{\mathrm{hr}} = \Sigma_{33} + \Sigma_{44} + 2\Sigma_{34} + P_{\mathrm{pl}}^2\Sigma_{66}, \tag{28.19}$$

in kW, the last by the block-diagonal covariance of Section 28.2.1: the variance of the total load, Chapter 20's, plus that of the delivered capacity. It is linear, so the delta method is exact under the Laplace posterior, and Chapter 20 found the total load's delta band within 0.04 kW of the full posterior's, Section 20.4.4. The heat meter reads both terms: the present demand through the row on $Q_1 + Q_2$ and the capacity through the duty record. If the meter carries a bias b_P, each row reads its quantity plus b_P, and the headroom is the difference of the two quantities the meter reads, $(P_{\mathrm{pl}}c_{\mathrm{ch}} + b_P) - (Q_1 + Q_2 + b_P)$, in which the bias cancels; Section 28.4.4 checks it with the bias free.

Building blocks. $\hat{Q}_1 + \hat{Q}_2$, $\hat{c}_{\mathrm{ch}}$, the block of $\boldsymbol{\Sigma}$ in (Q_1, Q_2) and Σ_{66}.

The chiller's headroom now. By equation (28.19), $P_{\mathrm{hr}}(\hat{\mathbf{z}}) = 54.692 - (30.688 + 22.231) = 1.77\,\mathrm{kW}$, with the variance $3.7153 + 60^2 \times 1.0922 \times 10^{-4} = 3.7153 + 0.3932 = 4.1085\,\mathrm{kW}^2$, the first term the total load's of Chapter 20, Section 20.3.4: a spread of 2.03 kW and the band $[-2.20,\ 5.75]\,\mathrm{kW}$.

28.3.7 The derate carried to the hall

Engineering question. What does the derate do to the margin at full demand? The operator asks what happens to hall 1's margin when the chiller is asked for its plate and delivers less.

General form. The threshold form (2.12) for a derived quantity that now depends on the hall's entries and on the chiller's, and its inversion for the assurance load, equation (28.12).

Specialized form. The question is equation (2.12) with $f(\mathbf{z}) = -T_1^{*\prime}(\mathbf{z}; Q_1^*)$, the scenario temperature with the derate below, $f^* = -T_{\max}$, and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of the six entries, and its inversion for the assurance load, equation (28.12). Four choices enter here. The chiller is taken to remove at most its delivered capacity from water that returns as it would with the supply held, so that the supply leaves warmer by the shortfall over the loop's heat-capacity rate,

$$\Delta T_s(\mathbf{z}; P^*) = \frac{\left(P^* - P_{\mathrm{pl}}c_{\mathrm{ch}}\right)^+}{\dot{m}c_p}, \qquad (x)^+ = \max(0, x), \tag{28.20}$$

the onset shortfall of the loop balance (21.3): it buys a supply rise linear in the shortfall, and it fails for a sustained shortfall, since a loop that receives more heat than the chiller removes has no steady state and warms until the demand falls or the chiller's capacity rises with the temperature of the water it receives; the onset rise is the least that a sustained shortfall does (Section 28.6.5). The derate enters the hall through the supply alone, $T_1^{*\prime} = T_1^* + \Delta T_s$, the path and its conductance unchanged: this is the modeling choice that the hall rejects its heat to whatever supply the chiller delivers, the choice Chapter 20 made with $\Delta T_s = 0$; it fails if the loop's control answers a warm supply by raising the flow, which would change $\dot{m}c_p$ and the coil's conductance together. The heat-capacity rate is pinned at Chapter 21's 10 kW/K, equation (21.4); a flow that changed with the demand would make it an entry (Section 28.6.5).

Hall 2's load is taken off this chiller, onto another source, so that with hall 1 at Q_1^* the demand on the chiller is $P^* = Q_1^*$, the header balance with $Q_2^* = 0$: this buys a demand at the scenario equal to hall 1's load, the plate demand of 60 kW at $Q_1^* = 60\,\mathrm{kW}$; it breaks if hall 2 stays on the chiller, a free entry at its present 22.23 kW, and it breaks equally if hall 2 is only moved off the shared supply header while its water still returns to this chiller, the drawing B of Chapter 24, Section 24.1.2, in which the chiller still removes hall 2's heat. The case of hall 2 kept on the chiller is derived beside it.

Hall 2 kept on the chiller. The demand is then $P^* = Q_1^* + Q_2$, with the spread of Q_2, and at $Q_1^* = 60\,\mathrm{kW}$ it exceeds the delivered capacity by a sizable fraction of it. A loop whose chiller removes less heat than it receives has no steady state, since the water stores the difference, $\dot{m}c_p\,\mathrm{d}T_s/\mathrm{d}t > 0$ in whole-loop terms, and every row of the model is a steady balance; the onset rise (28.20) then describes only the first pass. The loop balance (21.4) gives a second reading of the same fact: with the supply held, the return would be $T_r = T_s + P^*/\dot{m}c_p$, to be set against the chiller's return limit of 18°C that Chapter 21 used. In this case the quantities computed are the shortfall beyond the plate and beyond the delivered capacity, the return temperature, and the hall-1 load at which the shortfall begins, $P_{\mathrm{pl}}c_{\mathrm{ch}} - Q_2$, a linear function with the variance $P_{\mathrm{pl}}^2\Sigma_{66} + \Sigma_{44}$.

The function is a hinge in the capacity factor. Under the posterior the shortfall $X = P^* - P_{\mathrm{pl}}c_{\mathrm{ch}}$ is normal, with mean $\mu_X = P^* - P_{\mathrm{pl}}\hat{c}_{\mathrm{ch}}$ and spread $\mathrm{sd}_X = P_{\mathrm{pl}}\sqrt{\Sigma_{66}}$, exactly, because c_{ch} is. Its positive part has the partial expectation

$$\mathbb{E}[X^+ \mid \mathbf{y}] = \int_0^\infty \frac{x}{\mathrm{sd}_X}\,\phi\Big(\frac{x-\mu_X}{\mathrm{sd}_X}\Big)\mathrm{d}x = \int_{-\mu_X/\mathrm{sd}_X}^\infty (\mu_X + \mathrm{sd}_X\,u)\,\phi(u)\,\mathrm{d}u,$$

with ϕ the standard normal density and $u = (x-\mu_X)/\mathrm{sd}_X$; since $\int_b^\infty \phi(u)\,\mathrm{d}u = \Phi(-b)$ and $\int_b^\infty u\,\phi(u)\,\mathrm{d}u = \phi(b)$, and ϕ is even, with $d_X = \mu_X/\mathrm{sd}_X$,

$$\mathbb{E}[X^+ \mid \mathbf{y}] = \mu_X\,\Phi(d_X) + \mathrm{sd}_X\,\phi(d_X), \qquad \Pr(X > 0 \mid \mathbf{y}) = \Phi(d_X). \tag{28.21}$$

The plug-in shortfall μ_X^+ is smaller than the expected one: the hinge is convex, and Jensen's inequality of Chapter 9 applies; the two agree when d_X is large, where the hinge is linear over the whole band. Where the shortfall is certain, $d_X \gg 1$, the hinge is the line $P^* - P_{\mathrm{pl}}c_{\mathrm{ch}}$ over the posterior and the hall's temperature with the derate is affine in the capacity factor, with the gradient of the pattern (28.4),

$$\mathbf{g}_\Delta = \Big(-Q_1^*\,\mathrm{e}^{-\hat{g}_1},\ 0,\ 0,\ 0,\ 1,\ -\frac{P_{\mathrm{pl}}}{\dot{m}c_p}\Big)^{\mathsf{T}}, \qquad \mathbf{g}_\Delta^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}_\Delta = \mathbf{g}_T^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}_T + \Big(\frac{P_{\mathrm{pl}}}{\dot{m}c_p}\Big)^2\Sigma_{66}, \tag{28.22}$$

the cross terms with the sixth entry vanishing because $\boldsymbol{\Sigma}$ is block diagonal (Section 28.2.1): the derate adds its own variance to the hall's and shifts the center by $\mathbb{E}[\Delta T_s \mid \mathbf{y}]$. The delta tail is equation (20.18) with the shifted center and this variance.

The assurance load with the derate, hall 2 off this chiller. With $P^* = Q_1^*$ the probability of staying under the limit becomes

$$p_{\mathrm{s}}'(Q_1^*) = \Pr\Big(T_s + \frac{(Q_1^* - P_{\mathrm{pl}}c_{\mathrm{ch}})^+}{\dot{m}c_p} + Q_1^*\,\mathrm{e}^{-g_1} < T_{\max}\ \Big|\ \mathbf{y}\Big),$$

nonincreasing in Q_1^* for the reason of Section 28.3.3, since the hinge is nondecreasing in the load; its root Q_κ' of $p_{\mathrm{s}}'(Q_\kappa') = \kappa$ is the assurance load of the derated plant with hall 2 off this chiller. It is found by bisection, with the delta tail of the gradient (28.22) where the shortfall is certain and with the reweighted draws for the full posterior. With hall 2 on the chiller the shortfall begins once hall 1 passes the onset load $P_{\mathrm{pl}}c_{\mathrm{ch}} - Q_2$, and the onset model does not carry the question to the loads of interest.

Building blocks. $\hat{c}_{\mathrm{ch}}$, Σ_{66}; the partial expectation (28.21); the gradient (28.22). For the values beside them, the reweighted hall draws joined with exact draws of c_{ch}.

The derate at the plant's demands. Table 28.6 gives the shortfall at five demands from equation (28.21), with $\mu_X = P^* - 54.692\,\mathrm{kW}$ and $\mathrm{sd}_X = 0.627\,\mathrm{kW}$. Each row is a demand; the columns are the plug-in shortfall, the probability of a shortfall, the expected shortfall and its spread, and the supply rise (28.20) with its spread. One row of the arithmetic, at 55 kW: $\mu_X = 0.308\,\mathrm{kW}$, $d_X = 0.491$, $\Pr(X > 0) = \Phi(0.491) = 0.688$, and $\mathbb{E}[X^+] = 0.308 \times 0.688 + 0.627 \times \phi(0.491) = 0.212 + 0.222 = 0.434\,\mathrm{kW}$.

Table 28.6: The chiller's shortfall and the onset rise of the supply at five demands P^*, from the posterior of the capacity factor, equations (28.20) and (28.21), with $\dot{m}c_p = 10\,\mathrm{kW/K}$.

P^* [kW]	plug-in μ^+ [kW]	Pr(shortfall)	$\mathbb{E}[X^+]$ [kW]	sd(X^+) [kW]	ΔT_s [K]
50	0	0.0000	0	0	0
55	0.31	0.688	0.43	0.46	0.043 ± 0.046
60	5.31	1.0000	5.31	0.63	0.531 ± 0.063
65	10.31	1.0000	10.31	0.63	1.031 ± 0.063
70	15.31	1.0000	15.31	0.63	1.531 ± 0.063

At the plate demand of 60 kW the derated chiller is 5.31 kW short and the supply rises by $5.31\,\mathrm{kW}/(10\,\mathrm{kW/K}) = 0.53\,\mathrm{K}$; at 55 kW the expected shortfall is 0.43 kW against the plug-in 0.31 kW.

Hall 1 at 60 kW, hall 2 off this chiller. With $Q_1^* = P^* = 60\,\mathrm{kW}$, hall 2's load off this chiller, the gradient (28.22) is $(-10.928\,\mathrm{K},\ 0,\ 0,\ 0,\ 1,\ -6\,\mathrm{K})^{\mathsf{T}}$, and

$$T_1^{*\prime}(\hat{\mathbf{z}}) = 23.040 + 0.531 = 23.571^{\circ}\mathrm{C},$$
$$\mathbf{g}_\Delta^{\mathsf{T}} \boldsymbol{\Sigma} \mathbf{g}_\Delta = 1.1543 + 6^2 \times 1.0922 \times 10^{-4} = 1.1543 + 0.0039 = 1.1582\,\mathrm{K}^2,$$

a spread of 1.076 K. The margin falls from 1.960 to 1.429 K, and

$$\Pr(T_1^{*\prime} < 25^{\circ}\mathrm{C} \mid \mathbf{y}) \approx \Phi\Big(\frac{1.429}{1.076}\Big) = \Phi(1.328) = 0.908.$$

The derate's term, $0.0039\,\mathrm{K}^2$, is 0.34 % of the variance. The assurance load $Q'_{0.95}$ of Section 28.3.7, hall 2 off this chiller, is 58.91 kW by the delta method. The values beside them come from the full posterior, the reweighted hall draws joined with exact draws of the capacity factor: $\Pr(T_1^{*\prime} < 25^{\circ}\mathrm{C} \mid \mathbf{y}) = 0.914 \pm 0.0006$, 0.006 above the reported 0.908, a median margin of 1.52 K with the band $[-0.69,\ 3.51]\,\mathrm{K}$, and $Q'_{0.95} = 58.96 \pm 0.02\,\mathrm{kW}$, 0.05 kW above. Both routes put the derated assurance load below the design load of 60 kW.

Hall 1 at 60 kW, hall 2 kept on the chiller. Chapter 20's scenario keeps hall 2 at its present load, a free entry at 22.23 kW. The demand is then $P^* = 60 + 22.23 = 82.23 \pm 2.35\,\mathrm{kW}$: 22.2 kW above the plate and $27.5 \pm 2.4\,\mathrm{kW}$ above the delivered capacity, an onset rise of 2.75 K. With the supply held the return would be $12.11 + 82.23/10 = 20.33^{\circ}\mathrm{C}$, against the 18°C limit and the 17.40°C of the present loads. The shortfall begins at the hall-1 load $54.69 - 22.23 = 32.46 \pm 2.43\,\mathrm{kW}$ (37.77 kW at the plate); hall 1 now draws 30.69 kW.

Figure 28.3 draws the derate against the demand. In panel (a) the horizontal axis is the chiller's demand P^* and the vertical axis the onset rise of the supply: the blue curve is the expected rise under the posterior of the capacity factor, with its central 95 % band shaded, and the gray line the rise a chiller at its plate would give. In panel (b) the vertical axis is the center of hall 1's temperature at 60 kW: the green line with the supply held, the gray line with a chiller

at its plate, the blue curve with the derate; the red line is the limit. The dashed verticals mark the plate demand of 60 kW and the demand of 82.2 kW with hall 2 kept on the chiller. The blue curve leaves the green at 54.7 kW, the delivered capacity, and reaches the limit at 74.3 kW; the gray line reaches it at 79.6 kW.

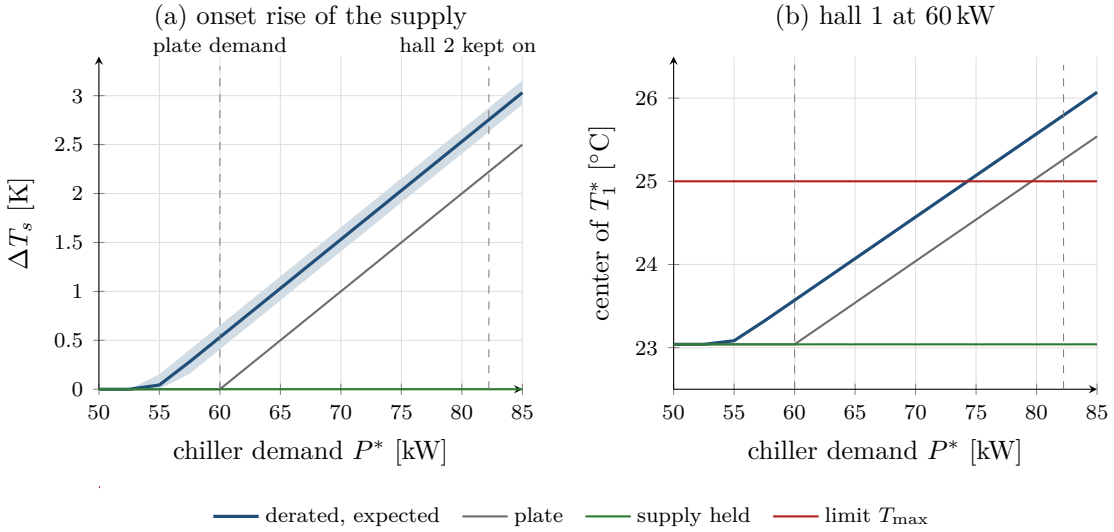


Figure 28.3: The derate against the chiller's demand. (a) The onset rise of the supply, equation (28.20): blue, its expectation under the posterior $\mathcal{N}(0.9115,\ 0.01045^2)$ of the capacity factor, equation (28.21), with the central 95 % band shaded; gray, a chiller at its plate. (b) The center of hall 1's temperature at 60 kW, 23.040°C + $\mathbb{E}[\Delta T_s]$: green, the supply held; gray, the plate; blue, the derate; red, the limit 25°C. Dashed: the plate demand 60 kW, hall 2 off this chiller, and the demand 82.2 kW with hall 2 kept on it. The onset model is the least a sustained shortfall does; beyond a few kilowatts of shortfall it is a lower bound on the rise.

28.4 The checks (step 5)

The delta answers rest on the Gaussian posterior of the hall's state and on the tangent plane of each output; the capacity factor's answers rest on the record's assumptions and on the plate prior. Each is set beside a calculation that does without it.

28.4.1 The margins against the full posterior

The hall's draws are Chapter 20's: 200 000 draws of the Laplace Gaussian from the PCG64 generator of NumPy, `default_rng(17)`, reweighted to the full posterior with $N_{\text{eff}} = 182\,933$ (Section 20.4.2). The capacity factor's draws, one per hall draw, come from its exact normal with `default_rng(28)`, and the standard errors of quantiles from 200 resamples of the weighted draws with `default_rng(5)`. Table 28.7 sets the delta method against the full posterior: each row is a load, with the delta probability of Table 28.2, the full posterior's with its standard error, the difference, whether the two agree within that standard error, and the full posterior's median and central 95 % band of the margin; the last row is hall 1 at 60 kW with the derate, hall 2 off this chiller.

The delta method agrees with the full posterior within the draws' standard error at 60 kW, as Chapter 20 found, and not at 50, 70 and 80 kW nor with the derate. At 70 kW, where the limit lies 0.11 standard deviations above the delta center, it understates the probability of staying under the limit by 0.034, thirty standard errors; at 80 kW, 1.16 standard deviations beyond the limit, by 0.011. The pattern is the one of Section 20.4.3: the full posterior's mean lies 0.03

Table 28.7: The probability that hall 1 stays under 25°C, by the delta method (reported) and by the full posterior (200 000 reweighted draws), whether they agree within the draws' standard error, and the full posterior's margin. The row at 60 kW is Chapter 20's, cited.

Case	delta	full (± s.e.)	delta − full	within s.e.	full median margin [K]	full 95 % band [K]
40 kW	1.0000	1.0000 ($< 10^{-5}$)	0.0000	yes	5.65	[4.21, 6.95]
50 kW	0.99999	0.99996 (0.00001)	0.00003	no	3.85	[2.03, 5.49]
60 kW	0.966	0.966 (0.0004)	0.0000	yes	2.05	[−0.16, 4.03]
70 kW	0.544	0.577 (0.0011)	−0.034	no	0.25	[−2.35, 2.58]
80 kW	0.124	0.135 (0.0009)	−0.011	no	−1.56	[−4.54, 1.13]
60 kW, derate	0.908	0.914 (0.0006)	−0.006	no	1.52	[−0.69, 3.51]

to 0.09 K, and its median 0.05 to 0.13 K, on the safe side of the delta center at every load, its margin has a skewness of −0.17, the temperature's +0.17 reflected, and the two effects cancel only near 1.8 standard deviations, where the question of Chapter 20 happened to lie. With the derate the limit lies 1.33 standard deviations above the center, short of the cancellation, and the delta method is low by 0.006.

Figure 28.4 draws the margin against the load. The horizontal axis is the load of hall 1 and the vertical axis the margin in kelvin; the blue line is the delta center $m_1^*(\hat{\mathbf{z}})$ with its central 95 % band shaded, the green dotted line the full posterior's median and the green dashed line its 5 % quantile; the red line at zero is the limit. The 5 % quantile of the margin crosses zero at the load where the probability of staying under the limit is 0.95, the assurance load, marked; the delta center crosses zero at the plug-in crossing, marked. The green curves lie above the blue at every load by a tenth of a kelvin or so, the shift of the center; the band's width grows with the load, as equation (28.10) says.

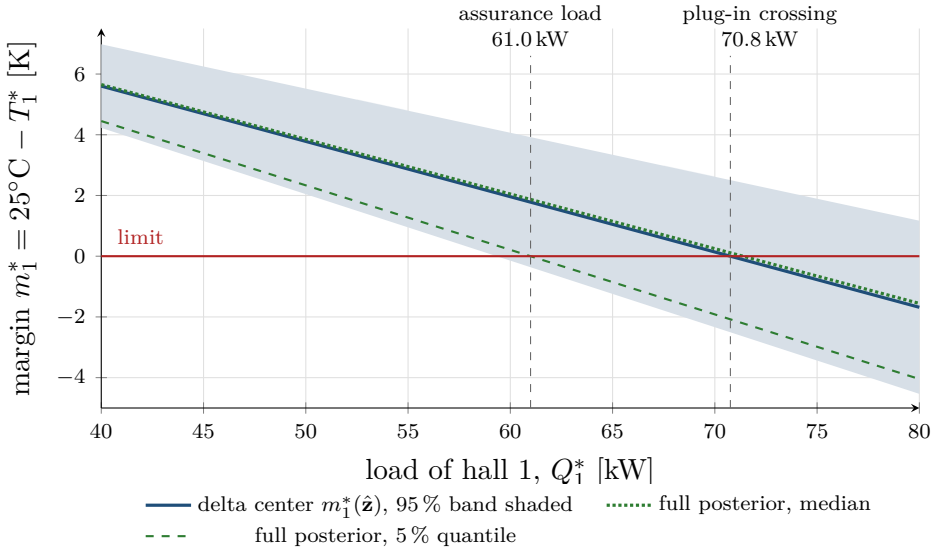


Figure 28.4: Hall 1's margin against its load with the supply held. Blue: the delta center $12.8885 - 0.18214\,Q_1^*$ and the delta band ± 1.960 sd, with sd^2 the quadratic (28.10). Green: the median (dotted) and the 5 % quantile (dashed) of the margin under the full posterior, 200 000 reweighted draws. Red: the limit. Dashed verticals: the assurance load $Q_{0.95} = 60.99$ kW of the full posterior, where its 5 % quantile crosses zero, and the plug-in crossing 70.76 kW.

The assurance load. The full posterior's 5 % quantile of the capacity, which by equation (28.16) is the assurance load, is $Q_{0.95} = 60.99\,\text{kW}$, with a standard error of 0.02 kW; the delta inversion's 60.90 kW lies 0.08 kW, four standard errors, below it, and the full posterior's probability at 60.90 kW is 0.9516 ± 0.0004, not 0.95. The reported value is the delta inversion's; the full posterior's 61.0 kW is set beside it. At the other levels the full posterior gives 57.32 kW at 0.99 (delta 57.61), 63.06 kW at 0.90 (delta 62.83) and 71.37 kW at 0.50, the capacity's median (delta 70.76), each with a standard error of 0.02 to 0.03 kW. The delta inversion is high at 0.99 and low at the others, for the same reason as the probabilities: it crosses the full posterior near 1.8 standard deviations. The derated assurance load is 58.96 ± 0.02 kW under the full posterior, against 58.91 kW by the delta method, the value reported.

28.4.2 The capacity against the full posterior

The same reweighted draws, pushed through equation (28.14), give the capacity a mean of 71.93 kW, a standard deviation of 7.30 kW, a median of 71.37 kW, the central 95 % band $[59.28,\ 87.80]$ kW and a skewness of 0.47. Draw by draw, the indicator of $Q_{\text{cap}} \geq 60\,\text{kW}$ equals the indicator of $T_1^*(60\,\text{kW}) \leq 25°\text{C}$, as equation (28.16) requires, so the full posterior's $\Pr(Q_{\text{cap}} \geq 60\,\text{kW} \mid \mathbf{y})$ is Chapter 20's 0.966, equation (20.19), the same number by the same draws; at 70 kW it is the 0.577 of Table 28.7, at 65 kW 0.832 and at 50 kW 0.99996. The delta method on the capacity departs from the full posterior beyond the draws' standard error at every design load: it is low by 0.029 at 60 kW, 0.038 at 65 kW and 0.034 at 70 kW, and its band $[57.00,\ 84.52]$ kW lies to the left of the full band at both ends, by 2.3 kW below and 3.3 kW above. Its assurance load of 59.21 kW is 1.8 kW low. The Laplace pushforward, the unweighted draws, gives the capacity a mean of 71.09 kW, which the second-order mean of Section 28.3.4 reproduces, and $\Pr(Q_{\text{cap}} \geq 60\,\text{kW}) = 0.952$, Chapter 20's Laplace pushforward of the same event.

Figure 28.5 draws the capacity's density. The horizontal axis is the capacity in kW and the vertical axis its density; the blue curve is the delta normal $\mathcal{N}(70.76,\ 7.02^2)$, the green curve the full posterior, a kernel estimate of the reweighted draws, and the shaded area under it to the left of 60 kW the probability of load–capacity interference at the design load. The verticals mark 60 and 70 kW. The green curve is skewed to the right, with its mode near 70 kW and a long upper tail, and below 60 kW it lies under the blue one: the area there is 0.034 under the full posterior and 0.063 under the delta normal.

Why the two delta answers differ. Chapter 20, Section 20.4.2, found the full posterior nearly symmetric in the thermal resistance e^{-g_1}, because the hall-1 reading row is linear in it, and the scenario temperature is linear in the resistance; the delta method on the temperature therefore lands close to the full posterior. The capacity is the room divided by the resistance, $(T_{\max} - T_s)/\text{e}^{-g_1}$: a reciprocal of a nearly symmetric quantity is skewed to the right, and the product with the room adds the curvature of the cross term. The delta method on the capacity puts a symmetric normal where the posterior has a skewness of 0.47, so its lower tail is too heavy, and it reports load–capacity interference at 60 kW as 0.063 where the full posterior gives 0.034. The question asked in kilowatts has the same answer as the question asked in degrees under any distribution of the state, 0.952 under the Laplace pushforward and 0.966 under the full posterior; only the delta method, which linearizes each differently, gives two.

28.4.3 The capacity factor's exact posterior

The capacity factor's rows are linear and its prior normal, so the Laplace posterior is the exact one and there is no other route to set against it; the checks are of the arithmetic. The closed-form evidence ratio of Section 28.3.5, the predictive densities of the record under the two

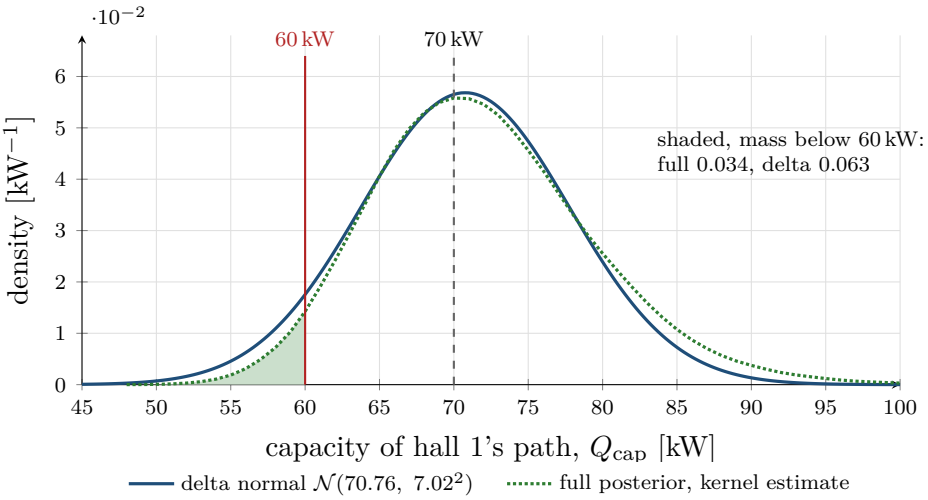


Figure 28.5: Density of the capacity of hall 1's path, $Q_{\text{cap}} = (25 - T_s)\,\mathrm{e}^{g_1}$. Blue: the delta normal, equation (28.15). Green dotted: the full posterior, a Gaussian kernel estimate (bandwidth 1 kW) of the 200 000 reweighted draws; shaded, its mass below the design load, the probability of load–capacity interference at 60 kW, 0.034, equal to Chapter 20's exceedance at 60 kW.

drawings, gives

$$\ln B_{\text{pf}} = \ln \mathcal{N}(54.6;\ 60,\ 0.4) - \ln \mathcal{N}(54.6;\ 60,\ 23.44) = -33.79,$$

the Savage–Dickey value (28.18) with the sign of B_{pf}, the reciprocal. Without the plate prior the record alone gives $c_{\text{ch}} = 54.6/60 = 0.910$ with the spread $0.632/60 = 0.0105$; the prior moves the center by 0.0015, a seventh of a spread.

28.4.4 A bias on the heat meter

The record reads the chiller through one meter, and the averaging over ten days removes its daily noise, not a calibration error. A meter that reads 5 kW low on every day produces the same record as a chiller that delivers 5 kW less, so the two are one column of the Jacobian, the footprint of Chapter 22. The check frees the bias as that chapter's bias drawings do, equation (22.4): a seventh entry b_P on the heat meter, with the prior row b_P/σ_b, $\sigma_b = 2$ kW, entering both rows the meter carries, the present reading y_P on $Q_1 + Q_2$ and the ten-day record on $P_{\text{pl}} c_{\text{ch}}$:

$$r_4 = \frac{Q_1 + Q_2 + b_P - y_P}{\sigma_P}, \qquad r_9 = \frac{P_{\text{pl}} c_{\text{ch}} + b_P - \bar{y}_P}{\sigma_P/\sqrt{N}}, \qquad r_{11} = \frac{b_P}{\sigma_b}.$$

The bias's column is shared by the hall's block and the chiller's, and the covariance is no longer block diagonal. The eleven-row fit gives $\hat{c}_{\text{ch}} = 0.929$ with $\sqrt{\Sigma_{66}} = 0.031$, three times the spread of the drawing without the bias, and $\text{corr}(c_{\text{ch}}, b_P) = -0.94$; the bias is $\hat{b}_P = -1.07 \pm 1.80$ kW, and the hall's entries barely move ($Q_2 = 23.17 \pm 2.86$ kW). With the meter free,

$$\Pr(c_{\text{ch}} < 0.95 \mid \mathbf{y}) = 0.748, \qquad \Pr(c_{\text{ch}} < 0.90 \mid \mathbf{y}) = 0.176, \qquad \ln B_{\text{fp}} = 1.63,$$

the last by the Savage–Dickey ratio with the same plate prior, whose condition (24.21) still holds because the bias's prior is independent of it. The red curve of Figure 28.2 is this posterior. The derate is no longer decisive: the evidence against the plate falls from 33.8 to 1.6 nats, and the probability that the chiller delivers less than 95 % of its plate from 0.9999 to 0.75. The present record cannot separate a chiller that delivers 54.7 kW from a meter that reads it low;

the fit divides the 5.4 kW between the two, $60 \times (1 - 0.929) = 4.3$ kW to the chiller and 1.1 kW to the meter, as the plate prior and the bias prior allow, and the record has no row that says which.

One quantity is untouched. The headroom of the plant, $P_{\mathrm{pl}} c_{\mathrm{ch}} - Q_1 - Q_2$, is 1.84 ± 2.03 kW with the bias free against 1.77 ± 2.03 kW without it: the meter reads both the capacity, on the days at full demand, and the present demand, and a bias common to both cancels in their difference. Whatever the meter's calibration, the chiller has about two kilowatts in hand at the present loads.

28.4.5 Sensitivity to the priors

Table 28.8 refits the plant under the priors of Chapter 20's Table 20.7 and under three plate priors. Each row is a prior; for the conductance and set-point priors the columns give the delta assurance load and the delta capacity with its spread, and for the plate priors the capacity factor, the probability that it is below 0.90 and the evidence against the plate.

Table 28.8: Sensitivity of the assurance load, the capacity and the capacity factor to the priors; the hall's quantities by the delta method, whose error against the full posterior, 0.08 kW, is smaller than their spread across priors.

Prior	$Q_{0.95}$ [kW]	$Q_{\mathrm{cap}}(\hat{\mathbf{z}})$ [kW]	sd [kW]
handbook, $g_i \sim \mathcal{N}(\ln 5,\ 0.3^2)$, $T_s \sim \mathcal{N}(12,\ 1)$	60.90	70.76	7.02
conductance prior widened, 0.6	61.19	71.60	7.47
conductance prior narrowed, 0.15	60.44	68.80	5.82
conductance prior centered on 4 kW/K	59.79	69.27	6.73
conductance prior centered on 6 kW/K	61.92	72.12	7.28
set-point prior widened, 2 K	60.92	70.80	7.04
Plate prior	$\hat{c}_{\mathrm{ch}}$	$\Pr(c_{\mathrm{ch}} < 0.90)$	$\ln B_{\mathrm{fp}}$
$\mathcal{N}(1,\ 0.08^2)$, the maker's tolerance	0.9115	0.135	33.79
$\mathcal{N}(1,\ 0.16^2)$, widened	0.9104	0.162	33.57
$\mathcal{N}(1,\ 0.04^2)$, narrowed	0.9158	0.060	32.72
$\mathcal{N}(0.95,\ 0.08^2)$, centered on 0.95	0.9107	0.153	34.29

The assurance load moves between 59.8 and 61.9 kW over the conductance priors, the center of the prior weighing more than its width, as for the threshold probability in Chapter 20; the design load of 60 kW meets the 95 % rule under every prior but the one centered on 4 kW/K, and by less than two kilowatts under any. The plate prior moves the capacity factor by less than half a spread and the evidence by a nat and a half: the record, with 98 % of the information, decides. The probability below 0.90 is the one number that moves, from 0.06 to 0.16, because 0.90 lies within a spread of the center.

28.4.6 The limit of exact balances

The margin, the capacity and the assurance load are functions of g_1 and T_s only, whose block of $\boldsymbol{\Sigma}$ Chapter 20 checked in the limit $\sigma_p \to 0$, Table 20.6: by the tolerance $\sigma_p = 0.01$ K the spread of the scenario temperature has stopped moving in its third decimal and the threshold probability moves by 10^{-4}, and at 0.001 K every printed number equals the reduced state's. The chiller adds no physics row: its two rows are a reading and a prior, and the loop balance enters only through the onset model (28.20) with $\dot{m} c_p$ pinned, a function asked about and not a row of the fit.

28.5 The report (step 6)

The plant's capacity is set by its chiller, not by hall 1's path. The path can take 60.9 kW with 95 % assurance (61.0 kW by the full posterior), ten kilowatts below the plug-in crossing of 70.8 kW; but the chiller delivers 54.7 kW, 91 % of its plate, and the plant has about two kilowatts in hand at its present loads. Chapter 20's move, hall 1 to 60 kW with hall 2 kept on the chiller at its present load, would ask 82 kW of a chiller that delivers 54.7 kW: a loop that receives more heat than its chiller removes has no steady state, so the held supply on which every margin of the path rests fails, by 27 kW, and would fail by 22 kW at the plate; the return header says the same, 20.3°C against the chiller's 18°C. Even with hall 2's load off this chiller, the derate takes half a kelvin from the margin at 60 kW and the probability of staying under the limit falls from 0.966 to 0.908 (0.914 by the full posterior). The derate itself is certain only if the heat meter is: with its calibration free, the evidence against the plate falls from 33.8 to 1.6 nats. And the capacity's delta band, asked in kilowatts, puts the probability of reaching 60 kW at 0.937, below the assurance of 0.95, for the event whose probability in degrees is 0.966, above it; the full posterior gives 0.966 in both. Each answer is reported from the posterior of the state, by the delta method, with the full posterior's value beside it and the gap, in the units in which the question was asked.

28.5.1 The answers and their bands

Table 28.9 is the report for the seven questions of Table 28.1, in their order: one row per question, with the answer from the posterior of the state in the second column, the full posterior's value beside it in the third, and in the fourth the band and what the checks of Section 28.4 found.

The report line. Hall 1's cooling path, with the supply held at its set point, can carry 61 kW with 95 % assurance (60.9 kW from the posterior of the state, 61.0 kW by the full posterior) and reaches its 25°C limit at 71 kW (delta center 70.8 kW, band 57 to 85 kW; full median 71.4 kW, band 59 to 88 kW); at 60 kW it stays under the limit with probability 0.97. The chiller does not hold the supply at those loads. Its ten-day record at full demand shows 54.7 kW delivered against a 60 kW plate, a derate the record establishes beyond doubt if the heat meter is calibrated and only weakly if it is not, and it leaves the plant about 2 kW of headroom at the present loads, whatever the meter's calibration. Hall 1 at 60 kW is possible only with hall 2's load off this chiller, and then with a probability of 0.91 of staying under the limit; with hall 2 kept on this chiller, the move needs more chiller capacity first.

28.6 The next measurement (step 7)

The report leaves two uncertainties that a measurement can narrow: the conductance of hall 1's path, which sets the band of every margin of the path, and the calibration of the heat meter, on which the derate's certainty rests. The candidates are rows added before load is moved, ranked by the preposterior of Chapter 20, equations (20.27) and (20.29), the form (2.17) of Chapter 2.

28.6.1 Candidates for the path's margin

The margin, the capacity and the assurance load are functions of g_1 and T_s, as the scenario temperature is, so Chapter 20's three candidates, Table 20.9, rank them the same way. A second hall-1 temperature sensor of 0.5 K, the one candidate whose line of **J** meets both columns, narrows the spread of the temperature at 60 kW from 1.074 to 0.874 K (that chapter's numbers) and the capacity's from 7.02 to 5.74 kW, a width removed of 18.2 %; with the center held, the delta assurance load rises from 60.90 to 62.51 kW, so the sensor buys about one and a half

Table 28.9: The answers to the questions asked of the two-hall plant and its chiller: answers from the posterior of the state, by the delta method (the capacity factor's posterior is exactly normal), with the full posterior by reweighted draws beside them. Bands are central 95 % posterior intervals; the hall's answers hold the supply, except the last.

Question	Answer (state route)	Full posterior	Band and check
Will the hall stay under the limit at 60 kW?	0.966 at 60 kW (Chapter 20); 1.0000, 0.99999, 0.544, 0.124 at 40, 50, 70, 80 kW	0.966; 1.0000, 0.99996, 0.577, 0.135	the delta method low by 0.034 at 70 kW and 0.011 at 80 kW
How much margin is there, and how well is it known?	1.960 K at 60 kW, spread 1.074 K; 5.603, 0.139, −1.683 K at 40, 70, 80 kW	median 2.05 K, spread 1.071 K; 5.65, 0.25, −1.56 K	$[-0.15,\ 4.07]$ K at 60 kW (full $[-0.16,\ 4.03]$); spread 0.70 to 1.45 K from 40 to 80 kW; full median 0.05 to 0.13 K above the delta center
How much load can the hall take with 95 % assurance?	$Q_{0.95} = 60.9$ kW; plug-in crossing 70.8 kW	61.0 ± 0.02 kW	57.6 and 62.8 kW at $\kappa = 0.99$ and 0.90 (full 57.3 and 63.1); 59.8 to 61.9 kW over the conductance priors; $Q'_{0.95} = 58.9$ kW (59.0 full) with the derate, hall 2 off this chiller
What is the path's capacity, and is it at least the design load?	70.76 ± 7.02 kW; $\Pr(Q_{\text{cap}} \geq 60\,\text{kW}) = 0.937$, below 0.95; $\Pr(\geq 70\,\text{kW}) = 0.543$	median 71.4 kW, mean 71.9 kW, spread 7.3 kW; 0.966, above 0.95; 0.577	$[57.0,\ 84.5]$ kW (full $[59.3,\ 87.8]$); the gap at 60 kW crosses the assurance level, from the skew of the capacity; the same event as the first question
Is the chiller delivering what its plate says?	no: $c_{\text{ch}} = 0.912 \pm 0.010$, 54.7 kW delivered; $\Pr(c_{\text{ch}} < 0.95) = 0.9999$; $\ln B_{\text{fp}} = 33.8$	the same (exact normal)	$[53.5,\ 55.9]$ kW; $\Pr(c_{\text{ch}} < 0.90) = 0.135$, 0.06 to 0.16 over the plate priors; with the meter's bias free 0.929 ± 0.031, 0.75 and 1.6 nats
How much load does the chiller have in hand now?	$P_{\text{hr}} = 1.77 \pm 2.03$ kW	– (linear; exact under the Laplace posterior)	$[-2.20,\ 5.75]$ kW; 1.84 ± 2.03 kW with the meter's bias free
What does the derate do to the margin at full demand?	hall 2 off this chiller: supply up 0.53 ± 0.06 K; margin at 60 kW 1.429 K; $\Pr(T_1^{*\prime} < 25°\text{C}) = 0.908$	median 1.52 K; 0.914	delta low by 0.006; with hall 2 kept on the chiller the demand is 82.2 kW, no steady state with the supply held, return 20.3°C against 18°C

kilowatts of assured load. A supply sensor of 0.1 K gives 61.08 kW and a load meter on hall 2 61.11 kW, a fifth of a kilowatt each. The sensor does not change which limit binds: it narrows the path's margin, and the path is not the plant's constraint.

28.6.2 Candidates for the chiller

The derate's question is asked of the drawing with the meter's bias free, Section 28.4.4, since that is where its answer is weak. Three candidates.

- *Twenty more days of the same meter* at full demand. Its row reads $P_{\mathrm{pl}}c_{\mathrm{ch}} + b_P$ with spread $2/\sqrt{20} = 0.447$ kW, the same combination as the record: the spread of c_{ch} goes from 0.0313 to 0.0304, a width removed of 3 %. More of the same meter cannot see its own bias.
- *An independent duty record* at full demand: the loop's heat-capacity rate times the rise from supply to return, $\dot{m}c_p(T_r - T_s)$, read on each load-bank day from the supply sensor and the return-header sensor of 0.3 K that Chapters 22 and 24 already carry (it read 17.50°C at the present loads), with $\dot{m}c_p$ pinned. Its spread on one day is $\dot{m}c_p\sqrt{\sigma_{T_r}^2 + \sigma_{T_s}^2} = 10\sqrt{0.3^2 + 0.3^2} = 4.24$ kW, and over the ten days, if the temperature errors are independent from day to day, $4.24/\sqrt{10} = 1.34$ kW. Its gradient is $\mathbf{j}_a = (0,\ 0,\ 0,\ 0,\ 0,\ 60/1.34,\ 0)^{\mathsf{T}}$ on the seven entries $(g_1, g_2, Q_1, Q_2, T_s, c_{\mathrm{ch}}, b_P)$, with no entry on the heat meter's bias, and by equation (20.28) the spread of c_{ch} falls from 0.0313 to 0.0182, a width removed of 41.8 % (8.5 % for a single day's reading), and $\mathrm{corr}(c_{\mathrm{ch}}, b_P)$ from -0.94 to -0.85. If the center stayed at 0.929, the probability of a derate below 95 % would rise from 0.75 to 0.88 and the evidence against the plate from 1.6 to 6.1 nats. The sensors are in place; what the candidate needs is their log over the load-bank days.
- *A duty record at a second supply temperature*, the same meter at full demand with the set point raised by two kelvin. A chiller that has lost capacity delivers more with warmer water returning to it, while a chiller held back by a control setting, a demand limiter at 91 %, delivers the same; the difference of the two records separates the two derates, and the meter's bias cancels in it (Exercise 28.6).

28.6.3 The recommendation

The decision about hall 1's load does not wait on a measurement: with hall 2 kept on this chiller, the move adds 29 kW to a headroom that the meter's calibration does not change. Before chiller capacity is bought or a derate is claimed under warranty, the next measurement is the independent duty record, the supply and return sensors already on the loop logged over the load-bank days: it is the one candidate that sees the chiller without the heat meter, and it removes about two fifths of the derate's spread. The second-supply-temperature record follows if the derate is confirmed, to decide between a repair and a control setting. A second hall-1 sensor is worth its cost only once the chiller can carry more load, when it buys one and a half kilowatts of assured load on the path. The general treatment of the next measurement, by its value for a decision, is the subject of Chapter 31, in Volume III.

28.6.4 What the analysis showed

The first readings, 10.8 kW to spare in hall 1 at 60 kW and 7 kW in hand at the chiller, said the plant had room; the analysis overturns both. The hall's path can take about a kilowatt beyond 60 kW with the customer's assurance, not ten, because the band of the conductance inferred at half the load widens with every kilowatt added. The chiller, read through its duty record instead of its plate, has about two kilowatts in hand, and the move that Chapter 20 analyzed would have asked of it half again what it delivers: that chapter's margin was the path's margin, valid only while the supply is held, and it was right to say so among its limits. The derate

is real if the heat meter is, and the record cannot tell the two apart; the headroom, which is what the operator needs, is the same either way, because the meter reads the demand and the capacity alike. Asked in kilowatts by a symmetric band, the capacity question gives a different answer from the same question asked in degrees, 0.937 against 0.966 at 60 kW, while the full posterior, which evaluates the event itself, gives 0.966 in both. What the engineer carries back to the site is specific: plan loads against the chiller's delivered duty, not its plate; read the plant's capacity as the smaller of the path's and the source's; read the chiller through the loop's own temperatures before claiming a derate; and do not move hall 1 to 60 kW while hall 2 stays on the same chiller.

28.6.5 Limits of the model

The answers are conditional on the steady-state model, on the scenario node's two choices, on the onset model of the shortfall and on the record's assumptions.

The conductance at the scenario load. As in Chapter 20, Section 20.6.6, the path's conductance inferred at 31 kW is used up to 80 kW; a coil whose conductance depends on the flow would need its own row, and every margin of the path would widen.

The onset shortfall. The supply rise (28.20) is the rise on the first pass of the water. A loop that receives more heat than its chiller removes has no steady state: the return warms, the supply with it, and the rise grows until the demand falls, the halls' own protection acts, or the chiller's capacity rises with the warmer water it receives. The onset rise is therefore the least that a sustained shortfall does, and the probability 0.908 at the plate demand (0.914 by the full posterior) an upper bound on the probability of staying under the limit, unless the chiller's capacity is known to rise with its entering temperature; a row for that dependence, from a record at a second supply temperature, would replace the bound by a steady state.

Hall 2 off this chiller. The derated answers at 60 kW take hall 2's load onto another source. A hall 2 moved only off the shared supply header, its water still returning to this chiller as in Chapter 24's drawing B, leaves its heat on the chiller, and those answers do not apply to it.

The pinned heat-capacity rate. $\dot{m}c_p$ is Chapter 21's 10 kW/K, pinned. A pump that answers a warm supply with more flow would lower the rise per kilowatt of shortfall and change the coils' conductances with it; the rate would then be an entry with the flow meter as its row.

Full demand on every day. The record is read as ten readings of the delivered capacity. A day on which the demand fell below the capacity reads the demand, and a record that mixes such days understates the capacity: the derate would be smaller than reported. The chiller's log of its load ratio on each day, if it exists, is the row that tests this. Conversely, on a day on which the load bank held the demand above the capacity, the supply was not held at 12°C: the loop warmed by the deficit, and the chiller delivered its capacity at a warmer supply than the plate's. A chiller whose capacity rises with the temperature of its water then delivered more than it would at 12°C, so the record overstates the capacity at the plate's conditions and the derate there is larger than reported; the supply sensor's log over the test days gives the temperature at which each reading was taken.

The meter's calibration. A bias common to the ten days is a derate in disguise (Section 28.4.4). The bias drawing with Chapter 22's prior spread of 2 kW is one choice of spread; a meter known to 0.5 kW would restore most of the record's evidence, one known only to 5 kW would leave the plate almost untested.

One limit, one hall. The margins are of one limit on one hall. Equipment with several limits has a joint margin, the intersection of the events, whose probability is below the smallest of the single ones (Exercise 28.5); and a plant with both halls loaded has both halls' margins and the chiller's shortfall moving together.

The steady state. Every margin is of the settled temperature. While load is raised, or while a shortfall persists, the hall passes through transients that the steady balance does not see; a storage term, as Chapter 1 wrote it, is needed for them.

28.7 What to retain

Definition 28.1 (Margins, capacity and derating as functions of the posterior)**.** Let a limit $T_{\max}$ apply to a temperature $T^*(\mathbf{z};Q) = T_s + Q\,\mathrm{e}^{-g}$ at a load Q. The margin is $m^* = T_{\max} - T^*$; the capacity $Q_{\mathrm{cap}} = (T_{\max} - T_s)\,\mathrm{e}^{g}$ is the load at which the margin vanishes; and because $Q_{\mathrm{cap}} \geq Q \iff T^*(Q) \leq T_{\max}$ for every state, $p_{\mathrm{s}}(Q) = \Pr(T^*(Q) < T_{\max} \mid \mathbf{y}) = \Pr(Q_{\mathrm{cap}} \geq Q \mid \mathbf{y})$ is the survival function of the capacity, nonincreasing in Q, and the assurance load Q_κ, the root of $p_{\mathrm{s}}(Q_\kappa) = \kappa$, is the $(1-\kappa)$ quantile of the capacity. Under the delta method on the temperature $Q_\kappa = x_\kappa \mathrm{e}^{\hat{g}}$, with x_κ the smaller root of $(1 - z_\kappa^2\Sigma_{gg})x^2 - 2(a - z_\kappa^2\Sigma_{gs})x + a^2 - z_\kappa^2\Sigma_{ss} = 0$, $a = T_{\max} - \hat{T}_s$, $z_\kappa = \Phi^{-1}(\kappa)$. A unit's delivered capacity is $P_{\mathrm{pl}}c$ with c an entry of the state, the plate a prior row $(c-1)/\sigma_{\mathrm{pl}}$, and a duty record at full demand a reading row $(P_{\mathrm{pl}}c - \bar{y})/(\sigma/\sqrt{N})$; the posterior is normal and exact, the threshold a normal tail, and the evidence against the plate the Savage–Dickey ratio of the prior to the posterior density at $c = 1$. A shortfall enters the hall through the supply, $\Delta T_s = (P^* - P_{\mathrm{pl}}c)^+/\dot{m}c_p$, with expectation $\mu_X\Phi(\mu_X/\mathrm{sd}_X) + \mathrm{sd}_X\phi(\mu_X/\mathrm{sd}_X)$ under a normal c.

Three conclusions carry beyond the two halls. A margin, a capacity and an assurance load are three readings of one event, and a question asked in kilowatts must give the answer the question asked in degrees gives; the full posterior does, a delta band in the wrong coordinate does not, and the coordinate in which the posterior is nearly symmetric is the one in which the delta method can be trusted. A rating is not a capacity: a plate enters as a prior, a duty record as a row, and the plant's capacity is the smaller of the capacities of its paths and of its source, so a margin computed with the source held is the margin of the path alone. And a derate read through one instrument is as certain as the instrument's calibration, which no averaging improves; the quantity the operator needs may be immune to that calibration, as the headroom is here, and it is worth asking which quantity that is before buying the measurement that would settle the other.

Table 28.10 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 28.10: The terms of operating margins, capacity and derating, and where each is defined.

Term	Expression	Units	Meaning	Defined in
Margin	$m_1^* = T_{\max} - T_s - Q_1^* \mathrm{e}^{-g_1}$	K	room to the limit	28.3.2
Margin variance	$Q_1^{*2}\mathrm{e}^{-2\hat{g}_1}\Sigma_{11} - 2Q_1^*\mathrm{e}^{-\hat{g}_1}\Sigma_{15} + \Sigma_{55}$	K^2	quadratic in the load	28.3.1
Capacity	$Q_{\mathrm{cap}} = (T_{\max} - T_s)\,\mathrm{e}^{g_1}$	kW	load at the limit	28.3.4
One event	$Q_{\mathrm{cap}} \geq Q_1^* \iff T_1^* \leq T_{\max}$	–	same probability	28.3.4
Load–capacity interference	$\Pr(Q_{\mathrm{cap}} < Q_1^* \mid \mathbf{y})$	–	capacity risk	28.3.4
Assurance load	$Q_\kappa = F_{\mathrm{cap}}^{-1}(1-\kappa)$	kW	tail inverted	28.3.3
Delta inversion	$x_\kappa = (B - \sqrt{B^2 - AC})/A$	K	rise at Q_κ	28.3.3
Plug-in crossing	$(T_{\max} - \hat{T}_s)\,\mathrm{e}^{\hat{g}_1}$	kW	$\kappa = \frac{1}{2}$, delta	28.3.3
Capacity factor	c_{ch}, delivered $P_{\mathrm{pl}}c_{\mathrm{ch}}$	–, kW	derate	28.1.4
Duty record row	$(P_{\mathrm{pl}}c_{\mathrm{ch}} - \bar{y}_P)/(\sigma_P/\sqrt{N})$	–	reading at full demand	28.1.4
Rank-one posterior	$\hat{c} = 1 + \varphi_{\mathrm{c}}(\bar{y}_P - P_{\mathrm{pl}})$, $\Sigma_{66} = \sigma_{\mathrm{pl}}^2 - V_{\mathrm{c}}\varphi_{\mathrm{c}}^2$	–	exact	28.3.5
Evidence against the plate	$\ln p(c=1) - \ln p(c=1 \mid \mathbf{y})$	nat	Savage–Dickey	28.3.5
Onset shortfall	$\Delta T_s = (P^* - P_{\mathrm{pl}}c_{\mathrm{ch}})^+/\dot{m}c_p$	K	supply rise	28.3.7
Partial expectation	$\mu_X\Phi(d_X) + \mathrm{sd}_X\phi(d_X)$	kW	mean shortfall	28.3.7
Headroom	$P_{\mathrm{hr}} = P_{\mathrm{pl}}c_{\mathrm{ch}} - Q_1 - Q_2$	kW	load in hand	28.3.6

This chapter closes Part VII and Volume II. The volume began with the state of a physical plant, formulated, fitted and measured for its spread; it read that state for its outputs, for the readings it does not have and for the ones that lie, for the faults and the drawings the evidence prefers; and Part VII turned it to risk: the expected loss of the enumerated configurations, the importance of each unit to that loss, the acceptance of a unit on its test record, and the margins a plant keeps against its limits. Every answer so far has been a probability, a band or a ratio of evidences. Volume III, Parts VIII and IX, turns these into decisions: the interval at which to inspect, the unit to repair or replace, the measurement worth its price and the reliability worth its investment, beginning with inspection intervals in Chapter 29, in Volume III, and then the whole method on integrated case studies of pumps, cooling systems and data centers.

28.8 Exercises

Exercise 28.1. *Formulation.* A hall's posterior on its log-conductance and supply temperature is normal with $\hat{g} = \ln 5$, $\hat{T}_s = 12\,^\circ\mathrm{C}$, spreads $\sqrt{\Sigma_{11}} = 0.10$ and $\sqrt{\Sigma_{22}} = 0.30\,\mathrm{K}$ and correlation 0.5; its limit is $21\,^\circ\mathrm{C}$. (a) Write its margin at 40 kW as a function of the state, with its gradient, and give its delta band and the probability that it is positive. (b) Find the load it can take with 95 % assurance by the delta inversion, and the plug-in crossing. ((a) $m^* = 21 - T_s - 40\,\mathrm{e}^{-g}$, 1 K at $\hat{\mathbf{z}}$, gradient $(8\,\mathrm{K},\ -1)$; variance $0.64 + 0.09 - 0.24 = 0.49\,\mathrm{K}^2$, band $1 \pm 1.372 = [-0.37,\ 2.37]\,\mathrm{K}$, $\Phi(1/0.7) = 0.923$. (b) $a = 9$, $z_{0.95}^2 = 2.7055$: $A = 0.97294$, $B = 8.95942$, $C = 80.75650$, $B^2 - AC = 1.69955$, $x = (8.95942 - 1.30367)/0.97294 = 7.8686\,\mathrm{K}$ and $Q_{0.95} = 5 \times 7.8686 = 39.34\,\mathrm{kW}$; the plug-in crossing is $9 \times 5 = 45.0\,\mathrm{kW}$, so the uncertainty costs 5.7 kW.)

Exercise 28.2. *Derivation.* (a) For $Q_{\mathrm{cap}} = (T_{\max} - T_s)\,\mathrm{e}^{g}$, derive the gradient and the delta variance in terms of the variances and the covariance of the two entries, and show that the event $Q_{\mathrm{cap}} \geq Q$ is the event $T_s + Q\,\mathrm{e}^{-g} \leq T_{\max}$. (b) For the hall of Exercise 28.1, compute the capacity and its spread and the delta probability that it reaches 40 kW, and set it against the answer of part (a) of that exercise and against the Laplace pushforward of the event, 0.909. ((a) $\mathbf{g} = (Q_{\mathrm{cap}},\ -\mathrm{e}^{g})$, variance ${Q_{\mathrm{cap}}}^2\Sigma_{gg} + \mathrm{e}^{2g}\Sigma_{ss} - 2Q_{\mathrm{cap}}\mathrm{e}^{g}\Sigma_{gs}$; dividing $(T_{\max} - T_s)\mathrm{e}^{g} \geq Q$

by $\mathrm{e}^{g} > 0$ gives $T_{\max} - T_s \geq Q\mathrm{e}^{-g}$. (b) $Q_{\mathrm{cap}} = 45\,\mathrm{kW}$, gradient $(45\,\mathrm{kW},\ -5\,\mathrm{kW/K})$, variance $20.25 + 2.25 - 6.75 = 15.75\,\mathrm{kW}^2$, spread $3.97\,\mathrm{kW}$; $\Phi(5/3.969) = \Phi(1.260) = 0.896$, against 0.923 through the temperature: one event, two delta answers, with the Laplace pushforward 0.909 between them.)

Exercise 28.3. *Derivation.* A twenty-day duty record of the same chiller at full demand averages $55.2\,\mathrm{kW}$ on a meter of $\sigma_P = 2\,\mathrm{kW}$. (a) With the plate prior $\mathcal{N}(1,\ 0.08^2)$, write the state, the residual vector and the 2×1 Jacobian, and derive the posterior by the rank-one update. (b) Compute $\Pr(c_{\mathrm{ch}} < 0.95 \mid \mathbf{y})$, $\Pr(c_{\mathrm{ch}} < 0.90 \mid \mathbf{y})$ and the Savage–Dickey evidence against the plate. ((a) $\mathbf{z} = (c_{\mathrm{ch}})$, $\mathbf{r} = ((60c_{\mathrm{ch}} - 55.2)/0.4472,\ (c_{\mathrm{ch}} - 1)/0.08)^{\mathsf{T}}$, $\mathbf{J} = (134.16,\ 12.5)^{\mathsf{T}}$; $V_{\mathrm{c}} = 23.04 + 0.20 = 23.24\,\mathrm{kW}^2$, $\varphi_{\mathrm{c}} = 0.384/23.24 = 0.016523\,\mathrm{kW}^{-1}$, $\hat{c}_{\mathrm{ch}} = 1 - 0.016523 \times 4.8 = 0.92069$, $\Sigma_{11} = 0.0064 - 23.24 \times 0.016523^2 = 5.508 \times 10^{-5}$, spread 0.00742. (b) $\Phi(3.950) = 0.99996$; $\Phi(-2.788) = 0.0027$; the plate lies 10.687 spreads above the mean, $\ln B_{\mathrm{fp}} = 1.6068 - (3.9844 - 57.104) = 54.73$ nats.)

Exercise 28.4. *Derivation.* (a) Show that $p_{\mathrm{s}}(Q) = \Pr(T_s + Q\,\mathrm{e}^{-g} < T_{\max} \mid \mathbf{y})$ is nonincreasing in Q for any posterior, and that the assurance load is the $(1-\kappa)$ quantile of the capacity. (b) Show that the plug-in crossing is the load at which both delta answers equal $\frac{1}{2}$, and explain why the full posterior's median capacity, $71.37\,\mathrm{kW}$ for hall 1, lies above the plug-in crossing of $70.76\,\mathrm{kW}$. ((a) For $Q' > Q$ the event at Q' is contained in the event at Q, because $\mathrm{e}^{-g} > 0$; by equation (28.16) $p_{\mathrm{s}}(Q) = 1 - F_{\mathrm{cap}}(Q)$, so $p_{\mathrm{s}}(Q_\kappa) = \kappa$ is $F_{\mathrm{cap}}(Q_\kappa) = 1 - \kappa$. (b) At $Q = a\,\mathrm{e}^{\hat{g}}$ both centers are at the limit, $T^*(\hat{\mathbf{z}}) = T_{\max}$ and $Q_{\mathrm{cap}}(\hat{\mathbf{z}}) = Q$, and $\Phi(0) = \frac{1}{2}$. The full posterior's median capacity is $Q_{\mathrm{cap}}(\hat{\mathbf{z}})$ only if the posterior puts half its mass on each side of the level set of the capacity through $\hat{\mathbf{z}}$; the full posterior of g_1 is skewed toward large conductances, its mean 0.013 above the MAP value (Chapter 20), the capacity increases with g_1, and more than half the mass lies on the side of larger capacities, so the median lies above $Q_{\mathrm{cap}}(\hat{\mathbf{z}})$.)

Exercise 28.5. *Application.* Suppose hall 1's chiller had ample capacity but its evaporator admitted a rise from supply to return of at most $8.5\,\mathrm{K}$, a second limit, $(Q_1^* + Q_2)/\dot{m}c_p \leq 8.5\,\mathrm{K}$ with $\dot{m}c_p = 10\,\mathrm{kW/K}$. At $Q_1^* = 62\,\mathrm{kW}$, with hall 2 at its present load and the fit of Section 28.2.2, the two margins have delta centers $1.596\,\mathrm{K}$ and $0.077\,\mathrm{K}$, spreads 1.112 and $0.2346\,\mathrm{K}$ and covariance $0.0682\,\mathrm{K}^2$. (a) Derive the gradient of the second margin and the covariance from the entries $\Sigma_{14} = -0.062742$ and $\Sigma_{45} = -0.026883\,\mathrm{K\,kW}$. (b) Compute each probability and the joint probability that both margins are positive, and compare the joint with the product. ((a) The second margin is $8.5 - (Q_1^* + Q_2)/10$, gradient $-0.1\,\mathrm{kW}^{-1}$ on Q_2 and zero elsewhere; the first has $(62\,\mathrm{e}^{-\hat{g}_1},\ 0,\ 0,\ 0,\ -1) = (11.293,\ 0,\ 0,\ 0,\ -1)$; covariance $11.293 \times (-0.1) \times (-0.062742) + (-1) \times (-0.1) \times (-0.026883) = 0.0709 - 0.0027 = 0.0682$, correlation 0.261. (b) $\Phi(1.435) = 0.924$ and $\Phi(0.328) = 0.629$; the bivariate normal gives 0.596 for both, above the product 0.581 because the margins are positively correlated, and below the smaller of the two: the joint margin is set by the rise, not by the hall.)

Exercise 28.6. *Application.* A capacity derate and a control setting both make the chiller deliver $54.7\,\mathrm{kW}$ at a $12^{\circ}\mathrm{C}$ supply. Take, as an illustrative figure, a chiller whose capacity rises by $3\,\%$ per kelvin of supply temperature, so that a capacity derate delivers $60\,c_{\mathrm{ch}}(1 + 0.03\,(T_s - 12))$, while a demand limiter delivers $60\,c_{\mathrm{ch}}$ at any supply. A second ten-day record is taken at full demand with the supply set at $14^{\circ}\mathrm{C}$, on the same meter ($0.632\,\mathrm{kW}$). (a) With the posterior $c_{\mathrm{ch}} \sim \mathcal{N}(0.9115,\ 0.01045^2)$, give the predictive mean and spread of the record under each drawing. (b) Compute the expected log Bayes factor in favor of the true drawing, the Kullback–Leibler divergence of Chapter 24, and say why the meter's bias does not spoil the comparison. ((a) Capacity derate: $60 \times 0.9115 \times 1.06 = 57.97\,\mathrm{kW}$, spread $\sqrt{(63.6 \times 0.01045)^2 + 0.632^2} = 0.918\,\mathrm{kW}$; limiter: $54.69\,\mathrm{kW}$, spread $\sqrt{0.627^2 + 0.632^2} = 0.891\,\mathrm{kW}$; the predictions are $3.28\,\mathrm{kW}$, 3.7 spreads, apart. (b) For normal predictives, $\mathrm{KL} = \ln(s_2/s_1) + (s_1^2 + \Delta^2)/2s_2^2 - \frac{1}{2}$: 6.8 nats if the derate is of capacity and 6.4 if it is a setting, decisive either way. The bias b_P adds the same amount

to both records, so it moves both predictions together and cancels in their difference, which is what separates the drawings.)

Appendix A

Atlas of probabilistic questions

This appendix can be read independently of the chapters. Each row is a question a reliability engineer asks, its mathematical form once the plant graph and its state $\mathbf{z}$ are fixed, the object that answers it, and the chapter in which the derivation is made. The rows grow as chapters are added.

Engineering question	Mathematical form	Derives to	Chapter
What is this quantity, and how well is it known?	$p(z_i \mid \mathbf{y})$	the marginal of the posterior: $\hat{z}_i$ and Σ_{ii}	4
What is a quantity that is not an entry of the state?	$p(f(\mathbf{z}) \mid \mathbf{y})$	$f(\hat{\mathbf{z}})$ and $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ by the delta method	4
Will a temperature stay below its limit?	$\Pr(T < T_{\max} \mid \mathbf{y})$	$\Phi\big((T_{\max} - \mu_T)/\sigma_T\big)$	4
What is the failure probability per demand?	$p(p \mid k, n)$	$\mathrm{Beta}(a + k,\ b + n - k)$	5
Is the item at least p^* reliable, with stated assurance?	$\Pr(p \le 1 - p^* \mid k, n)$	the Beta cumulative distribution at $1 - p^*$	5
Will the next m demands succeed?	$\Pr(\text{no failure in } m \mid k, n)$	the Beta–binomial predictive	5
How many tests demonstrate a reliability requirement?	n such that $\Pr(p \le 1 - p^* \mid 0, n) \ge \gamma$	the smallest n satisfying the Beta tail	5
What is the failure rate behind a count over an exposure?	$p(\lambda \mid k, T_{\mathrm{tot}})$	$\mathrm{Gamma}(a + k,\ b + T_{\mathrm{tot}})$	6
Is the rate below a limit, with stated assurance?	$\Pr(\lambda \le \lambda_{\max} \mid k, T_{\mathrm{tot}})$	the gamma cumulative distribution at $\lambda_{\max}$	6
How many failures in the next τ hours?	$\Pr(K_f = j \mid k, T_{\mathrm{tot}})$	the negative-binomial predictive	6

Engineering question	Mathematical form	Derives to	Chapter
Will the item survive a mission of length t?	$\Pr(T > t \mid \mathbf{y})$	the posterior expectation of $R(t)$: $\left(b/(b+t)\right)^{a}$ under a gamma posterior	6
What lifetime distribution does a record with survivors support?	$p(\ln\eta, \ln\beta \mid \mathbf{y})$	the censored likelihood, its MAP and Hessian	7
How long will a unit that has reached age a last?	$\Pr(T > a+\tau \mid T > a, \mathbf{y})$	the ratio of survivals averaged over the posterior	7
Does the record show wear-out?	$p(\mathbf{y} \mid \beta > 1)/p(\mathbf{y} \mid \beta = 1)$	an evidence ratio on the shape	7
What is the repair rate, and the mean time to repair?	$p(\mu \mid r, D)$	$\text{Gamma}(a+r,\ b+D)$	8
What fraction of the time is the item available, and how well is that known?	$p(A \mid \mathbf{y})$, $A = \mu/(\lambda+\mu)$	a Beta on the exact route; the delta method on the logit	8
Is the availability at least A_{req}?	$\Pr(A \geq A_{\text{req}} \mid \mathbf{y})$	a Beta tail, or a normal tail on the logit	8
Which record tightens the availability more?	$\boldsymbol{\Sigma}_a$ for each candidate record	the preposterior on the logit	8
Will the unit run through a mission of length t?	$\Pr(T > t \mid \mathbf{y}) = \mathbb{E}[R(t;\lambda) \mid \mathbf{y}]$	the Lomax predictive $\left(b/(b+t)\right)^{a}$; the plug-in understates it	9
How well is the mission reliability itself known?	$p\big(R(t;\lambda) \mid \mathbf{y}\big)$	the quantile map of the gamma posterior	9
Will the unit start and run?	$\mathbb{E}[(1-p)\,\mathrm{e}^{-\lambda t} \mid \mathbf{y}]$	a product of two entries; the budget on $\ln R$	9
How many zero-failure hours demonstrate a mission reliability?	T such that $\Pr(R(t) \geq R^{*} \mid 0, T) \geq \gamma$	the gamma tail at the required rate	9
Will a unit of age t_0 survive the next period?	$\Pr(T > t_0+\tau \mid T > t_0, \mathbf{y})$	the ratio of two predictives; the record's own rows	10
How much life is left, on average?	$\mathbb{E}[T - t_0 \mid T > t_0, \mathbf{y}]$	the integral of the conditional survival (restricted horizon)	10
Will any survivor of a fleet fail before the next test?	$\Pr(\text{no failure in } \tau \mid \mathbf{y})$	one expectation over the shared state, not a product	10
How many failures will a fleet have next year?	$p(K_f \mid \mathbf{y})$	the negative binomial; the variance split into Poisson and epistemic parts	11

Engineering question	Mathematical form	Derives to	Chapter
How many spares cover the year with stated probability?	smallest s with $\Pr(K_f \le s \mid \mathbf{y}) \ge q$	a quantile of the predictive	11
Was last year's count consistent with the record before it?	$\Pr(K \ge k \mid \mathbf{y}_{\text{prior years}})$	the leave-one-out predictive tail; an evidence ratio for an alternative rate	11
Is the rate rising?	$\Pr(\gamma > 0 \mid \mathbf{y})$; $p(\mathbf{y} \mid \gamma \text{ free})/p(\mathbf{y} \mid \gamma = 0)$	a normal tail on the trend entry; Savage–Dickey	12
When does the rate cross the limit?	$p(t^* \mid \mathbf{y})$, $t^* = \bar{t} + (\ln \lambda_{\max} - \alpha)/\gamma$	draws or the inversion, not $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ (a pole)	12
What is the service availability of several units under a drawing?	$\mathbb{E}[R_{\text{sys}}(\mathbf{q}) \mid \mathbf{y}]$	the 2^n enumeration; plug-in exact under independent posteriors, a gap under a shared one	13
Which unit matters most to the service?	$\partial R_{\text{sys}}/\partial q_i$	the Birnbaum gradient, the $\mathbf{J}$ row of the band	13
Which combinations take the service down, and how likely is each?	the minimal cut sets; $\mathbb{E}[\prod_{i \in c} q_i \mid \mathbf{y}]$	a filter on the enumeration; ranked with bands	14
How likely is the top event, and how good is the rare-event sum?	$\mathbb{E}[\Pr(\text{top}) \mid \mathbf{y}]$	exact by enumeration; inclusion–exclusion orders	14
What share of the failures is a common cause?	$p(\beta_{\text{cc}} \mid k_1, k_2)$	the grid posterior on (rate, share)	15
Did one cause take two units out?	$p(\mathbf{y} \mid \text{shared cause})/p(\mathbf{y} \mid \text{independence})$; $\Pr(\text{common cause} \mid \text{the double})$	an evidence ratio with a coincidence window; an event posterior	15
What is a redundant pair's unavailability with a common cause?	$\mathbb{E}[U_{\text{pair}} \mid \mathbf{y}]$	the two-state model with the shared process	15
Which of several explanations is the plant?	$\Pr(M_j \mid \mathbf{y})$	the evidences $p(\mathbf{y} \mid M_j)$ and the prior model probabilities	2
What does a measurement reduce?	$\mathbb{E}[\text{posterior spread} \mid a]$	the preposterior $\boldsymbol{\Sigma}_a = (\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Lambda}_a)^{-1}$	2
What action minimizes expected loss?	$\arg\min_a \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$	the expected loss of each action under the posterior	2

Engineering question	Mathematical form	Derives to	Chapter
What fraction of the time is a repairable system out, with shared repair and a common cause?	$\mathbb{E}[\pi_{\text{down}}(\mathbf{z}) \mid \mathbf{y}]$	the stationary distribution of a Markov chain, $\pi Q = 0$; its gradient by implicit differentiation	16
How soon is the system whole again after an outage?	$p(t) = p(0)\,\mathrm{e}^{Qt}$	the matrix exponential; time constants from the eigenvalues	16
How long until the system's first outage?	$\mathbb{E}[\text{first passage} \mid \mathbf{y}]$	a linear solve on the transient states	16
What can these readings say about the plant's state?	$r(\mathbf{z})$, $\mathbf{J}$: the rows and their design columns	the formulation: physical, reading and prior rows over one state vector	17
What are the plant's conductances, loads and supply temperature?	$\hat{\mathbf{z}} = \arg\max p(\mathbf{z} \mid \mathbf{y})$	the objective $\frac{1}{2}\sum r_k^2$ minimized by Gauss–Newton	18
Do the readings agree with the physics and with each other?	$2\mathcal{L}(\hat{\mathbf{z}})$ against χ^2 with rows minus entries degrees of freedom	the normalized residuals and the redundancy check	18
Is the estimate physically admissible?	$\Pr(G_i > 0,\ Q_i > 0,\ T_i > T_s \mid \mathbf{y})$, a joint event	the joint tail under the Laplace normal; distances from the boundary in standard deviations	18
How well is each entry known, and which are correlated?	$\boldsymbol{\Sigma} = \boldsymbol{\Lambda}^{-1}$	the Laplace approximation for n entries; marginals and correlations	19
Can my sensors see this parameter at all?	rank of the design columns; prior spread against posterior spread	observability analysis: the singular values of $\mathbf{J}$ and the weak directions	19
With instruments out, how likely is a load above a limit?	$\Pr(Q_2 > 45\,\text{kW} \mid \mathbf{y}')$ with the meter rows removed	the exact tail as a one-dimensional integral over the conductance; the prior's spread swept	19
What will the hall's temperature be at a future load, and how well is that known?	$p(f(\mathbf{z}) \mid \mathbf{y})$ at a scenario node: location and central interval	the delta method $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ against the full posterior by reweighted draws	20
Will the hall stay under the limit at a future load?	$\Pr(f(\mathbf{z}) < T_{\max} \mid \mathbf{y})$ at a scenario node	the delta method $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ and a normal tail; a check against reweighted draws	20

Engineering question	Mathematical form	Derives to	Chapter
Which entry's uncertainty drives an output?	the budget of $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$	one term per entry plus the cross terms, which may be negative	20
What are a ratio, a total and a conductance, with bands?	$p(f(\mathbf{z}) \mid \mathbf{y})$ for nonlinear f	the delta band against the full posterior; the delta band shifted where f is curved	20
What is the temperature where there is no sensor?	$p(f(\mathbf{z}) \mid \mathbf{y})$ for a node quantity with no reading row	a virtual sensor: the same $\boldsymbol{\Sigma}$ pushed through f; the downdate of a removed row; the predictive of an unseen reading	21
Is the return temperature below the chiller's limit?	$\Pr(T_r(\mathbf{z}) < 18°\mathrm{C} \mid \mathbf{y})$	a derived quantity's tail by the delta method and by reweighted draws	21
Do the readings disagree with the model?	$\Pr(\chi^2_\nu > 2\mathcal{L}(\hat{\mathbf{z}}_0))$, a tail of the predictive	bad-data detection: the misfit against its chi-square law	22
Is this reading lying? Which one?	the normalized residuals $r_k/\sqrt{1-K_{kk}}$; the largest against r^*	bad-data identification: the hat matrix, the redundancy number, the Bonferroni threshold	22
Which sensor is biased, by how much, and how sure is that?	$\Pr(M_k \mid \mathbf{y}) \propto p(\mathbf{y} \mid M_k)\,\Pr(M_k)$ over a family of bias entries; $\hat{b}_k$ and its band	diagnosis by evidence: one freed entry per sensor, the posterior over the family	22
Can a bias on this sensor be detected at all?	$\Pr(\lvert r^{\mathrm{N}}_k\rvert > r^* \mid$ bias $\Delta)$; $1-K_{kk}$	the minimum detectable bias b^*_k from the predictive of the other rows; the critical (unredundant) measurement	22
What does a biased reading do to the answers if it is kept?	$p(f(\mathbf{z}) \mid \mathbf{y})$ under the drawings with and without the bias	the scenario output refitted under each drawing	22
Which row makes a sensor checkable?	$1-K_{kk}$ and b^*_k after a candidate row, from $\boldsymbol{\Sigma}_a$	the redundancy number by the preposterior; a candidate ranked by the bias it would detect	22
Has the plant changed since the last period?	a tail of the predictive under the healthy model	fault detection: the earlier period's marginals as prior rows, the misfit and the held-out departures	23
Which component?	$\Pr(M_j \mid \mathbf{y}) \propto p(\mathbf{y} \mid M_j)\,\Pr(M_j)$ over the candidates the graph generates	fault isolation: a widened prior or an added entry per candidate; the posterior over the family, conditional on the family	23

Engineering question	Mathematical form	Derives to	Chapter
How large is the fault, and how well is it known?	$p(z_i \mid \mathbf{y}, M_j)$ within the leading drawings	fault identification: the freed entry's posterior and its band	23
Which parameter moved between two periods?	$(\hat{\mathbf{z}}_2 - \hat{\mathbf{z}}_1)_i$ against $\sqrt{(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)_{ii}}$; the joint statistic	drift attribution: two fits compared entry by entry	23
How long a record would have told?	the answers as functions of the record length N	the averaged rows $\bar{y}_k$ with spread $\sigma_k/\sqrt{N}$; the limits as N grows	23
Is it really parallel? Is this hall really on the loop?	$p(\mathbf{y} \mid M_\mathrm{A})/p(\mathbf{y} \mid M_\mathrm{B})$	model selection: the evidence as a normalizing constant, the Occam factor, the Laplace evidence and what cancels	24
Does the record contradict the handbook value?	$B_{01} = p(z_j^\mathrm{h} \mid \mathbf{y}, M_1)/p(z_j^\mathrm{h} \mid M_1)$	a point null by the Savage–Dickey ratio; the closed-form evidence of a count	24
What availability is reported when the drawing is uncertain?	$\sum_d \Pr(M_d \mid \mathbf{y})\, \mathbb{E}[U_\mathrm{c} \mid M_d, \mathbf{y}]$	the mixture over drawings and its between-drawing variance	24
Which reading settles the structure, before it is installed?	$\sum_d \Pr(M_d \mid \mathbf{y})\, \mathbb{E}[\ln B_{dd'}(\tilde{y}_a) \mid M_d]$	the expected evidence gain: the Kullback–Leibler divergence of the predictives	24
What is the expected loss per year, and how well is it known?	$\mathcal{R}(a) = \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$, $\mathcal{R}_\mathrm{yr} = t_\mathrm{yr} \sum_\mathbf{s} \Pr(\mathbf{s} \mid \mathbf{q})\, C(\mathbf{s})$	probabilistic risk assessment: the enumeration weighted by consequence, and its band	25
Where does the risk come from?	$\mathbb{E}[f_k(\mathbf{z}) \mid \mathbf{y}]$ for the parts of the risk	the sum split by state class and by minimal cut set	25
Is the annual loss below a limit with a stated assurance?	$\Pr(\mathcal{R}_\mathrm{yr}(\mathbf{q}) < \mathcal{R}^* \mid \mathbf{y}) \geq \kappa$	risk-informed acceptance: a tail of the risk posterior; the plug-in risk against the expected risk	25
What is the loss in a bad year?	the 95th percentile and the tail mean of the predictive year's loss	a quantile of a compound predictive: aleatory and epistemic variance split	25
How much risk does each design change remove?	$\mathcal{R}(a)$ ranked over the actions; $\Pr(L(a, \mathbf{z}) < L(a', \mathbf{z}) \mid \mathbf{y})$	the alternative drawings' risks, draw by draw	25
Which unit's unavailability matters most to the top event?	$I_{\mathrm{B},i} = \partial U/\partial q_i = U_i^+ - U_i^-$	the Birnbaum importance by the pivotal decomposition, with its band	26

Engineering question	Mathematical form	Derives to	Chapter
What share of the top event runs through each unit?	$I_{\mathrm{FV},i} = (U - U_i^-)/U$	the Fussell–Vesely importance, equal to the criticality importance; the cut-set forms bound it	26
How much worse if the unit were down, how much better if it never failed?	U_i^+/U, U/U_i^-	the risk achievement and reduction worths, and their identities with the other two measures	26
Are the rankings firm?	$\Pr(I_{\mathrm{FV},K} > I_{\mathrm{FV},C} \mid \mathbf{y})$	a threshold on the log-ratio of two shares; one exact integral	26
Why does the band of the top event use the same numbers?	$\mathrm{Var}(\ln U \mid \mathbf{y}) = \sum_j (g_j/U)^2 \Sigma_{jj}$	the gradient of the band as the importances summed over the units sharing an entry	26
Does a capacity read through a model meet its requirement?	$\Pr(G_1 \geq G^* \mid \mathbf{y}) \geq \kappa$	the acceptance rule as a normal tail of one entry, with the full posterior beside it; the one-sided bound	27
How strongly does the record favor the requirement over its complement?	posterior odds over prior odds of $G_1 \geq G^*$	the Bayes factor of two halves of one prior, a ratio of tails	27
How well does a test fix a capacity, and what limits it?	$\mathrm{sd}(g_1 \mid \tilde{y}_a)$ as n grows	reading noise over n against the supply's error over the rise	27
How likely is a test to accept a good or a bad path?	$\mathrm{OC}(\mathbf{z}_{\mathrm{t}}; a) = \Pr(\mathrm{accept} \mid \mathbf{z}_{\mathrm{t}}, a)$	the operating characteristic, a preposterior at a stated truth; producer's and consumer's risks	27
How long a test, at what load?	the smallest test meeting both risks	planning by the operating characteristic; the load is the lever	27
What does accepting on a point estimate get wrong?	the operating characteristic at $\kappa = \frac{1}{2}$	the consumer's risk of the point rule: one half at the requirement	27
Will the hall stay under the limit at a scanned load?	$\Pr(T_1^*(Q_1^*) < T_{\max} \mid \mathbf{y})$	the tail by load, on the route the propagation rule selects	28
How much margin is there, and how well is it known?	$p(T_{\max} - T_1^* \mid \mathbf{y})$	the margin's band, its variance quadratic in the load	28
How much load can be taken with an assurance?	$Q_\kappa = \max\{Q_1^* : \Pr(T_1^* < T_{\max} \mid \mathbf{y}) \geq \kappa\}$	the tail inverted: a quantile of the path's capacity	28

Engineering question	Mathematical form	Derives to	Chapter
What is the path's capacity, and is it at least the design load?	$p(Q_{\text{cap}} \mid \mathbf{y})$, $\Pr(Q_{\text{cap}} \geq Q \mid \mathbf{y})$	load–capacity interference: the same event as the temperature tail	28
Is the unit delivering what its plate says?	$p(c_{\text{ch}} \mid \mathbf{y})$; the evidence against $c_{\text{ch}} = 1$	a rank-one update and the Savage–Dickey ratio; a meter bias as the competing drawing	28
How much load does the chiller have in hand?	$p(P_{\text{pl}} c_{\text{ch}} - Q_1 - Q_2 \mid \mathbf{y})$	a linear derived quantity in which the meter bias cancels	28
What does a derate do to the margin at full demand?	$\Pr(T_1^* + (P^* - P_{\text{pl}} c_{\text{ch}})^+ / \dot{m} c_p < T_{\max} \mid \mathbf{y})$	the shortfall's partial expectation propagated through the supply	28