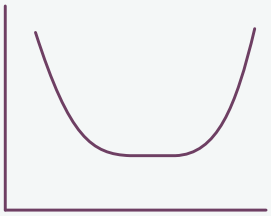
LESSON BOOK — VOLUME I



Reliability Engineering Volume I

Foundations, lifetime models, system logic, risk assessment and repairable systems

Twenty lessons · 509 pages

Ricardo Marquez, Ph.D.

ricardo-marquez.com

Reliability Engineering

Volume I

Foundations, lifetime models, system logic, risk assessment and repairable systems

A lesson sequence, written up as a book

Ricardo Marquez

October 14, 2026

How to read this book

This book is about one question: how an engineered system fails, how likely that is, what it costs, and what should be done about it. Each chapter expands one lesson of a sequence that builds the answer in the order the discipline itself uses,

function and failure → lifetime models → system logic → dependencies and risk → repair and availability → degradation and prognosis → power-system adequacy → measurements and credibility → decisions.

The sequence has nine modules and thirty-six lessons, published in two volumes. This first volume holds Parts I to V, Chapters 1 to 20: the classical reliability theory as the textbooks state it, from failure definitions and lifetime models through system logic, probabilistic risk assessment and repairable systems. The second volume holds Parts VI to IX, Chapters 21 to 36, which extend that theory to degradation and prognostics, power-system adequacy, the sufficiency of measurements, and engineering decisions. Chapter numbers run continuously across the two volumes, and a reference to a chapter of the other volume is written as such.

The book is written for a reader who builds or uses physics-based models of engineered systems and wants the concepts and the vocabulary of reliability engineering stated precisely. Every new quantity is therefore introduced in the same order: its name, its meaning, its units, its equation, and how it is estimated. Where the field uses a shorthand term, a *Jargon* box expands it once and the text then uses the expanded meaning. Every symbol is defined in the notation chapter that follows. The vocabulary is the standard one of the discipline, as fixed by Rausand, Barros and Høyland's *System Reliability Theory*, Modarres, Kaminskiy and Krivtsov's *Reliability Engineering and Risk Analysis*, Billinton and Allan's *Reliability Evaluation of Power Systems*, the NRC's *Fault Tree Handbook* and the other references each chapter names; no term is coined where a standard one exists.

Three systems recur throughout, with growing complexity as the methods allow: an industrial cooling-water pump, a data center cooling plant with three chillers of which two are required, and an electric transmission network whose line ratings depend on the weather. The numbers in the examples are illustrative, chosen to make each relation concrete, and are labeled as such.

One distinction recurs in every chapter: the physical system, which fails or performs; the model, which represents it under stated assumptions; and the evidence, which is never complete. Three kinds of uncertainty are kept apart throughout: aleatory variability in which unit fails and when, epistemic uncertainty about the parameters of the model, and uncertainty about whether the model's structure is adequate at all.

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 four exercises that can be solved from the chapter alone.

Contents

How to read this book **ii**

Notation and vocabulary **viii**

I Foundations of Reliability Engineering **1**

1 Introduction **2**
1.1 Reliability as an engineering question 2
1.2 Definition of reliability 4
1.3 Component and system reliability 7
1.4 Risk and availability 9
1.5 Reliability of the physical system 12
1.6 Worked example: chiller plant 16
1.7 What to retain 20
1.8 Exercises 21

2 Functions, boundaries and failure **23**
2.1 Required function and performance criterion 23
2.2 System boundaries and hierarchical failure 26
2.3 Multiple requirements and time horizons 29
2.4 Worked examples of boundaries and failure definitions 34
2.5 Functional decomposition and failure-mode analysis 38
2.6 Conditional models and the probabilistic formulation 39
2.7 What to retain 42
2.8 Exercises 43

3 Reliability, availability, maintainability and risk **45**
3.1 Four measures of dependability 45
3.2 Reliability as survival without failure 47
3.3 Availability of a repairable item 49
3.4 Mean times and maintainability 52
3.5 Risk and the frequency of interruptions 56
3.6 Power systems and the RAMS framework 59
3.7 What to retain 62
3.8 Exercises 64

4 Failure mechanisms and FMEA **65**
4.1 Failure cause, mechanism, mode and effect 65
4.2 Physical failure mechanisms 68
4.3 Mathematical failure models 70

4.4 Fatigue as physics-based prediction . . . 75
4.5 Failure mode and effects analysis . . . 79
4.6 Criticality and quantitative risk . . . 83
4.7 What to retain . . . 88
4.8 Exercises . . . 89

II Statistical Reliability Models 91

5 Lifetime distributions and hazard functions 92
5.1 Lifetime as a random variable . . . 92
5.2 Distribution, density and reliability . . . 94
5.3 The hazard function . . . 97
5.4 Conditional failure probability over an interval . . . 100
5.5 Hazard shapes and worked models . . . 102
5.6 Mean life and condition-dependent hazard . . . 107
5.7 What to retain . . . 111
5.8 Exercises . . . 112

6 Exponential, Weibull and lognormal models 114
6.1 The choice of a lifetime distribution . . . 114
6.2 The exponential distribution . . . 116
6.3 The Weibull distribution . . . 119
6.4 The lognormal distribution . . . 124
6.5 Comparing the three models . . . 128
6.6 Estimating and selecting a lifetime model . . . 130
6.7 What to retain . . . 135
6.8 Exercises . . . 136

7 Reliability data and statistical estimation 138
7.1 Reliability from incomplete operating records . . . 138
7.2 Censored lifetime data . . . 141
7.3 The Kaplan–Meier estimator . . . 143
7.4 Likelihood and maximum likelihood . . . 147
7.5 Statistical uncertainty . . . 151
7.6 Diagnostics and engineering credibility . . . 157
7.7 What to retain . . . 161
7.8 Exercises . . . 162

8 Bayesian reliability analysis 164
8.1 Bayesian inference of a failure rate . . . 164
8.2 Gamma–exponential conjugacy . . . 168
8.3 Posterior uncertainty and credible intervals . . . 171
8.4 Posterior predictive reliability . . . 174
8.5 Reliability estimation with zero observed failures . . . 176
8.6 Updating, populations and model credibility . . . 179
8.7 What to retain . . . 183
8.8 Exercises . . . 184

III System Reliability and Fault Trees 186

9 Reliability block diagrams and redundancy 187
9.1 System reliability from component reliability . . . 187
9.2 Series systems . . . 189
9.3 Parallel and k-out-of-n systems . . . 192
9.4 Active and standby redundancy . . . 197
9.5 Independence and common-cause failure . . . 202
9.6 Combined structures, importance and capacity . . . 203
9.7 What to retain . . . 208
9.8 Exercises . . . 209

10 Fault tree construction and logic 210
10.1 Fault tree analysis and the top event . . . 210
10.2 The logic of AND and OR gates . . . 213
10.3 Worked example: cooling-water flow . . . 217
10.4 Boolean algebra and fault-tree reduction . . . 219
10.5 Coherence and other gates . . . 222
10.6 Fault-tree construction in practice . . . 228
10.7 What to retain . . . 230
10.8 Exercises . . . 233

11 Minimal cut sets and quantification 235
11.1 Cut sets and minimal cut sets . . . 235
11.2 Exact quantification of the top event . . . 239
11.3 Rare-event approximation and bounds . . . 242
11.4 Minimal path sets . . . 245
11.5 Computational methods for large fault trees . . . 248
11.6 Application to the data center cooling plant . . . 253
11.7 What to retain . . . 257
11.8 Exercises . . . 259

12 Reliability importance measures 261
12.1 Component importance . . . 261
12.2 Birnbaum importance . . . 265
12.3 Fussell–Vesely, RAW and RRW . . . 267
12.4 Comparison of the four measures . . . 271
12.5 Importance and engineering decisions . . . 274
12.6 Limitations of importance measures . . . 278
12.7 What to retain . . . 280
12.8 Exercises . . . 282

IV Dependencies and Probabilistic Risk Assessment 284

13 Common-cause and dependent failures 285
13.1 Redundancy and the independence assumption . . . 285
13.2 Mathematical dependence . . . 287
13.3 Types of dependence . . . 292
13.4 Shared environmental conditions . . . 297
13.5 Load sharing and conditional hazards . . . 299
13.6 Engineering practice and quantitative risk . . . 304

13.7 What to retain . . . 308
13.8 Exercises . . . 310

14 Beta-factor and alpha-factor models 311
14.1 Common-cause failure in redundant groups . . . 311
14.2 Failure rates and failure probabilities . . . 313
14.3 The beta-factor model . . . 315
14.4 The alpha-factor model . . . 319
14.5 Beta against alpha . . . 326
14.6 Estimating the parameters from data . . . 329
14.7 What to retain . . . 333
14.8 Exercises . . . 335

15 Event trees and accident sequences 336
15.1 Event trees and initiating events . . . 336
15.2 Event-tree branching and sequence probabilities . . . 339
15.3 Worked example: data center power continuity . . . 342
15.4 Worked example: transmission-line outage . . . 346
15.5 Fault-tree links, dependence and recovery timing . . . 348
15.6 Expected consequence, risk and model credibility . . . 353
15.7 What to retain . . . 357
15.8 Exercises . . . 359

16 Integrated probabilistic risk assessment 360
16.1 The integrated assessment . . . 360
16.2 System, functions and dependencies . . . 363
16.3 The integrated event tree . . . 366
16.4 Risk importance and decisions . . . 371
16.5 Uncertainty propagation . . . 375
16.6 Credibility and other systems . . . 379
16.7 What to retain . . . 383
16.8 Exercises . . . 385

V Repairable Systems and Availability 387

17 Repairable systems and counting processes 388
17.1 Repairable systems and recurrent failures . . . 388
17.2 Recurrent failure data and ROCOF . . . 391
17.3 The homogeneous Poisson process . . . 397
17.4 The nonhomogeneous Poisson process . . . 400
17.5 Repair assumptions . . . 403
17.6 Estimation, model selection and applications . . . 406
17.7 What to retain . . . 413
17.8 Exercises . . . 414

18 Renewal processes and maintenance models 416
18.1 The effect of repair on future failures . . . 416
18.2 Perfect repair and renewal processes . . . 419
18.3 Minimal repair . . . 422
18.4 Imperfect repair and virtual age . . . 424
18.5 Renewal-reward theory and age replacement . . . 428

18.6 Optimal age replacement 431
18.7 What to retain 439
18.8 Exercises 440

19 Markov models and state transitions **441**
19.1 From failure counting to system states 441
19.2 Markov property and transition rates 443
19.3 The single component and the Kolmogorov equations 445
19.4 Two repairable pumps 448
19.5 Steady state, repair resources and failure frequency 453
19.6 Extensions, applications and validation 458
19.7 What to retain 463
19.8 Exercises 465

20 Availability and maintenance optimization **467**
20.1 Availability as an engineering objective 467
20.2 The cooling-plant model and repair capacity 471
20.3 Spare parts and repair logistics 475
20.4 Comparison of maintenance strategies 476
20.5 Scheduling and spare-parts inventory 479
20.6 Investment choices under uncertainty 486
20.7 What to retain 489
20.8 Exercises 492

Notation and vocabulary

Every symbol below is used with the same meaning in every chapter. Time is measured in operating hours unless a chapter states otherwise.

Symbol	Unit	Meaning
T	h	Random time to failure of a nonrepairable item, measured from a defined starting condition; $T \geq 0$.
t	h	A specified operating duration, or the running time variable.
H	h	Operating horizon or mission time over which a requirement must hold.
$R(t)$	–	Reliability (survivor) function, $\Pr(T > t)$.
$F(t)$	–	Cumulative failure probability, $\Pr(T \leq t) = 1 - R(t)$.
$f(t)$	h^{-1}	Failure density, $\mathrm{d}F/\mathrm{d}t$.
$\lambda(t)$	h^{-1}	Hazard function (instantaneous failure rate), $f(t)/R(t)$.
$\Lambda(t)$	–	Cumulative hazard, $\int_0^t \lambda(u)\,\mathrm{d}u$, so that $R(t) = \exp[-\Lambda(t)]$.
λ	h^{-1}	Constant failure rate of the exponential model.
μ	h^{-1}	Constant repair rate, 1/MTTR.
$R(t + \Delta t \mid t)$	–	Conditional reliability: survival over a further Δt given survival to t, $R(t + \Delta t)/R(t)$.
MTTF, MTBF, MTTR	h	Mean time to failure (nonrepairable, or operating duration between restorations); mean time between failures (repairable, including restoration); mean time to repair or restore.
MRL(t)	h	Mean residual life at age t, $\mathbb{E}[T - t \mid T > t]$.
A, $A(t)$, A_{mission}	–	Steady-state availability MTTF/(MTTF + MTTR); instantaneous (point) availability at time t; point availability at the end of a mission of length H, $A(H)$, for an item that starts in the operating state.
R_{system}, R_i	–	Reliability of the system; of component i.
k-out-of-n	–	A system of n units that functions when at least k of them function.
$\mathcal{R}$	R$/yr	Risk, $\sum_i \Pr(S_i)\, C(S_i)$ over a partition of outcomes S_i; expected annual loss when C is monetary.
$C(F)$	R$	Consequence of failure event F, conditional on its occurrence.
$X(t)$	varies	Uncertain system state or inputs at time t (load, weather, equipment condition).
$g(x)$	varies	Performance-margin function; $g \geq 0$ means the requirement is met.

Symbol	Unit	Meaning
$L(t)$, $C(t)$	varies	Required load or demand; available capacity, in the units of the function (m^3/h, kW, A).
$C_{\text{available}}(t)$	varies	Available capacity of a set of units, $\sum_i S_i(t)\, C_i(t)$.
$S_i(t)$	–	Status of unit i at time t: 1 available, 0 unavailable.
$T_{\max}$	°C	Upper temperature limit of a conductor, a room or a component.
I, $R(T)$	A, Ω	Current; temperature-dependent resistance of a conductor (never the reliability function, which is R).
mc_p	J/K	Heat capacity of the modeled conductor length.
q_{solar}, q_{conv}, q_{rad}	W	Solar heat gain; convective and radiative heat loss of a conductor.
S, O, D	–	FMEA severity, occurrence and detection ratings on a 1–10 scale.
RPN	–	Risk priority number, $S \cdot O \cdot D$.
β, η	–, h	Shape and scale parameters of the Weibull distribution.
σ	varies	Standard deviation; stress amplitude, by context.
N	–	Number of components, of load cycles, of trials, by context.
Statistical estimation (Chapters 6–8)		
$\hat{R}(t)$, $\hat{\lambda}$	–, h^{-1}	Estimates from data (Kaplan–Meier survival estimate; maximum likelihood estimate of the failure rate).
r, T_{tot}	–, h	Number of observed failures; total exposure (time on test) of all units, censored ones included.
$L(\theta)$, $\ell(\theta)$, $\mathcal{I}(\theta)$	varies	Likelihood, log-likelihood and Fisher information of a parameter θ.
Gamma(a, b)	–	Gamma distribution with shape a and rate b; prior and posterior of a failure rate.
μ, σ (lognormal)	–	Mean and standard deviation of $\ln T$ in the lognormal model; in that use μ is not the repair rate.
$\Gamma(\cdot)$, $\Phi(\cdot)$	–	Gamma function; standard normal cumulative distribution function.
System logic and probabilistic risk assessment (Chapters 9–15)		
$\phi(\mathbf{x})$, x_i	–	Structure function and component state indicators (1 operating, 0 failed; or the failure indicator, as stated).
q_i	–	Probability of basic event i (component failure or unavailability).
Q_{top}, Q_{sys}	–	Probability of the top event; system failure probability or unavailability.
C_j, P_j	–	Minimal cut set j; minimal path set j (chapter text says when C and P name components instead).
I_i^{B}, I_i^{FV}	–	Birnbaum and Fussell–Vesely importance of component i.
RAW_i, RRW_i	–	Risk achievement worth; risk reduction worth.
β (beta factor), α_k	–	Common-cause beta factor (fraction of a component's failure rate that is common cause; not the Weibull shape); alpha factor, the fraction of failure events involving exactly k components.

Symbol	Unit	Meaning
λ_T, λ_I, λ_C	h^{-1}	Total, independent and common-cause failure rates of one component in a common-cause group.
f_{IE}, f_{S_k}	yr^{-1}	Initiating-event frequency; frequency of accident sequence k.
$\Pr(S_k \mid I)$	–	Probability of sequence k given the initiating event.
LOLE, EUE	h/yr, MWh/yr	Loss-of-load expectation; expected unserved energy.
Repairable systems and availability (Chapters 17–20)		
$N(t)$, T_j, X_j	–, h, h	Number of failures of a repairable item by time t; the j-th failure epoch; the j-th interfailure time.
$W(t)$, $w(t)$	–, h^{-1}	Cumulative intensity (mean function) $\mathbb{E}[N(t)]$ of a counting process; its rate of occurrence of failures (ROCOF), $\mathrm{d}W/\mathrm{d}t$, which is not the hazard of one lifetime.
$M(t)$, $m(t)$	–, h^{-1}	Renewal function (mean number of renewals) and renewal density of a perfect-repair process.
b, η (power law)	–, h	Shape and scale of the power-law intensity $w(t) = (b/\eta)(t/\eta)^{b-1}$ of a nonhomogeneous Poisson process.
V_n	h	Virtual age of an item after its n-th repair (Kijima models).
$P_i(t)$, $\boldsymbol{\pi}$	–	Probability of being in state i at time t of a Markov model; its stationary distribution.
$\mathbf{Q}$, q_{ij}	h^{-1}	Generator (transition-rate) matrix of a continuous-time Markov chain and its off-diagonal rates; not the top-event probability Q_{top} nor the basic-event probability q_i.
ρ	–	Ratio of failure rate to repair rate, λ/μ; by context, a discount rate.
f_{F}, MDT	yr^{-1}, h	Frequency of entering the failed state; mean down time per failure, so that unavailability equals $f_{\mathrm{F}} \cdot \mathrm{MDT}$.
T_p, $C(T_p)$	h, R$/h	Preventive replacement age; long-run cost rate of the age-replacement policy.
c_p, c_f, c_D	R$, R$, R$/h	Cost of a preventive replacement, of a failure replacement, of one hour of downtime.
s, $\nu\tau$	–, –	Spare-parts stock level; expected spare demand over a lead time τ at demand rate ν.
Degradation, prognostics and health management (Chapters 21–24)		
$D(t)$, D_0, D_{crit}	varies	Degradation (damage) state, its initial value and the failure threshold; failure is the first crossing. Not the FMEA detection rating D.
μ_D, σ_D, $B(t)$	varies	Drift and diffusion of the Wiener degradation model $D(t) = D_0 + \mu_D t + \sigma_D B(t)$; standard Brownian motion. μ_D is not the repair rate.
α_D, β_D	–, varies	Shape rate and scale of a gamma degradation process.
T_{f}, $\mathrm{RUL}(t)$, $\mathrm{RUL}_\alpha(t)$	h, d	First-passage (failure) time; remaining useful life at time t, $T_{\mathrm{f}} - t$ given survival; its α-quantile.

Symbol	Unit	Meaning
e, z	varies, –	Residual, observed minus model-predicted value; standardized residual.
$H_0, H_1, \ldots$	–	Diagnostic hypotheses (healthy, each fault, each sensor fault) with posterior probabilities $\Pr(H_i \mid \mathbf{y})$.
$\mathcal{H}$	bit	Shannon entropy of the hypothesis probabilities; expected information gain of a test.
Y^+, Y^-	–	Positive and negative outcome of an inspection, characterized by sensitivity and specificity.
τ_{L}, p^*	d, –	Action lead time; decision threshold on the failure probability.
EVSI, EVPI	R$	Expected value of sample (test) information; of perfect information.
Power-system adequacy (Chapters 25 onward)		
L_t, C_t, S_t	MW	Hourly demand; available capacity; shortfall $(L_t - C_t)^+$.
FOR_i, U_i	–	Forced outage rate of unit i: its unavailability, a probability, not a rate.
LOLP_t, LOLH	–, h/yr	Loss-of-load probability in interval t; loss-of-load hours, the LOLE on an hourly basis.
X, $P(X \geq x)$	MW, –	Capacity on outage and its cumulative probability in a capacity outage probability table (COPT).
K_i, $\mathrm{PTDF}_{\ell,k}$, $\mathrm{LODF}_{\ell,m}$	MW, –, –	Rating of unit i; power transfer and line outage distribution factors of line ℓ.
θ_i, V_i, x_{ij}, P_{ij}, $P_{ij}^{\max}$	rad, pu, pu, MW	Bus voltage angle and magnitude; line reactance; DC line flow $(\theta_i - \theta_j)/x_{ij}$; line rating.
$\mathbf{B}$	pu	DC susceptance matrix of the network.
λ_{b}, N_{c}	–, –	Branching ratio (mean offspring number) of a cascade model; cascade size.
τ_{th}	min	Thermal time constant of a conductor.
State estimation and identifiability (Chapters 29–31)		
$\mathbf{x}$, $\mathbf{y}$	varies	State vector; measurement vector (written $\mathbf{z}$ in the power-system literature), $\mathbf{y} = h(\mathbf{x}) + \boldsymbol{\varepsilon}$.
$\mathbf{R}$, $\mathbf{W}$	varies	Measurement error covariance (bold; not the reliability function) and its inverse, the weight matrix.
$\mathbf{H}$, $\mathbf{G}$	varies	Measurement Jacobian $\partial h/\partial \mathbf{x}$ (bold; not the horizon) and the gain (information) matrix $\mathbf{H}^{\mathsf{T}}\mathbf{W}\mathbf{H}$.
$J(\mathbf{x})$, $\hat{\mathbf{x}}$	–, varies	Weighted least-squares objective; its minimizer, the state estimate.
$\mathbf{e}$, $\mathbf{S}$, $\boldsymbol{\Omega}$, e_i^{N}	varies	Residual $\mathbf{y} - h(\hat{\mathbf{x}})$; residual sensitivity matrix; residual covariance $\mathbf{SR}$; normalized residual $e_i/\sqrt{\Omega_{ii}}$.
$\boldsymbol{\theta}$, $\mathbf{S}_\theta$	varies	Parameter vector (bold; not the bus angles) and its sensitivity matrix $\partial h/\partial \boldsymbol{\theta}$.
$\sigma_1 \geq \sigma_2 \geq \ldots$, κ	–	Singular values of the scaled sensitivity matrix; condition number $\sigma_1/\sigma_{\min}$.
$\delta(\cdot)$, B_{12}	varies, –	Model discrepancy term; Bayes factor between models 1 and 2.
Value of information and engineering decisions (Chapters 32–36)		
$L(a, H)$, $\rho(a)$	R$	Loss of action a when hypothesis H holds; its expected loss under the current probabilities.

Symbol	Unit	Meaning
C_{now}, C_{sample}, C_{perfect}	R$	Minimum expected loss with current information; with the outcome of a test; with the true hypothesis known.
EIG, $I(H;Y)$, D_{KL}	bit or nat	Expected information gain of a test, equal to the mutual information between hypothesis and outcome; Kullback–Leibler divergence.
$\mathcal{R}(a)$, $\Delta\mathcal{R}(a)$	varies	Risk after intervention a; its reduction from the baseline $\mathcal{R}_0$.
x_i, c_i, B	–, R$/yr, R$/yr	Selection variable of project i (1 selected, 0 not); its annualized cost; the budget.
v, H_{rel}	R$/MWh, h/yr	Consequence value of unserved energy; hours per year for which an analyzed operating condition applies.
CVaR_α	varies	Conditional value at risk: the expected loss in the upper $(1-\alpha)$ tail.
A_{IT}, $Y(t)$	–	Availability of the required IT service; its indicator at time t.
C_{eff}, t_{crit}	MJ/K, min	Effective thermal capacitance of a room; time for its temperature to reach the limit after a loss of cooling.
λ_{DU}, T_{p}	h^{-1}, h	Dangerous undetected failure rate of a safety function; proof-test interval (not the replacement age of Chapter 18).
PFD_{avg}, PFH	–, h^{-1}	Average probability of dangerous failure on demand; average frequency of dangerous failure per hour.

Vocabulary

The book uses the standard vocabulary of reliability engineering and probabilistic risk assessment, as fixed by the textbooks named in the preface. Where a term has a common shorthand, the shorthand is given in parentheses once and the full term is used thereafter.

Item A component, subsystem or system treated as a unit of analysis.

Required function What an item must do for the system to count as operating; stated as a measurable performance criterion.

Failure The termination of an item's ability to perform its required function; a functional event, which need not coincide with physical breakage.

Fault The state of an item after failure, or a defect that may lead to failure; a fault can exist without a failure having occurred.

Failure mode The manner in which the failure is observed (no flow, low flow, leak, spurious trip).

Failure mechanism The physical, chemical or electrical process that produces the failure (fatigue, corrosion, wear, electromigration).

Failure cause The circumstance that activates the mechanism (overload, contamination, misuse).

Failure effect The consequence of the mode at the next level of the system hierarchy.

System boundary The set of items and interfaces inside the reliability model; what is outside is an input or an assumption.

Reliability Probability that an item performs its required function under specified conditions for a specified period.

Availability Probability that an item is in an operating state at a given time, or the long-run fraction of time it is operating, restoration included.

Maintainability Probability that a failed item is restored within a specified time under specified procedures.

Risk The combination of the probability of outcomes and their consequences.

RAMS Reliability, availability, maintainability and safety, considered together.

Hazard function (failure rate) The conditional failure intensity at age t given survival to t; constant for the exponential model, increasing with wear-out, decreasing with early failures.

Bathtub curve The hazard shape with an early decreasing part, a constant middle and an increasing wear-out part.

Mission reliability Reliability over a stated mission time without credit for repair.

Repairable system An item that is restored to operation after failure; described by counting and renewal processes rather than a single lifetime.

FMEA, FMECA Failure mode and effects analysis; the same with a criticality analysis.

Fault tree analysis (FTA) A deductive logic model of the combinations of events that produce a defined top event.

Minimal cut set A smallest set of component failures sufficient to cause the top event.

Common-cause failure Failure of several items from one shared cause, which defeats the independence that redundancy assumes.

Importance measure A quantity that ranks components by their contribution to system unreliability (Birnbaum, Fussell–Vesely, risk achievement worth, risk reduction worth).

Probabilistic risk assessment (PRA) The scenario-based quantification of risk: initiating events, event trees, fault trees and consequences.

Aleatory and epistemic uncertainty Variability inherent in the process, and lack of knowledge about its parameters or structure.

Censoring A lifetime observation that ends before failure, so that only a lower bound of T is known (right-censoring); interval and left censoring bound T from both sides or from above.

Kaplan–Meier estimator The nonparametric product-limit estimate of the reliability function from censored data.

Maximum likelihood estimate (MLE) The parameter value that maximizes the likelihood of the observed failures and censored records.

Credible interval A Bayesian interval that contains the parameter with stated posterior probability; distinct from a confidence interval.

Posterior predictive reliability The reliability of a new unit obtained by averaging the model over the posterior of its parameters.

Structure function The Boolean function that gives the system state from the component states.

Reliability block diagram A success-logic diagram in which the system operates when a path of operating blocks connects input to output.

Cut set, path set A set of component failures that causes system failure; a set of operating components that keeps the system operating. Minimal when no proper subset qualifies.

Rare-event approximation The sum of minimal-cut-set probabilities as an estimate of the top-event probability; an upper bound that is accurate when the probabilities are small.

Inclusion–exclusion The exact expansion of the probability of a union of events in alternating sums over intersections.

Coherent system A system in which no component failure can restore operation and every component matters.

Common-cause component group A set of redundant, similar components exposed to the same common causes; the unit to which beta and alpha factors apply.

Coupling factor The shared property (design, environment, procedure) through which one root cause affects several components.

Initiating event The event that starts an accident sequence; characterized by a frequency per year.

Event tree An inductive model that branches on the success or failure of each protection function after an initiating event.

Accident sequence, end state One path through an event tree; the consequence category in which it ends.

Frequency–consequence curve The annual frequency of exceeding each consequence level; the risk profile beyond its expectation.

Integrated PRA A risk assessment that links initiating-event frequencies, fault trees, event trees, consequences and uncertainty into one traceable model; organized by the risk triplet (what can go wrong, how likely, with what consequences).

Counting process, ROCOF The number of failures of a repairable item by time t; the rate of occurrence of failures, the derivative of its mean, as distinct from the hazard of one lifetime.

Homogeneous and nonhomogeneous Poisson process (HPP, NHPP) Counting processes with a constant, respectively time-varying, intensity; the NHPP with a power-law intensity is the Crow–AMSAA model.

Perfect, minimal and imperfect repair Repair that returns an item to as-good-as-new, to as-bad-as-old, or to a state between the two (virtual age).

Renewal process The counting process of a perfectly repaired item: independent identically distributed interfailure times.

Age replacement The maintenance policy that replaces an item at failure or at age T_p, whichever comes first; optimized through the renewal-reward theorem.

Continuous-time Markov chain (CTMC) A state model with exponential holding times and a generator matrix of transition rates; solved by the Kolmogorov forward equations or its stationary distribution.

Frequency and duration The description of unavailability as the product of the failure frequency and the mean down time.

Degradation, failure threshold The gradual loss of a property that fixes an item's ability to function; failure is the first crossing of the threshold by the degradation state.

Prognostics and health management (PHM) Detection, diagnosis and prognosis of equipment condition from measurements, and the maintenance decisions built on them.

Remaining useful life (RUL) The time from now until the degradation state reaches its threshold, a random variable conditional on the present evidence.

Condition indicator, residual A quantity computed from measurements that tracks condition; the difference between a measurement and its model prediction, which detection and diagnosis test.

Fault detection and isolation (FDI) Deciding that a fault is present, and which one; with bad-data detection, the discipline's names for separating physical faults from sensor faults.

Value of information The expected reduction in decision loss from a test or inspection before deciding; sample information (EVSI) and perfect information (EVPI).

Generation (resource) adequacy Whether installed resources suffice to meet demand at every hour under random outages; measured by LOLP, LOLE and EUE.

Forced outage rate (FOR) The unavailability of a generating unit, the fraction of time it is out of service on forced outage; a probability.

Capacity outage probability table (COPT) The distribution of capacity on outage for a fleet, built by recursive convolution of the units' two-state models.

Multistate unit, derated state A generating unit modeled with partial-capacity states between full availability and outage.

Contingency, N−1 criterion The outage of one or more network elements; the requirement that the system withstand the loss of any single element without violating limits.

Cascading failure A sequence of dependent outages in which each failure raises the stress on, and the failure probability of, the remaining elements.

Observability Whether the states of interest can be inferred from the available measurements; the rank condition on the measurement Jacobian.

State estimation The weighted least-squares (or robust, or Bayesian) inference of the system state from redundant measurements.

Critical measurement A measurement whose removal makes the system unobservable; its residual is always zero, so its error cannot be detected.

Bad-data detection and identification The chi-square test on the weighted residual sum, and the largest normalized residual test that names the suspect measurement.

Topology error A mismatch between the network model's switching state and the physical network, which a residual test can confuse with a sensor fault.

Identifiability Whether the parameters of a model can be determined from the measurements, structurally (from the model) and practically (given the noise and the design).

Model discrepancy The systematic difference between a model's prediction and the physical system, distinct from measurement error.

Expected information gain The expected reduction in the entropy of the hypothesis probabilities from a test; a measure of knowledge, not of decision value.

Counterfactual risk reduction The change in risk from a hypothetical intervention on the system, as distinct from conditioning on an observed state.

Portfolio optimization The selection of projects under a budget, with system risk recomputed for each combination because project benefits interact.

Verification, validation Whether a model is implemented and solved correctly; whether it represents the physical system adequately for its intended use.

Traceability The documented connection from a requirement through the model, its inputs and evidence to the result and the decision.

Safety integrity level (SIL) The band of required failure measures for a safety function, from 1 to 4.

Thermal ride-through The time a room's thermal inertia allows between a loss of cooling and the crossing of the temperature limit.

Part I

Foundations of Reliability Engineering

Chapter 1

Introduction to Reliability Engineering

1.1 Reliability as an engineering question

Engineered systems are built to do something: move water, remove heat, carry electric power. They are expected to keep doing it, and they sometimes stop. Whether a system will keep doing what it was built for, over a period that matters to someone, is a question asked every day by plant managers, facility operators and system operators, and the answer is never certain. Reliability engineering is the discipline that states such questions precisely and answers them with probabilities. The question of this chapter is

what it means for an engineered system to be reliable, and how that reliability is quantified.

The chapter answers it in six steps. It first poses three engineering situations and extracts their common structure. It then defines reliability, in words and as a probability, and separates the reliability of a component from that of a system. It distinguishes the probability of failure from risk, which adds consequences, and from availability, which adds repair. It shows that reliability is a property of the physical system, which can fail by missing a performance requirement without breaking, and lists the ways in which failure probabilities are obtained. Finally, it works one example through, a cooling plant of three chillers of which two are required, and states what that calculation leaves out.

Jargon: item, required function, failure, reliability, availability, risk, MTTF, MTTR

An *item* is a component, subsystem or system treated as a unit of analysis. Its *required function* is what it must do for the system to count as operating, stated as a measurable performance criterion, such as "deliver at least $100\,\mathrm{m}^3/\mathrm{h}$ of cooling water." A *failure* is the termination of the item's ability to perform its required function; it is a functional event and need not coincide with physical breakage. *Reliability* is the probability that an item performs its required function under specified conditions for a specified period. *Availability* is the probability that an item is in an operating state at a given time, or the long-run fraction of time it is operating, restoration after failure included. *Risk* is the combination of the probabilities of outcomes and their consequences. *MTTF* (mean time to failure) is the expected operating duration before a failure; *MTTR* (mean time to repair or restore) is the expected duration of the restoration that follows. Both are measured in hours.

Key idea

Reliability is not a number attached to a piece of equipment. It is the probability of a precisely defined event, the item performing a stated function under stated conditions over a stated period, and it changes when any of the four changes; system reliability then follows from the arrangement of the components, and risk and availability are different quantities built on top of it.

1.1.1 Three engineering situations

Three situations, which recur throughout the book with growing complexity, show the kind of question the discipline answers. In the first, a manufacturing facility runs a pump continuously to supply cooling water to a process. The pump has been operating for five years, and the plant manager asks for the probability that it will continue to function for another year. The question is about one item, about a future period, and about an item that has already survived a long time, which matters for the answer (Section 1.2.5).

In the second, a data center is cooled by three chillers, of which only two are required to satisfy the cooling demand. The operator asks whether the facility will continue operating if one chiller fails, and what the probability is of losing sufficient cooling capacity. The question is about a system of identical units with a spare, and about a capacity requirement rather than the survival of any single unit.

In the third, an electric transmission network consists of generators, lines, substations and protective devices. The system operator asks for the probability that equipment failures will prevent the network from supplying its required electrical demand. The question is about a large system whose components interact through the network, and whose ability to carry power depends on the weather as well as on the equipment (Section 1.5.1).

1.1.2 The common structure

The three situations differ in scale and in physics, but they share a structure of four elements. A system is expected to perform a function. Some event, or some combination of events, can prevent it from performing that function. The occurrence of those events is uncertain. Engineering decisions can change their probability or their consequences. Table 1.1 lists each element for the three situations. Its rows are the four elements and its columns the three systems; the entries in the last two rows are typical examples, not an exhaustive list.

Table 1.1: The common structure of three reliability questions. The events, uncertainties and decisions are typical examples.

Element	Cooling-water pump	Chiller plant (two of three)	Transmission network
Function	deliver cooling water to the process	remove the data center's cooling load	supply the required electrical demand
Events that prevent it	pump failure; flow below the requirement	failure of two or more chillers, or of a dependency they share	failures of generators, lines, substations or protective devices; ratings reduced by weather
Uncertainty	when the pump will fail, given five years of service	which chillers fail, when, and the load at the time	which equipment fails; weather; demand
Decisions	maintenance, replacement, a spare pump	redundancy, maintenance schedule, load management	reinforcement, operating limits, redispatch of generation

Reliability engineering provides the methods for analyzing questions of this structure. Each element of the structure becomes a part of the analysis: the function becomes a success criterion, the preventing events become failure events, the uncertainty becomes a probability model, and the decisions are compared by how much they change the probabilities and the consequences. The rest of the chapter defines these parts in turn.

1.2 Definition of reliability

Reliability is defined in words before it is defined as a probability, because the words fix which event the probability is about. This section states the definition and its four elements, writes it as a function of time, and separates survival from a starting condition from survival given that an item has already lasted some time.

1.2.1 The four elements of the definition

A definition that serves engineering work is the following: *reliability is the probability that an item will perform its required function under specified conditions for a specified period of time.* The definition has four elements, and a reliability statement is complete only when each of them is specified. Table 1.2 lists them. Each row is one element; the middle column is the question the element answers, and the right column gives its value for the cooling-water pump.

Table 1.2: The four elements of the definition of reliability, with the values for the cooling-water pump (illustrative).

Element	Engineering question	Pump example
Item	what is being evaluated	the cooling-water pump
Required function	what it must accomplish	deliver at least $100\,\mathrm{m^3/h}$
Operating conditions	under what circumstances	specified temperature, pressure and fluid
Time	over what period	the next 1000 operating hours

It follows that reliability is not an intrinsic numerical property of a piece of equipment, independent of its application. The same pump may be highly reliable under moderate operating conditions and considerably less reliable when it is exposed to high temperatures, to cavitation (the formation and collapse of vapor bubbles where the local pressure falls below the vapor pressure of the liquid) or to abrasive particles in the fluid. Changing the conditions changes the probability, because it changes the rate at which the physical processes that end the pump's function proceed.

1.2.2 Functional failure without physical breakdown

The required function is the element most often left implicit, and it is the one that decides what counts as a failure. A pump may continue to rotate and still fail to deliver the required flow: a worn impeller, a partly blocked suction strainer or a leaking seal can reduce its flow to $85\,\mathrm{m^3/h}$ while the motor runs normally. Against a requirement of $100\,\mathrm{m^3/h}$, that pump has failed, although an inspection that checks only whether it runs would report it as working.

A component therefore does not need to stop operating to fail its required function. Failure is defined relative to the function, and the function is defined by a measurable performance criterion. The distinction is central to the study of physical systems, in which performance degrades continuously before any part breaks, and it returns in Section 1.5.2 in the form of a

performance-based failure event. Chapter 2 treats the definition of functions, system boundaries and failure events in full.

Principle 1.1. *Reliability is function-dependent: successful operation must be defined, as a measurable criterion, before failure can be defined, and an item that runs but misses its criterion has failed.*

1.2.3 The reliability function and the cumulative failure probability

Let T denote the random time to failure of a nonrepairable item, measured in operating hours from a defined starting condition, such as the moment it is put into service in working order; $T \geq 0$. For a specified operating duration t, the reliability of the item is the probability that it survives beyond t,

$$R(t) = \Pr(T > t). \tag{1.1}$$

The function $R(t)$ is called the reliability function, or the survivor function. Its complement is the cumulative failure probability, the probability that the item has failed by time t,

$$F(t) = \Pr(T \leq t) = 1 - R(t). \tag{1.2}$$

The events $\{T > t\}$ and $\{T \leq t\}$ are mutually exclusive and together certain, so their probabilities add to one,

$$R(t) + F(t) = 1. \tag{1.3}$$

In the language of probability, F is the cumulative distribution function of T and R its complement; both are dimensionless and lie between 0 and 1.

As an example, suppose a pump operates successfully for 1000 hours with probability 0.95. Then $R(1000\,\mathrm{h}) = 0.95$ and $F(1000\,\mathrm{h}) = 1 - 0.95 = 0.05$: there is a 5 % probability that the pump fails during those 1000 hours. The number 0.95 is attached to the duration 1000 h; over 2000 h the same pump has a different reliability, which cannot be larger (equation (1.4)). Reliability always refers to a time horizon, and a statement that a pump "has 95 % reliability" without the interval is incomplete.

Reliability cannot increase with the duration. For two durations $t_1 < t_2$, an item that survives beyond t_2 has also survived beyond t_1, so the event $\{T > t_2\}$ is contained in the event $\{T > t_1\}$ and

$$R(t_2) = \Pr(T > t_2) \leq \Pr(T > t_1) = R(t_1). \tag{1.4}$$

The reliability function of a lifetime model is therefore nonincreasing. It starts at $R(0) = 1$ when the item starts in working order, and it tends to zero as t grows if the item is certain to fail eventually.

1.2.4 Reliability as a function of time

Figure 1.1 draws one reliability function. The horizontal axis is the operating time t in hours, from 0 to 10 000; the vertical axis is a probability. The solid curve is the reliability $R(t)$, the dashed curve the cumulative failure probability $F(t) = 1 - R(t)$, and the dots mark $R(t)$ at every 1000 hours. The curve is an illustrative exponential model,

$$R(t) = \mathrm{e}^{-t/10\,000\,\mathrm{h}}, \tag{1.5}$$

which describes an item whose failure rate does not change with age, a constant hazard of $10^{-4}\,\mathrm{h}^{-1}$; its mean time to failure is 10 000 h. The model is the simplest standard lifetime model, and Chapter 5 derives it from the hazard function.

The figure shows the two functions mirror each other about 0.5, as equation (1.3) requires. The reliability falls steadily: 0.905 at 1000 h, 0.607 at 5000 h, and 0.368 at 10 000 h, the mean time to failure. More than a third of such items therefore outlast their mean life, and almost two thirds fail before it: for the exponential model, $F(\text{MTTF}) = 1 - \text{e}^{-1} = 0.632$.

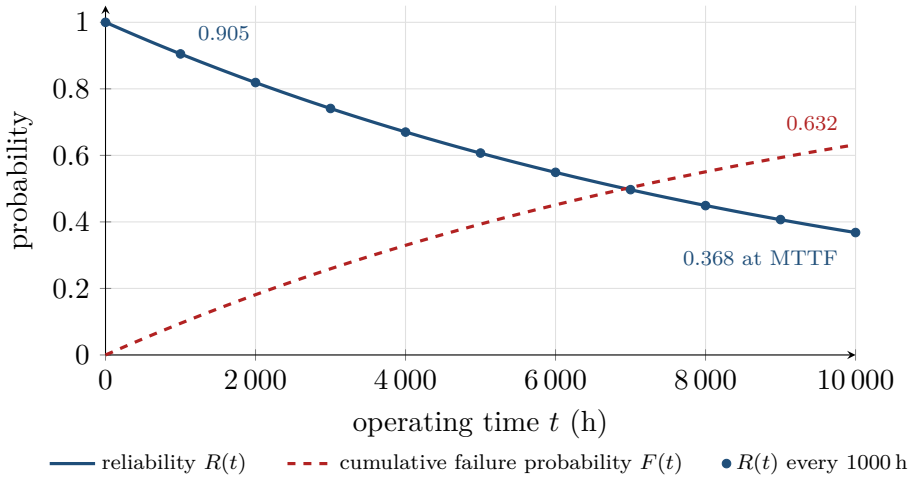


Figure 1.1: Illustrative reliability function $R(t) = \text{e}^{-t/10\,000\,\text{h}}$ (equation (1.5)), an exponential model with constant hazard $10^{-4}\,\text{h}^{-1}$ and MTTF = 10 000 h, and its complement $F(t) = 1 - R(t)$. Reliability decreases with operating time; at $t =$ MTTF it is 0.368.

1.2.5 Conditional reliability

A reliability function cannot increase with the duration, but a reliability prediction can change when new information about the item's condition arrives. The commonest piece of information is that the item has already survived for some time. Two quantities must then be kept apart. The reliability $R(t)$ describes survival over a specified duration measured from the starting condition. The conditional reliability $\Pr(T > t + \Delta t \mid T > t)$ describes survival over a further duration Δt given that the item has already survived to time t. By the definition of conditional probability, and because $\{T > t + \Delta t\}$ is contained in $\{T > t\}$,

$$R(t + \Delta t \mid t) = \Pr(T > t + \Delta t \mid T > t) = \frac{\Pr(T > t + \Delta t)}{\Pr(T > t)} = \frac{R(t + \Delta t)}{R(t)}. \tag{1.6}$$

The two answer different questions. The plant manager of the first situation in Section 1.1.1 asks the conditional one: the pump has survived five years, and the question is about the next year.

Table 1.3 evaluates both quantities for two illustrative lifetime models with the same mean time to failure, 50 000 h, for a pump in continuous operation (1 yr = 8760 h). The first is the exponential model $R(t) = \text{e}^{-t/50\,000\,\text{h}}$; the second is a Weibull model with shape $\beta = 2$ and scale $\eta = 56\,419$ h, $R(t) = \exp[-(t/\eta)^{\beta}]$, a standard model of an item that wears out, treated in a later chapter on lifetime distributions. Each row is an age t. For each model the table gives the unconditional reliability $R(t)$ and the conditional reliability over the next year. As an example of one entry, the Weibull pump at five years has

$$R(5\,\text{yr}) = \exp\left[-\left(\frac{43\,800}{56\,419}\right)^2\right] = 0.5473, \qquad R(6\,\text{yr}) = \exp\left[-\left(\frac{52\,560}{56\,419}\right)^2\right] = 0.4198,$$

and so a conditional reliability for the sixth year of $0.4198/0.5473 = 0.7671$.

Table 1.3: Unconditional reliability $R(t)$ and conditional reliability over the next year, $R(t + 8760\,\text{h} \mid t)$ (equation (1.6)), for two illustrative pump models with MTTF = 50 000 h: exponential, and Weibull with $\beta = 2$, $\eta = 56\,419$ h.

	Exponential		Weibull, $\beta = 2$	
Age t (years)	$R(t)$	next year, given t	$R(t)$	next year, given t
0	1.0000	0.8393	1.0000	0.9762
1	0.8393	0.8393	0.9762	0.9302
3	0.5912	0.8393	0.8050	0.8447
5	0.4164	0.8393	0.5473	0.7671

The table shows two behaviors. Under the exponential model the conditional reliability for the next year is $\mathrm{e}^{-8760/50\,000} = 0.8393$ at every age: from equation (1.6), $\mathrm{e}^{-(t+\Delta t)/\mathrm{MTTF}}/\mathrm{e}^{-t/\mathrm{MTTF}} = \mathrm{e}^{-\Delta t/\mathrm{MTTF}}$, independent of t. A surviving item is as good as new, a property called memorylessness. Under the wear-out model the conditional reliability falls with age, from 0.9762 for a new pump to 0.7671 for a five-year-old one. In both models the conditional value at five years is much larger than the unconditional probability of surviving six years from new (0.3495 and 0.4198), because the first five years are already known to have been survived. The answer to the plant manager's question therefore depends on which lifetime model describes the pump, and the evidence that selects the model is a subject of the statistical chapters that follow.

1.3 Component and system reliability

Systems are built of components, and the reliability of a system follows from the reliabilities of its components and from the way they are arranged. This section derives the two basic arrangements, series and parallel, for a cooling system of two pumps, and states the assumption on which both derivations rest.

1.3.1 Two pumps in series

Consider a cooling system of two pumps, A and B, arranged in series, so that both must function for the water to circulate. Panel (a) of Figure 1.2 draws it as a reliability block diagram: each block is a component, the line runs from the input on the left to the output on the right, and the system functions if a path of functioning blocks connects the two ends. In a series arrangement there is one path, through both blocks. Each pump has reliability $R_A = R_B = 0.95$ over the same period.

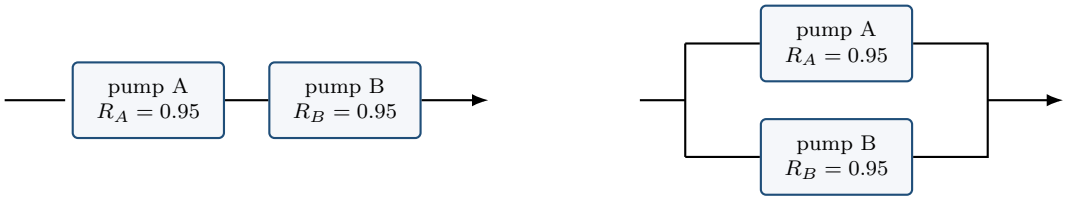
The system functions only if A functions and B functions. If the failures of A and B are independent, the probability of the intersection of the two events is the product of their probabilities,

$$R_{\text{system}} = \Pr(\text{A functions} \cap \text{B functions}) = R_A\, R_B, \tag{1.7}$$

and, for n components in series, $R_{\text{system}} = \prod_{i=1}^{n} R_i$. For the two pumps,

$$R_{\text{system}} = 0.95 \times 0.95 = 0.9025.$$

The system reliability is 90.25 %, although each pump alone has 95 %. The system fails if either pump fails, and its failure probability, $1 - 0.9025 = 0.0975$, is almost the sum of the two component failure probabilities, $0.05 + 0.05 = 0.10$; the difference, $0.05^2 = 0.0025$, is the probability that both fail, which the sum counts twice. A series system is less reliable than its least reliable component, and every component added in series lowers its reliability further.



(a) series: both pumps required
$R_{\text{system}} = 0.95^2 = 0.9025$

(b) parallel: either pump sufficient
$R_{\text{system}} = 1 - 0.05^2 = 0.9975$

Figure 1.2: Reliability block diagrams of a two-pump cooling system, each pump of reliability 0.95 over the same period, with independent failures. (a) Series, equation (1.7): one path through both blocks. (b) Parallel, equation (1.8): two paths, one through each block. The system functions if functioning blocks connect the left end to the right end.

1.3.2 Two pumps in parallel

Suppose now that either pump can provide sufficient cooling on its own, as in panel (b) of Figure 1.2. There are two paths, one through each pump, and the system fails only if both pumps fail. This arrangement is called redundancy: equipment is installed beyond what the function requires, so that the function survives the loss of some of it. The probability that a pump fails is $1 - R_i$; if the failures are independent, the probability that both fail is the product, and the system reliability is its complement,

$$R_{\text{system}} = 1 - \Pr(\text{A fails} \cap \text{B fails}) = 1 - (1 - R_A)(1 - R_B), \tag{1.8}$$

and, for n components in parallel, $R_{\text{system}} = 1 - \prod_{i=1}^{n}(1 - R_i)$. For the two pumps,

$$R_{\text{system}} = 1 - (0.05)(0.05) = 1 - 0.0025 = 0.9975.$$

The same two pumps give 90.25 % in series and 99.75 % in parallel. The failure probability falls from 0.0975 to 0.0025, a factor of 39, and relative to one pump alone from 0.05 to 0.0025, a factor of 20. In series, failure probabilities approximately add; in parallel, they multiply.

1.3.3 The independence assumption

Both results rest on one assumption: the failures of the two pumps are independent, so that the probability of a joint event is the product of the individual probabilities. The assumption fails whenever the pumps share something whose failure affects both: the same power supply, a common cooling-water source, or one controller. Then the failure of the shared element removes both pumps at once, and the multiplication in equation (1.8) no longer describes the system.

An illustrative calculation shows the size of the effect. Let both pumps be fed from one power supply of reliability 0.99 over the same period, and let the pumps otherwise fail independently as before. The supply is in series with the parallel pair, since the system needs the supply and at least one pump, and

$$R_{\text{system}} = 0.99 \times 0.9975 = 0.987525.$$

The system failure probability rises from 0.0025 to 0.012475, five times higher, and four fifths of it comes from the shared supply alone. A component shared by redundant paths sits in series with all of them and caps the reliability of the whole at its own. Failures of several items from one shared cause are called common-cause failures; later chapters on fault trees and on common-cause failure models treat them in detail.

Principle 1.2. *System reliability depends on architecture: series arrangements multiply reliabilities, parallel arrangements multiply failure probabilities, and both rules hold only when the component failures are independent.*

1.4 Failure probability, risk and availability

The probability of failure is one of several quantities used to judge a system, and it is often confused with two others. This section distinguishes failure probability from risk, which weighs outcomes by their consequences, and reliability from availability, which credits restoration after failure.

1.4.1 Failure probability and consequence

Consider two equipment failures in the same facility over one year. The first is the failure of pump A, with an annual probability of 10 %, whose consequence is the temporary loss of a minor process. The second is the failure of the main cooling system, with an annual probability of 1 %, whose consequence is a major data center outage. The first event is ten times more probable; the second may be far more consequential. Ranking the two by probability alone puts the pump first, and that ranking says nothing about which deserves attention, because it ignores what each failure costs.

The two events illustrate why the probability of failure is not the same as risk. Risk combines the probability of an outcome with its consequence, and a rare event with a large consequence can carry more risk than a frequent event with a small one.

1.4.2 Risk as expected loss

A simple quantitative measure of risk for one failure event F is

$$\mathcal{R} = \Pr(F)\, C(F), \tag{1.9}$$

where $\Pr(F)$ is the probability of the specified failure event over a stated period and $C(F)$ is its consequence, conditional on the event occurring, here in Brazilian reais (R$). When the period is one year, $\mathcal{R}$ is an expected annual loss in R$ per year. Table 1.4 applies equation (1.9) to the two events, with illustrative losses. Each row is one event; the columns give its annual probability, the loss if it occurs, and the product. For the pump, $\mathcal{R} = 0.10 \times \text{R\$}\,10\,000 = \text{R\$}\,1000$ per year; for the main cooling system, $\mathcal{R} = 0.01 \times \text{R\$}\,1\,000\,000 = \text{R\$}\,10\,000$ per year. The main cooling system has one tenth of the failure probability and ten times the expected annual loss.

Table 1.4: Failure probability and risk of two events, illustrative losses. The expected annual loss is the product of the annual probability and the loss, equation (1.9).

Event	Consequence	Probability per year	Loss if it occurs	Expected annual loss
Pump A fails	temporary loss of a minor process	0.10	R$ 10 000	R$ 1000
Main cooling fails	major data center outage	0.01	R$ 1 000 000	R$ 10 000

Figure 1.3 places the two events on a plane of probability against consequence. The horizontal axis is the loss if the event occurs, in reais, and the vertical axis the annual probability, both on logarithmic scales. The two dashed lines are lines of constant risk, $\Pr \cdot C = \text{R\$}\,1000$ and R$ 10 000 per year; on logarithmic axes they are straight lines of slope -1. The pump lies on the lower line and the main cooling system on the upper one: it sits lower on the probability axis and higher in risk. Every event on one line carries the same risk, so a hundredfold rarer event with a hundredfold larger loss is equivalent in expectation.

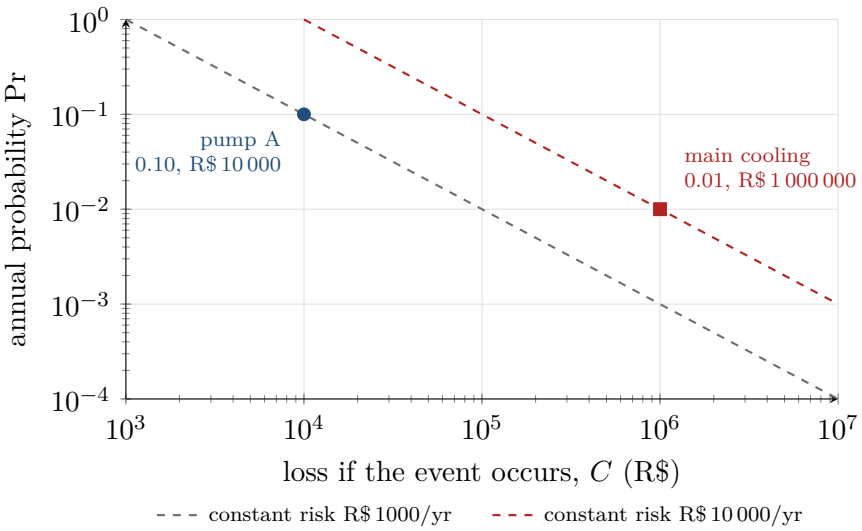


Figure 1.3: Annual probability against consequence for the two events of Table 1.4, logarithmic axes. Dashed lines are lines of constant risk $\Pr \cdot C$. The main cooling system is ten times less probable than the pump failure and carries ten times its risk. Values illustrative.

Several outcomes. When a failure can lead to several outcomes S_i with different consequences, such as a short outage, a long one and equipment damage, the risk is the probability-weighted sum

$$\mathcal{R} = \sum_i \Pr(S_i)\, C(S_i). \tag{1.10}$$

The expression is the expected value of the consequence, and it is correct when the outcomes form a partition, mutually exclusive and together exhaustive (the outcome "no failure" included, with consequence zero), and when the consequences are defined on one scale and over one period. If the outcomes overlapped, a scenario belonging to two of them would be counted twice. Two qualifications sharpen equation (1.9). First, for separate failure events whose losses add, linearity of expectation gives the expected total loss as the sum of the separate products even when the events can occur together: the facility of Table 1.4 has an expected loss of R$ 1000 + R$ 10 000 = R$ 11 000 per year from the two events. Second, an event that can occur more than once in a year contributes the expected number of its occurrences times the loss per occurrence; for a rare event the expected number and the probability of at least one occurrence are nearly equal, which is the case equation (1.9) represents. Risk analysis is therefore concerned not only with the likelihood of failure but with what happens after it, including the cases in which several outcomes are possible.

1.4.3 Reliability and availability

Consider two pumps. Pump P fails every few months but is repaired within a day; pump Q rarely fails, but each repair takes three weeks, the time needed to obtain a spare part. Which pump is more reliable and which is more available are two different questions, and they can have different answers. Reliability concerns successful operation without failure over a specified interval. Availability concerns whether the item is operational at a particular time, or the fraction of time it is operational, and it credits the restoration that follows each failure.

Figure 1.4 draws one illustrative year of operation of each pump. The horizontal axis is operating time over one year, 8760 h; each row is one pump; green bars are operating intervals and red bars repair intervals. Pump P fails four times and is down for 24 h after each failure, 96 h in all; its repair intervals are too short to show at this scale and are drawn at the narrowest

visible width. Pump Q fails once and is down for 504 h. Pump Q runs for long stretches without interruption, so it is the more reliable; pump P is operating for a larger fraction of the year, $8664/8760 = 0.989$ against $8256/8760 = 0.942$ for Q, so in this year it is the more available.

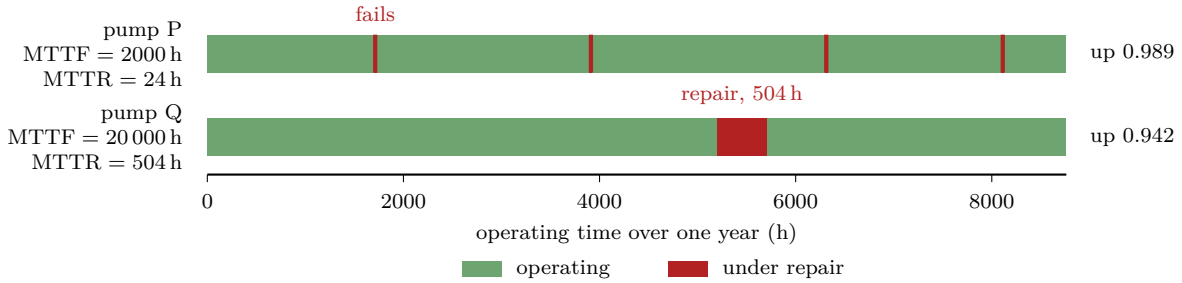


Figure 1.4: One illustrative year of operation of two repairable pumps. Pump P fails four times and is repaired in 24 h each time (red marks drawn at the narrowest visible width); pump Q fails once and is repaired in 504 h. Pump Q is more reliable over any interval; pump P is operating for a larger fraction of the time.

1.4.4 Steady-state availability

For a repairable item that alternates between operating intervals, which end in a failure, and repair intervals, which end in a restoration, the long-run fraction of time it operates follows from a short argument. Let U_k be the length of the k-th operating interval and D_k that of the k-th repair interval, and suppose the cycles are statistically alike and independent, with means $\mathbb{E}[U_k] = \mathrm{MTTF}$ and $\mathbb{E}[D_k] = \mathrm{MTTR}$. Over the first n cycles the fraction of time operating is

$$\frac{\sum_{k=1}^{n} U_k}{\sum_{k=1}^{n} U_k + \sum_{k=1}^{n} D_k} = \frac{\frac{1}{n}\sum_{k} U_k}{\frac{1}{n}\sum_{k} U_k + \frac{1}{n}\sum_{k} D_k},$$

and, by the law of large numbers, each average tends to its mean as n grows. The limit is the steady-state availability,

$$A = \frac{\mathrm{MTTF}}{\mathrm{MTTF} + \mathrm{MTTR}}, \tag{1.11}$$

a standard result for this model, which is called an alternating renewal process. Here MTTF is the mean operating duration before a failure in the alternating model, and MTTR is the mean time to repair or restore, as defined for the application (detection, travel, spare parts and testing included or not). Equation (1.11) is also the long-run probability that the item is operating at a randomly chosen time.

As an example, let $\mathrm{MTTF} = 1000\,\mathrm{h}$ and $\mathrm{MTTR} = 10\,\mathrm{h}$. Then

$$A = \frac{1000\,\mathrm{h}}{1000\,\mathrm{h} + 10\,\mathrm{h}} = \frac{1000}{1010} = 0.990099,$$

a long-run availability of about 99.01 %. The number does not mean that the probability of failure over the next 1000 hours is 0.99 %. With a constant failure rate, $R(t) = \mathrm{e}^{-t/\mathrm{MTTF}}$, and the probability of at least one failure in 1000 h is $1 - \mathrm{e}^{-1000/1000} = 0.632$: the item fails, on average, $8760/1010 = 8.67$ times a year, and it is available 99 % of the time only because each failure is repaired in ten hours.

Table 1.5 compares this item with the pumps P and Q of Figure 1.4. Each row is one item; the columns give its mean times, its steady-state availability from equation (1.11), its reliability over 1000 h with a constant failure rate, and the long-run number of failures and hours of downtime per year, $8760/(\mathrm{MTTF} + \mathrm{MTTR})$ and $(1 - A) \times 8760\,\mathrm{h}$. Pump P has the

higher availability, 0.9881 against 0.9754, and much the lower reliability, 0.607 against 0.951. The ranking of the two pumps reverses between the two measures, so a statement about one says nothing about the other. In the long run pump Q is down about 215 h a year, more than in the single year of Figure 1.4, which is one realization of a random process.

Table 1.5: Steady-state availability (equation (1.11)) and reliability over 1000 h with a constant failure rate, for three illustrative repairable items. Failures and downtime are long-run averages per year of 8760 h.

Item	MTTF (h)	MTTR (h)	A	$R(1000\,\text{h})$	Failures per year	Downtime (h/yr)
Example item	1000	10	0.9901	0.368	8.67	86.7
Pump P	2000	24	0.9881	0.607	4.33	103.9
Pump Q	20 000	504	0.9754	0.951	0.43	215.3

Reliability and availability therefore quantify different properties. Reliability is the right measure when any interruption is costly, as for a mission that cannot be resumed; availability is the right measure when the cost is set by the total time out of service. Chapter 3 treats the two together with maintainability and risk, and later chapters on repairable systems, Markov models and maintenance return to availability as a function of time.

Principle 1.3. *Reliability, availability and risk are different quantities: reliability is the probability of operating without failure over an interval, availability the probability of being in operation, restoration included, and risk weighs the probabilities of outcomes by their consequences.*

1.5 Reliability as a property of the physical system

The calculations so far took the failure probability of each component as a given number. This section asks where such numbers come from, shows with a transmission line that a failure can be the violation of a physical performance limit computed from a model, sets the lifetime and the performance-based formulations of reliability side by side, and lists the approaches by which failure probabilities are obtained.

1.5.1 The conductor heat balance

Consider one overhead conductor of an electric transmission line. The current it may carry is limited by its temperature: a hot conductor expands and sags toward the ground and, held hot for long enough, loses strength. The permissible current, called the ampacity or thermal rating of the line, therefore depends on the conductor's thermal behavior, which is affected by the conductor temperature, the ambient temperature, the wind speed and direction, the solar radiation, the properties of the conductor, and the electrical loading.

Figure 1.5 draws the heat balance of a length of conductor. The box is the conductor, which stores heat; the arrows entering it are heat gains, the arrows leaving it heat losses, and the labels on the outside name the external conditions that set each term. The balance is

$$mc_p \frac{\mathrm{d}T}{\mathrm{d}t} = I^2 R(T) + q_{\text{solar}} - q_{\text{conv}} - q_{\text{rad}}, \tag{1.12}$$

where $T(t)$ is the conductor temperature in °C, mc_p is the heat capacity of the modeled length in J/K, I is the current in amperes, $R(T)$ is the electrical resistance of the length in ohms,

which rises with temperature, q_{solar} is the solar heat gain, and q_{conv} and q_{rad} are the convective and radiative heat losses, all in watts. The first term on the right is the Joule heating; the convective loss grows with the wind speed and with the difference between the conductor and the ambient temperature. The temperature rises when the gains exceed the losses and settles where they balance.

Remark 1.1. In equation (1.12), $R(T)$ is the temperature-dependent electrical resistance of the conductor. It is not the reliability function, which is always written $R(t)$ and has time, not temperature, as its argument. Likewise $T(t)$ here is a temperature, not the time to failure T.

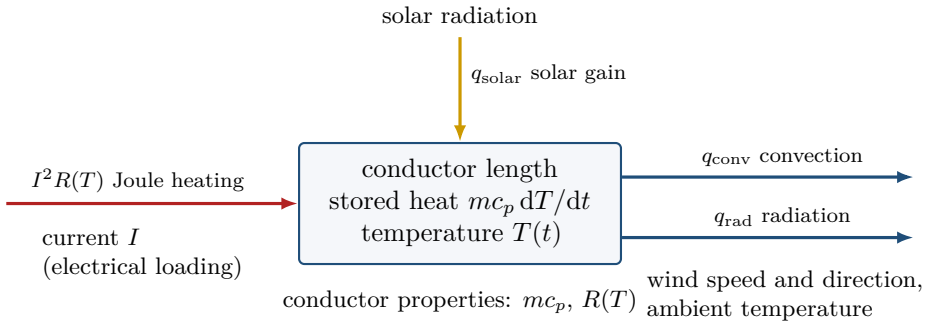


Figure 1.5: Heat balance of a length of overhead conductor, equation (1.12). Joule heating and solar gain enter; convection and radiation leave; the difference changes the stored heat and hence the conductor temperature. The labels outside the arrows name the conditions that set each term.

1.5.2 Failure as a performance event

Suppose the required operating criterion of the line is that the conductor temperature stays at or below a limit, $T(t) \le T_{\max}$, over an operating horizon H. A failure event is then

$$F = \left\{ \max_{0 \le t \le H} T(t) > T_{\max} \right\}, \tag{1.13}$$

the event that the temperature exceeds its limit at some time in the horizon, and its probability is

$$\Pr(F) = \Pr\left(\max_{0 \le t \le H} T(t) > T_{\max} \right). \tag{1.14}$$

The probability arises because the inputs of equation (1.12), the current, the wind, the ambient temperature and the sun, are uncertain over the horizon, so $T(t)$ is a random trajectory.

Figure 1.6 shows three such trajectories. The horizontal axis is the time of day over a horizon $H = 24\,\text{h}$; the vertical axis is the conductor temperature. The dashed red line is an illustrative limit $T_{\max} = 75°\text{C}$. Each curve is the temperature of the same conductor under a different realization of weather and load: all three rise in the afternoon, when load and sun peak. Realizations 1 and 2 stay below the limit; realization 3 crosses it between about 13.3 and 16.7 h (shaded), and in that realization the event F occurs. $\Pr(F)$ is the probability, over all realizations, of a trajectory like the third.

This is a different type of reliability problem from predicting when a mechanical bearing will fail. The conductor may remain structurally intact while its operating temperature exceeds the allowable threshold; the event represents a failure to meet an operating requirement, not necessarily permanent damage. When the temperature falls again, the line is again within its rating, and nothing has to be repaired.

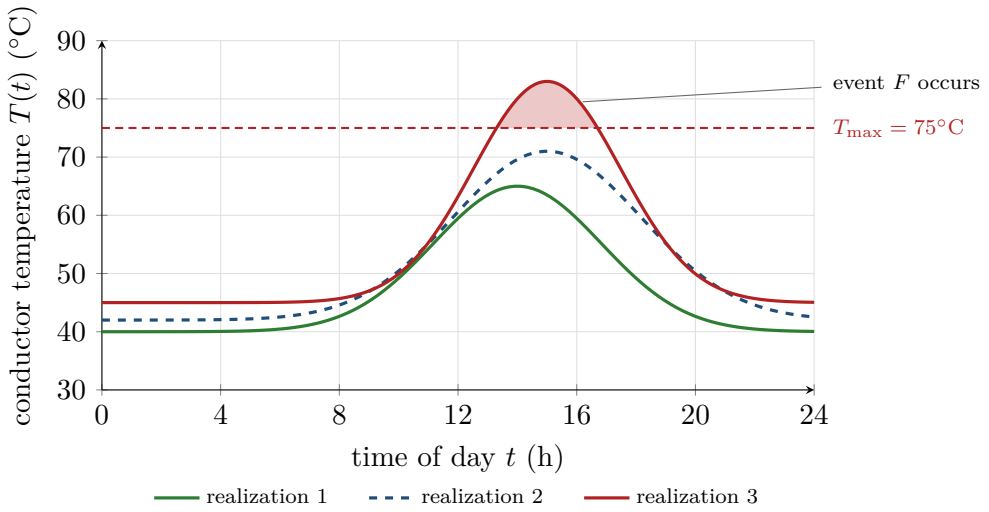


Figure 1.6: Three illustrative trajectories of the conductor temperature over a horizon $H = 24\,\text{h}$, for three realizations of weather and load, against an illustrative limit $T_{\max} = 75\,^{\circ}\text{C}$. Realization 3 exceeds the limit (shaded), so the failure event of equation (1.13) occurs in it. The curves are drawn to show the event, not computed from equation (1.12).

1.5.3 Lifetime and performance-based formulations

The two kinds of problem lead to two formulations of reliability. Table 1.6 sets them side by side. Each column is one formulation; the rows give its defining probability, the random object it is about, the failure event, an example, and the evidence it typically rests on. In the component lifetime formulation, reliability is the probability that a component survives a specified duration,

$$\Pr(T > H) = R(H). \tag{1.15}$$

In the performance-based formulation, reliability is the probability that a system continuously satisfies a specified performance requirement over the horizon,

$$\Pr\big(g(X(t)) \geq 0 \text{ for all } t \in [0, H]\big), \tag{1.16}$$

where $X(t)$ represents the uncertain system states or inputs at time t (load, weather, equipment condition) and g is a performance-margin function whose nonnegative values indicate acceptable performance. In structural reliability, g is called the limit-state function. For the conductor, $g = T_{\max} - T(t)$, with $T(t)$ obtained from equation (1.12) for the inputs $X(t)$, and equation (1.16) is $1 - \Pr(F)$ with F from equation (1.13).

Table 1.6: The component lifetime and the performance-based formulations of reliability.

	Component lifetime reliability	Performance-based reliability
Probability	$\Pr(T > H)$	$\Pr\big(g(X(t)) \geq 0,\ \forall t \in [0, H]\big)$
Random object	the time to failure T	the trajectory of states and inputs $X(t)$
Failure event	the component stops performing, once	the margin becomes negative at some time
Example	a bearing survives H	a conductor stays below $T_{\max}$ over H
Typical evidence	failure and survival times of similar items	a physical model and the distributions of its inputs

The lifetime formulation is a special case of the performance-based one. Let the state $X(t)$ be the condition of the component, and let $g(X(t)) = +1$ while the component works and -1 after it has failed. For a nonrepairable component the margin stays negative once it becomes negative, so it is nonnegative throughout $[0, H]$ exactly when $T > H$, and equation (1.16) reduces to equation (1.15). The two approaches can also be combined. A system may have components that fail at random times while its physical performance changes with weather, load and equipment condition; its margin then depends on both, and Section 1.6.4 writes such a margin for the cooling plant.

1.5.4 Sources of failure probabilities

Reliability engineering obtains failure probabilities by several complementary approaches. Table 1.7 lists seven. Each row is one approach; the columns give its main idea, an example, and the evidence it mainly rests on. Empirical reliability estimates failure behavior from historical observations. Statistical lifetime modeling fits probability distributions to failure times. Physics of failure models the degradation and failure mechanisms themselves. Logic-based system modeling combines component failure events into system failure events. State-based stochastic modeling represents transitions between operating and failed states. Simulation evaluates many possible realizations of the system, as in Figure 1.6. Bayesian inference updates the estimates as observations arrive.

Table 1.7: Approaches by which reliability engineering obtains failure probabilities.

Approach	Main idea	Example	Main evidence
Empirical reliability	estimate failure behavior from historical observations	pump failure records	counts of failures and operating hours
Statistical lifetime modeling	fit probability distributions to failure times	Weibull model for bearings	failure and survival times
Physics of failure	model the degradation and failure mechanisms	fatigue crack growth	material properties, load histories
Logic-based system modeling	combine component failure events	fault trees	component probabilities and system logic
State-based stochastic modeling	represent transitions between operating states	Markov availability model	failure and repair rates
Simulation	evaluate many possible realizations of the system	Monte Carlo analysis	models of components and of uncertain inputs
Bayesian inference	update estimates using observations	posterior distribution of a failure rate	a prior distribution and new observations

No approach is universally superior; the choice depends on the evidence available and on the structure of the system. With extensive historical data on thousands of identical components operated under similar conditions, empirical estimation may be sufficient. For a new design with little operational history, physical modeling and engineering assumptions play a larger role. For a complex system with multiple dependencies, component reliability estimates alone are unlikely to be sufficient, because the system's failure depends on how the components interact. Reliability engineering therefore requires both statistical and physical reasoning, and most studies combine several rows of Table 1.7.

Principle 1.4. *Physical models and statistical models complement each other: reliability may involve component lifetimes, physical performance thresholds, stochastic operating conditions and dependencies between components, and no single approach covers all four.*

1.6 Worked example: the data center cooling plant

The concepts of the chapter come together in the second situation of Section 1.1.1. This section computes the reliability of a cooling plant of three chillers of which two are required, shows how it depends on the reliability of one chiller, lists the assumptions behind the number, and writes the capacity formulation that removes some of them.

1.6.1 The two-out-of-three success criterion

A data center uses three chillers, and any two of them are sufficient to meet the cooling demand. Each chiller has reliability $R = 0.95$ over a one-year mission. The analysis assumes that each unit starts in working order, that no repair is credited during the mission, so that a failed chiller stays failed for the rest of the year, and that the failures are independent. The cooling system succeeds when at least two chillers survive. A system of n units that functions when at least k of them function is called a k-out-of-n system; this plant is 2-out-of-3.

Panel (a) of Figure 1.7 draws it as a reliability block diagram: three chiller blocks in parallel feed a voting element, marked 2/3, that passes the function on when at least two of its inputs are working. The success event splits into two mutually exclusive cases. In the first, all three chillers survive, with probability

$$\Pr(3 \text{ survive}) = R^3. \tag{1.17}$$

In the second, exactly two survive. There are three ways to choose the chiller that fails, and each way has probability $R^2(1 - R)$, so

$$\Pr(2 \text{ survive}) = 3R^2(1 - R). \tag{1.18}$$

The two cases cannot occur together, so their probabilities add,

$$R_{\text{system}} = R^3 + 3R^2(1 - R) = 3R^2 - 2R^3. \tag{1.19}$$

Substituting $R = 0.95$,

$$R_{\text{system}} = 0.95^3 + 3\,(0.95)^2(0.05) = 0.857375 + 0.135375 = 0.99275.$$

The cooling plant has a reliability of 99.275 % over the one-year mission under these assumptions. Its failure probability, $1 - 0.99275 = 0.00725$, is about one seventh of the failure probability of one chiller, 0.05. Equation (1.19) is the case $k = 2$, $n = 3$ of the binomial sum $R_{\text{system}} = \sum_{j=k}^{n} \binom{n}{j} R^j (1 - R)^{n-j}$, which a later chapter on reliability block diagrams and redundancy develops.

Panel (b) of Figure 1.7 draws R_{system} against R. The horizontal axis is the reliability of one chiller and the vertical axis the system reliability. The solid curve is the 2-out-of-3 plant, equation (1.19); the dotted line is one chiller alone, $R_{\text{system}} = R$; the dashed curve is a plant of two chillers with no spare, both required, a series system with $R_{\text{system}} = R^2$. The spare raises the plant from 0.9025 to 0.99275 at $R = 0.95$. The 2-out-of-3 curve crosses the single-unit line at $R = 0.5$, since $3R^2 - 2R^3 = R$ has the roots $R = 0$, 0.5 and 1: the arrangement improves on one unit only when each unit is more likely to survive than to fail.

1.6.2 Sensitivity to component reliability

Table 1.8 evaluates equation (1.19) for six values of the chiller reliability, from 0.60 to 0.99. Each row is one value of R; the columns give the two terms of the sum, the system reliability and the system failure probability, the approximation $3(1 - R)^2$ explained below, and the ratio of the failure probability of one chiller to that of the plant. The row $R = 0.95$ repeats the calculation of Section 1.6.1. For $R = 0.97$, as a second example,

$$R_{\text{system}} = 0.97^3 + 3\,(0.97)^2(0.03) = 0.912673 + 0.084681 = 0.997354,$$

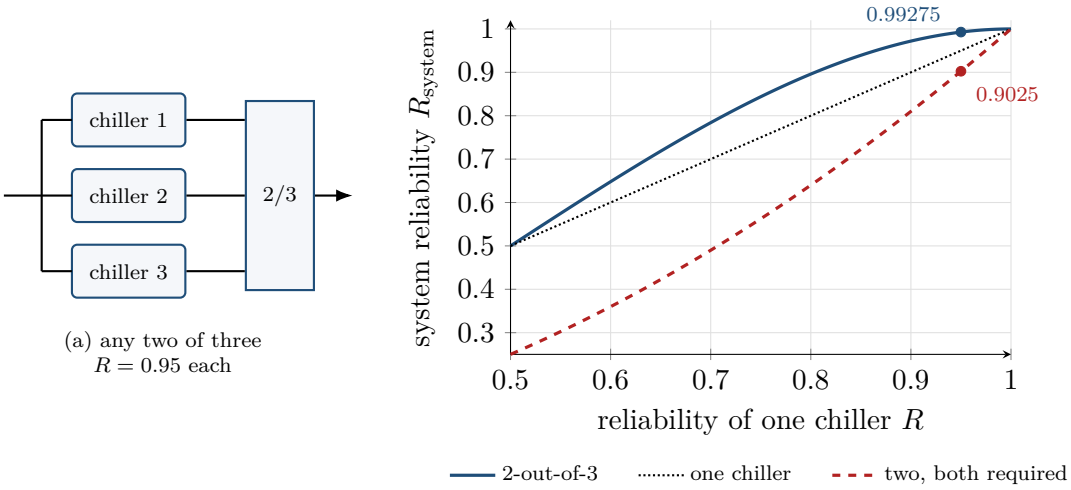


Figure 1.7: The 2-out-of-3 chiller plant with independent failures. (a) Reliability block diagram: three chillers in parallel feed a voting element that requires any two. (b) System reliability against the reliability of one chiller: $R_{\text{system}} = 3R^2 - 2R^3$ (equation (1.19)), compared with one chiller ($R_{\text{system}} = R$) and with two chillers both required ($R_{\text{system}} = R^2$). Marked: $R = 0.95$.

Table 1.8: Reliability of the 2-out-of-3 chiller plant over a one-year mission, equation (1.19), independent failures and no repair. The last column is the ratio of one chiller's failure probability to the plant's.

R	R^3	$3R^2(1-R)$	R_{system}	$1-R_{\text{system}}$	$3(1-R)^2$	Ratio
0.60	0.216000	0.432000	0.648000	0.352000	0.480000	1.1
0.80	0.512000	0.384000	0.896000	0.104000	0.120000	1.9
0.90	0.729000	0.243000	0.972000	0.028000	0.030000	3.6
0.95	0.857375	0.135375	0.992750	0.007250	0.007500	6.9
0.97	0.912673	0.084681	0.997354	0.002646	0.002700	11.3
0.99	0.970299	0.029403	0.999702	0.000298	0.000300	33.6

and the plant failure probability is $1 - 0.997354 = 0.002646$.

The table shows how strongly redundancy rewards good components. Writing $q = 1 - R$ for the failure probability of one chiller, the plant fails when at least two chillers fail,

$$1 - R_{\text{system}} = 3q^2(1 - q) + q^3 = 3q^2 - 2q^3 \approx 3q^2 \quad \text{for small } q,$$

so the plant's failure probability falls with the square of the chiller's. Halving q from 0.10 to 0.05 divides the plant failure probability by almost four, from 0.028 to 0.00725, and at $R = 0.99$ the plant fails 34 times less often than one chiller. The approximation $3q^2$ is within 10 % of the exact value for $R \geq 0.90$ and always overestimates it, because it counts the case of all three chillers failing three times instead of once, an excess of $2q^3$; at $R = 0.60$ it is useless. The same quadratic dependence makes the result sensitive to any effect that the model leaves out, which is the subject of the next two subsections.

1.6.3 Assumptions of the calculation

The number 0.99275 is the probability of a precisely stated model, and the model leaves out several features of a real plant. The calculation assumes that:

- each chiller is equally reliable;
- the failures are independent;
- two chillers always provide sufficient capacity;
- all three chillers start the mission in working order;
- no repair is credited during the mission;
- shared electrical, control and cooling-water dependencies create no additional failure paths.

Each assumption can fail, and each failure moves the number. A shared electrical feeder or a common control system is a common-cause failure path of the kind treated in Section 1.3.3, and it sits in series with the whole plant. A chiller that is out of service for maintenance at the start of the mission turns the plant into a series system of two, with reliability 0.9025 at $R = 0.95$ (Figure 1.7). Crediting repair works in the other direction: a failed chiller restored within days leaves only a short window in which a second failure causes a loss of cooling. The third assumption is the least visible one, and the next subsection removes it.

1.6.4 Capacity and load formulation

In a real data center the cooling demand changes with the IT load and with the environmental conditions, and the capacity of a chiller varies with its operating conditions: on a hot afternoon its condenser rejects heat to hotter outdoor air, and it delivers less. A more complete assessment represents the available cooling capacity at time t as

$$C_{\text{available}}(t) = \sum_{i=1}^{3} S_i(t)\, C_i(t), \tag{1.20}$$

where $S_i(t) \in \{0, 1\}$ indicates whether chiller i is available (1) or not (0), and $C_i(t)$ is the cooling capacity of chiller i under the current conditions, in kilowatts. The system fails to satisfy its cooling requirement if

$$C_{\text{available}}(t) < L(t) \tag{1.21}$$

at some time in the mission, where $L(t)$ is the required cooling load in kilowatts. In the terms of Section 1.5.3, the performance margin is $g = C_{\text{available}}(t) - L(t)$, and the uncertain state $X(t)$ contains both the random statuses $S_i(t)$ and the physical conditions that set $C_i(t)$ and $L(t)$. The formulation combines random equipment failures with physical operating conditions.

Figure 1.8 shows what the third assumption hides. The horizontal axis is time over two days; the vertical axis is cooling capacity and load in kilowatts. In this illustrative case each

chiller delivers $C_i(t) = 430 - 20\sin[2\pi(t - 8\,\text{h})/24\,\text{h}]\,\text{kW}$, between 410 and 450 kW, least in the afternoon, and the load is $L(t) = 725 + 125\sin[2\pi(t - 8\,\text{h})/24\,\text{h}]\,\text{kW}$, between 600 and 850 kW, most in the afternoon. The solid blue line is $C_{\text{available}}(t)$; the gray line is the load. Chiller 3 fails at $t = 20\,\text{h}$, so S_3 drops to 0 and $C_{\text{available}}$ falls from three units to two. With two chillers the plant still meets the load at night and in the morning, but on the afternoon of the second day, from about 35.7 to 40.3 h (shaded), the load exceeds the capacity of two derated chillers by up to $850 - 820 = 30\,\text{kW}$. By the counting rule of Section 1.6.1 the plant has not failed, since two chillers survive; by equation (1.21) it has. The same weather that raises the load lowers the capacity, a dependence that no count of surviving units represents.

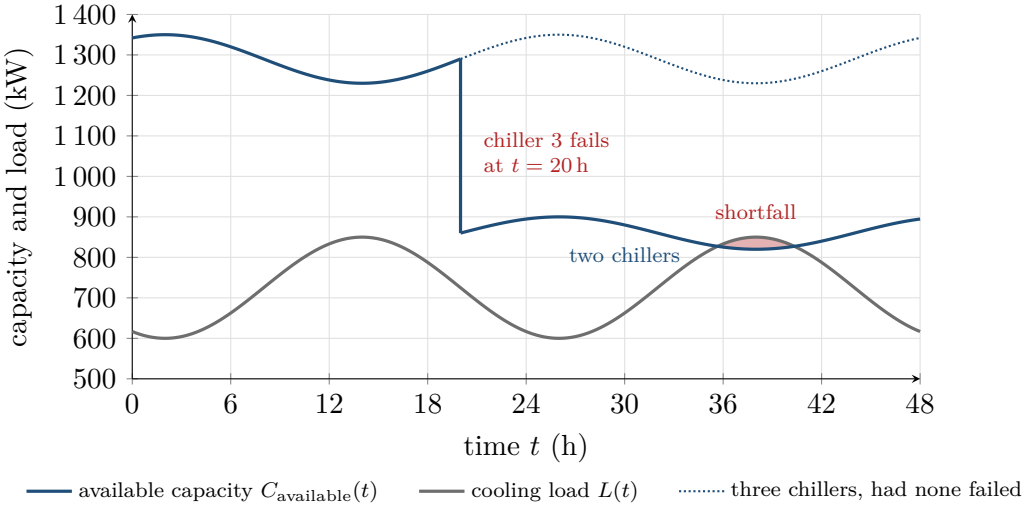


Figure 1.8: Available capacity $C_{\text{available}}(t) = \sum_i S_i(t)C_i(t)$ (equation (1.20)) and cooling load of an illustrative 2-out-of-3 plant over two days, with $C_i(t) = 430 - 20\sin[2\pi(t - 8)/24]\,\text{kW}$ and $L(t) = 725 + 125\sin[2\pi(t - 8)/24]\,\text{kW}$, t in hours. After chiller 3 fails at 20 h, two surviving chillers fall short of the afternoon load (shaded, up to 30 kW).

The example shows why understanding the system function must precede constructing its reliability model. The counting model of equation (1.19) answers the question "how many chillers survive"; the operator's question, whether sufficient cooling capacity is lost, is the capacity question of equation (1.21). The two coincide only when two chillers always suffice.

1.6.5 Common conceptual mistakes

The distinctions of this chapter are easy to lose in practice, and seven errors recur. Table 1.9 lists them. Each row is one mistaken belief; the middle column states why it matters, and the right column the subsection that treats the point.

The last row deserves emphasis. The value 0.99275 of Section 1.6.1 is exact to five digits for its model, and Figure 1.8 shows a plant for which the model's answer to the operator's question is wrong. A probability is as credible as the definition of failure, the system boundary and the assumptions behind it, and the precision of the arithmetic adds nothing to that credibility. Chapter 2 turns to the first of these: how the system being evaluated is defined, how its required functions and boundaries are stated, and how component failures are distinguished from the loss of a system function.

Table 1.9: Common conceptual mistakes about reliability and the correction of each.

Mistake	Why it matters	See
Reliability is an intrinsic property of equipment	it depends on the function, the conditions and the duration	1.2.1
A component must physically break to fail	failure can mean not meeting a performance requirement	1.2.2, 1.5.2
Redundancy always provides independent protection	shared dependencies may defeat redundancy	1.3.3
Reliability and availability are equivalent	availability includes repair and restoration	1.4.3
The most failure-prone component creates the greatest risk	system effects and consequences matter	1.4.2
Historical failure rates completely determine future reliability	conditions, degradation, uncertainty and model assumptions matter	1.2.5, 1.5.4
A precise probability implies a credible model	the model structure and the input evidence must also be credible	1.6.3

1.7 What to retain

Definition 1.1 (Reliability)**.** Reliability is the probability that an item performs its required function under specified conditions for a specified period, $R(t) = \Pr(T > t)$; it is defined only once the function, the conditions and the period are, it depends on the system's architecture, and it is distinct from availability, which credits repair, and from risk, which weighs outcomes by their consequences.

Table 1.10 lists the chapter's terms with their meaning, their units and the subsection that defines each.

Table 1.10: The terms of the chapter and where each is defined.

Term	Meaning	Units	Defined in
Item	component, subsystem or system treated as a unit of analysis	–	1.2.1
Required function	what the item must do, as a measurable criterion	–	1.2.1
Failure	termination of the ability to perform the required function	–	1.2.2
Reliability $R(t)$	$\Pr(T > t)$, survival beyond t	–	1.2.3
Cumulative failure probability $F(t)$	$\Pr(T \le t) = 1 - R(t)$	–	1.2.3
Conditional reliability	$R(t+\Delta t)/R(t)$, survival of a further Δt given survival to t	–	1.2.5
Series system	all components required; $R_{\text{system}} = \prod_i R_i$	–	1.3.1
Parallel system (redundancy)	any one component suffices; $R_{\text{system}} = 1 - \prod_i (1 - R_i)$	–	1.3.2
Common-cause failure	failure of several items from one shared cause	–	1.3.3
Risk $\mathcal{R}$	$\sum_i \Pr(S_i) C(S_i)$, expected loss over a partition of outcomes	R$/yr	1.4.2
Availability A	MTTF/(MTTF + MTTR), long-run fraction of time operating	–	1.4.4
MTTF, MTTR	mean operating duration before failure; mean time to repair or restore	h	1.4.4
Performance-margin function g	$g(X(t)) \ge 0$ means the requirement is met	varies	1.5.3
Performance-based failure event	$\{\max_{t \le H} T(t) > T_{\max}\}$, a limit exceeded in the horizon	–	1.5.2
k-out-of-n system	functions when at least k of n units function	–	1.6.1
Available capacity $C_{\text{available}}(t)$	$\sum_i S_i(t) C_i(t)$; shortfall if below $L(t)$	kW	1.6.4

Further reading. Rausand, Barros and Høyland, *System Reliability Theory*, 3rd ed., for the introductory material on reliability concepts, systems and failure definitions; O'Connor and Kleyner, *Practical Reliability Engineering*, 5th ed., for reliability terminology and engineering applications; and Modarres, Kaminskiy and Krivtsov, *Reliability Engineering and Risk Analysis*, 3rd ed., for the foundational distinctions between reliability analysis and risk assessment.

1.8 Exercises

Exercise 1.1. A pump operates continuously but delivers less flow than required. (a) Has it failed? (b) A specification requires at least $100\,\text{m}^3/\text{h}$. Three pumps of that specification, all running, are measured at 120, 100 and $85\,\text{m}^3/\text{h}$. Which of them have failed their required function? (Answers: (a) yes, if the required flow defines its function, since failure is defined relative to the required function and running alone does not imply functional success; (b) only the pump delivering $85\,\text{m}^3/\text{h}$.)

Exercise 1.2. Two independent components each have reliability 0.90 over a stated period. (a) Both must work: compute the system reliability. (b) Either suffices: compute it again. (c) A third identical component is added and any two of the three suffice: compute the reliability with equation (1.19). (d) In arrangement (b), both components draw power from one supply of reliability 0.98; compute the system reliability. (Answers: (a) $0.90 \times 0.90 = 0.81$, since a series system multiplies the reliabilities; (b) $1 - 0.1^2 = 0.99$; (c) $3(0.81) - 2(0.729) = 0.972$; (d) $0.98 \times 0.99 = 0.9702$.)

Exercise 1.3. A repaired pump is operational 99 % of the time. (a) Does that imply a reliability of 99 % over a year? (b) The pump has MTTF = 990 h and MTTR = 10 h. Verify its availability with equation (1.11), and compute, for a constant failure rate, its reliability over one year of continuous operation (8760 h) and its long-run number of failures per year. (Answers: (a) no, because availability includes restoration after failures, whereas mission reliability requires continuous success without failure; (b) $A = 990/1000 = 0.99$, $R(8760\,\mathrm{h}) = \mathrm{e}^{-8760/990} = 1.4 \times 10^{-4}$, and $8760/1000 = 8.76$ failures per year.)

Exercise 1.4. (a) Of two failure events, which necessarily has the greater risk: the one with the higher failure probability, the one with the more expensive consequence, or neither without evaluating both? (b) Event X has an annual probability of 0.2 and a loss of R$ 5000; event Y has an annual probability of 0.002 and a loss of R$ 2 000 000. Compute the expected annual loss of each with equation (1.9), and the ratio of the two. (Answers: (a) neither, because risk combines the probabilities and the consequences of the outcomes and neither alone is sufficient; (b) R$ 1000 and R$ 4000 per year: Y is a hundred times less probable and carries four times the risk.)

Chapter 2

System Functions, Boundaries, and Failure Definitions

2.1 Required function and performance criterion

Chapter 1 defined reliability as the probability that an item performs its required function under specified conditions for a specified period. The definition contains a difficulty that must be resolved before any probability is computed: it presupposes that success and failure have already been defined. Take a data center cooling plant with three chillers, of which any two carry the cooling demand. One chiller stops. Whether a failure has occurred depends on the item being assessed: the chiller has failed; the cooling system has not necessarily failed, since two chillers remain; and the data center may have experienced nothing of consequence. The three statements are all correct, because each refers to a different item and a different function. The question of this chapter is therefore

how the function a system must perform is stated precisely enough that success and failure can be decided, which items belong in the reliability model, and how the failure event is defined before its probability is estimated.

The chapter answers it in five steps. It first states the required function and converts it into a measurable performance criterion. It then fixes the system boundary and shows how failure at one level of a system hierarchy relates to failure at the next. It treats systems with several requirements and the role of the time horizon. Two worked examples follow, one in which the boundary changes and one in which the failure definition changes. The chapter closes with functional decomposition, the conditional character of every reliability result, and a general probabilistic formulation that contains the classical lifetime models as a special case.

Jargon: required function, performance criterion, limit-state function, functional failure, system boundary, mission time

The *required function* of an item is what it must do for the system to count as operating, stated as a measurable requirement (for example, "deliver at least $100\,\mathrm{m}^3/\mathrm{h}$"). A *performance criterion* is the measurable condition that separates acceptable from unacceptable performance of that function. A *limit-state function* (also called a performance function or performance-margin function) is a scalar function g of the system state, inputs and parameters, constructed so that $g \geq 0$ when the criterion is met and $g < 0$ when it is violated; the set $g = 0$ is the *limit state*. A *functional failure* is the loss of the ability to perform the required function, whether or not any component has physically broken. The *system boundary* is the set of items and interfaces included in the reliability model; everything outside it is an input or an assumption. The *mission time* H is the duration over which the required function must be maintained.

Key idea

Reliability is a probability of an event, and the event must be defined first: the required function, the performance criterion, the system boundary, the time horizon and the level of the hierarchy at which failure is assessed together determine the number, and two studies of the same equipment that fix them differently answer different questions.

2.1.1 Component status and system function

The three-chiller plant shows why a list of equipment is not a sufficient starting point. Its components are three chillers, the pumps and pipes that carry chilled water, the controls, and the electrical supply. Each component has an operational status at every moment: running, stopped, or out of service. The plant has a function: to remove the heat produced by the IT equipment. The status of one chiller and the success of the plant's function are different variables, and the reliability model has to say which of them its failure event refers to.

The distinction is the starting point of reliability engineering as a discipline. A reliability model begins with a functional requirement, and only then asks which equipment contributes to it. Starting from the equipment list leads to statements such as "the plant has failed because a chiller has failed", which may be false for the plant and irrelevant for the facility.

2.1.2 The required function

The three running cases of this book have the following required functions.

- **Industrial pumping system.** Deliver at least $100\,\mathrm{m}^3/\mathrm{h}$ of cooling water at a specified discharge pressure.
- **Data center cooling system.** Remove the facility's thermal load while maintaining critical equipment temperatures within limits.
- **Electric transmission network.** Deliver electrical power to specified delivery points while satisfying thermal, voltage, stability and other operating constraints.

None of the three says that the equipment must be running. A pump can run without delivering enough water: its impeller may be worn, a valve may be partly closed, or the system resistance may have risen. A chiller can operate without supplying enough cooling, because its condenser is fouled or the outdoor air is hot. A transmission line can remain energized while the network as a whole violates a thermal or voltage limit.

Two different questions are therefore asked of every item. The first is its *operational status*: whether the component is physically operating. The second is its *functional performance*:

whether the required engineering function is being fulfilled. The two are related, since a stopped pump delivers no flow, but they are not equivalent, since a running pump may deliver too little. A reliability model whose failure event is defined on operational status alone misses every functional failure of running equipment.

Common mistake

Equating "the equipment is running" with "the function is performed". A status signal reports that a motor is energized; it does not report the flow, the cooling or the delivered power. The failure event of a reliability model is defined on the function.

2.1.3 The limit-state function

A reliability model needs a measurable condition that separates acceptable from unacceptable performance. The common construction is a limit-state function g of three groups of variables: the system states x, such as temperatures, pressures and stored energy; the operating inputs or controls u, such as pump speeds, valve positions or the electrical dispatch; and the model parameters θ, such as efficiencies, resistances and heat-transfer coefficients. The performance criterion is

$$g(x, u, \theta) \geq 0, \tag{2.1}$$

with the sign convention

$$\begin{aligned} g > 0 &: \text{ acceptable performance,} \\ g = 0 &: \text{ performance boundary (the limit state),} \\ g < 0 &: \text{ requirement violated.} \end{aligned} \tag{2.2}$$

The value of g is the performance margin: the distance, in the units of the criterion, between the present performance and the boundary.

The sign convention is a modeling choice; the opposite convention, with failure at $g > 0$, appears in some texts. What matters is that one convention is used throughout a model, because every later operation, the combination of several criteria, the probability of violation and the sensitivity of the margin to a parameter, depends on it. This book uses $g \geq 0$ for success everywhere.

2.1.4 Margins of flow and temperature

Two examples from the running cases show the construction. The first criterion is a lower bound on a delivered quantity, the second an upper bound on a state.

Example 2.1 (Pump flow)**.** The pump must supply at least $Q_{\min} = 100\,\mathrm{m^3/h}$, where Q is the delivered flow in $\mathrm{m^3/h}$. The limit-state function is

$$g(Q) = Q - Q_{\min}.$$

At an actual flow of $Q = 120\,\mathrm{m^3/h}$, $g = 120 - 100 = 20\,\mathrm{m^3/h}$ and the flow requirement is met. At $Q = 85\,\mathrm{m^3/h}$, $g = 85 - 100 = -15\,\mathrm{m^3/h}$ and the requirement is violated, even if the pump is running.

Example 2.2 (Inlet temperature)**.** The inlet temperature T_{in} of the critical IT equipment must remain below $T_{\max} = 27°\mathrm{C}$. The limit-state function is

$$g(T_{\mathrm{in}}) = T_{\max} - T_{\mathrm{in}},$$

a temperature difference, expressed in kelvin. At $T_{\mathrm{in}} = 25°\mathrm{C}$, $g = 27 - 25 = 2\,\mathrm{K}$ and the requirement is met. At $T_{\mathrm{in}} = 30°\mathrm{C}$, $g = 27 - 30 = -3\,\mathrm{K}$ and it is violated.

Figure 2.1 draws both functions. In each panel the horizontal axis is the performance variable, flow on the left and inlet temperature on the right, and the vertical axis is the margin g in the units of the criterion. The straight line is the limit-state function; the horizontal line at $g = 0$ is the limit state, and the two dots are the operating points of the examples. Points above zero are acceptable and points below it violate the requirement. The two lines slope in opposite directions because one criterion is a lower bound and the other an upper bound, yet the success test is the same in both, $g \geq 0$. In neither example does the failure require a broken component: a flow of $85\,\mathrm{m}^3/\mathrm{h}$ from a running pump and an inlet at 30°C with every chiller running are both failures of the required function.

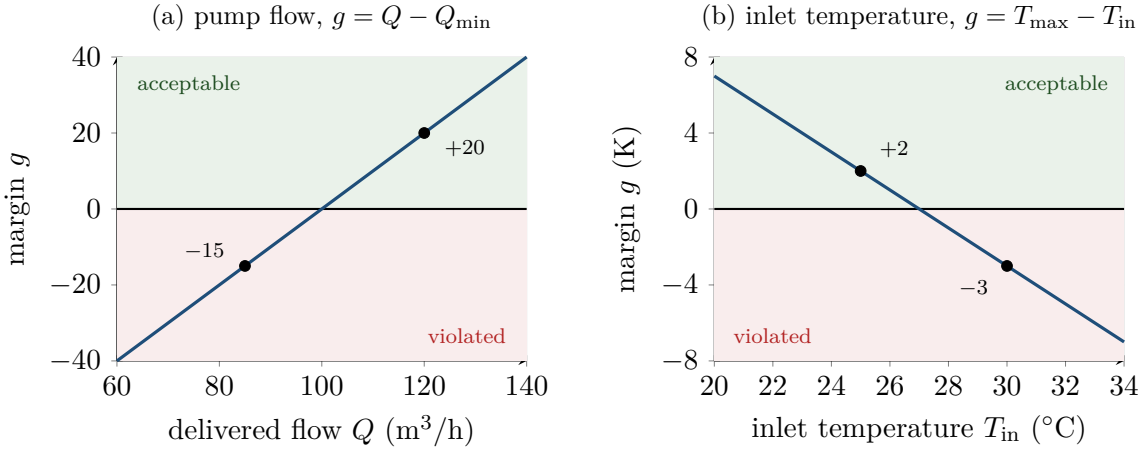


Figure 2.1: Limit-state functions of the two examples. (a) Pump flow with $Q_{\min} = 100\,\mathrm{m}^3/\mathrm{h}$, operating points $Q = 120$ and $85\,\mathrm{m}^3/\mathrm{h}$. (b) Inlet temperature with $T_{\max} = 27$°C, operating points 25 and 30°C. The shaded upper half is acceptable ($g > 0$), the lower half violates the requirement ($g < 0$).

Principle 2.1. *Failure is defined by the loss of performance margin, $g < 0$, not by the physical breakage of a component. A system with every component running can fail its required function, and a system with a failed component can still perform it.*

2.2 System boundaries and hierarchical failure

A performance criterion says what success is; the system boundary says which items the model must represent in order to compute it. This section defines the boundary, shows that it is chosen by the purpose of the study, separates a boundary from a dependency, and then follows failure through the levels of a hierarchical system.

2.2.1 The system boundary

A system boundary identifies the components, processes and dependencies included in an analysis. Inside the boundary, items are modeled: their states, failure modes and interactions are variables of the model. Outside it, items appear only as inputs to the model, with stated values or stated probabilities, or are assumed away.

Figure 2.2 draws an illustrative boundary for a data center cooling system. The dashed rectangle encloses the equipment being evaluated: the chillers, which generate the cooling; the pumps and piping, which carry it; and the controls and sensors, which operate both. Two items sit outside the boundary: grid power on the left, which supplies the equipment, and the IT

thermal load on the right, which the cooling serves. The lines are conceptual dependencies, not a hydraulic or electrical schematic. The figure makes the modeling decision visible: grid power and the IT thermal demand are external inputs, but their influence on the cooling service does not disappear by being drawn outside the line.

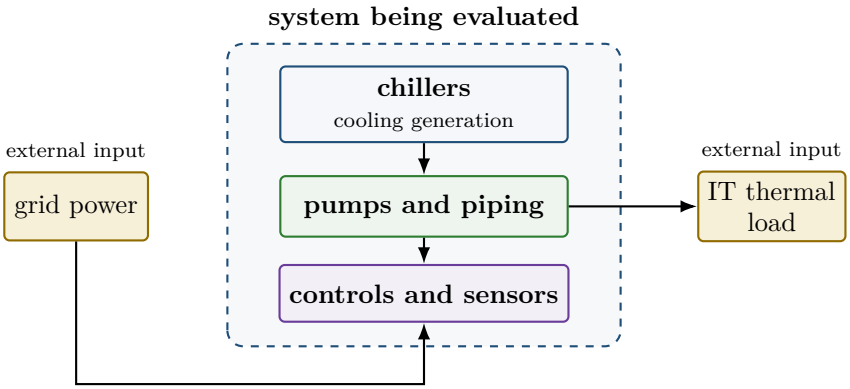


Figure 2.2: An illustrative boundary for a data center cooling system. The dashed boundary encloses the modeled equipment (chillers, pumps and piping, controls and sensors). Grid power and the IT thermal load remain outside as external inputs, but their influence cannot be ignored. The lines are conceptual dependencies, not a hydraulic schematic.

If external grid power is lost and the cooling system shuts down as a result, the figure does not say whether that event counts as a failure of the cooling system. No answer is correct in general; the answer is fixed by the purpose of the assessment, which is the subject of the next subsection.

2.2.2 The purpose of the study sets the boundary

Three studies of the same cooling plant draw the boundary in three places, and each is legitimate.

- **Equipment reliability study.** Electrical power is assumed available, and the study evaluates failures of the cooling equipment itself. Its question is how reliable the chillers, pumps and controls are.
- **Cooling-service reliability study.** Loss of external power is included as a possible cause of failure to deliver cooling. Its question is how reliably cooling reaches the IT equipment.
- **Whole-facility reliability study.** The power infrastructure, the cooling equipment, the IT load, the protective systems and their interactions are all inside the boundary. Its question is how reliably the facility delivers its service.

The three studies answer different questions, and their results are not interchangeable. An equipment study is the right tool for comparing two chiller models; a service study is the right tool for deciding whether a second electrical feeder is worth its cost; a facility study is the right tool for a statement to the facility's customers.

The choice of boundary is therefore part of the problem statement, not a detail of the calculation. A study report that does not state its boundary leaves the reader unable to tell which of the three questions it answers.

2.2.3 Boundary and dependency

A system boundary does not make an external dependency disappear; it converts it into an assumption. Let R_C be the probability that the cooling equipment succeeds over the mission

and R_E the probability that the electrical supply succeeds. Cooling service requires both. By the multiplication rule of probability,

$$\begin{aligned} R_{\text{service}} &= \Pr(\text{E succeeds and C succeeds}) \\ &= \Pr(\text{E succeeds})\ \Pr(\text{C succeeds} \mid \text{E succeeds}) = R_E\, R_{C|E}, \end{aligned} \tag{2.3}$$

which holds without any assumption. If the success events of the equipment and of the supply are independent, the conditional probability equals the unconditional one, $R_{C|E} = R_C$, and

$$R_{\text{service}} = R_C\, R_E. \tag{2.4}$$

A study that assumes power is always available reports R_C, or more precisely $R_{C|E}$. That number is a valid result conditional on the power-supply assumption; it is not the reliability of the delivered cooling service, which is smaller by the factor $R_E \leq 1$. Equation (2.3) also shows what the independence assumption hides: if the supply and the equipment share a cause, for example a voltage disturbance that both interrupts the supply and trips a chiller on restart, then $R_{C|E}$ differs from R_C and the product form is wrong.

Common mistake

Reporting a result computed inside a narrow boundary as the reliability of the whole service. A defensible report states every assumption that replaces an excluded dependency, such as "power is assumed available", and labels the result as conditional on it.

2.2.4 Hierarchical systems

Engineered systems are organized in hierarchies: a facility consists of subsystems, subsystems of components, and each component has its own failure modes. Figure 2.3 draws a simplified data center as such a hierarchy. The top box is the data center and its facility function. The second row holds two subsystems, cooling and electrical. The third row holds four components, chillers and pumps under cooling, the uninterruptible power supply (UPS) and the switchgear under electrical. The band at the bottom represents the individual component failure modes, grouped as mechanical, electrical, controls and sensors, to which every component is linked by a dashed line. The labels on the left name the four levels.

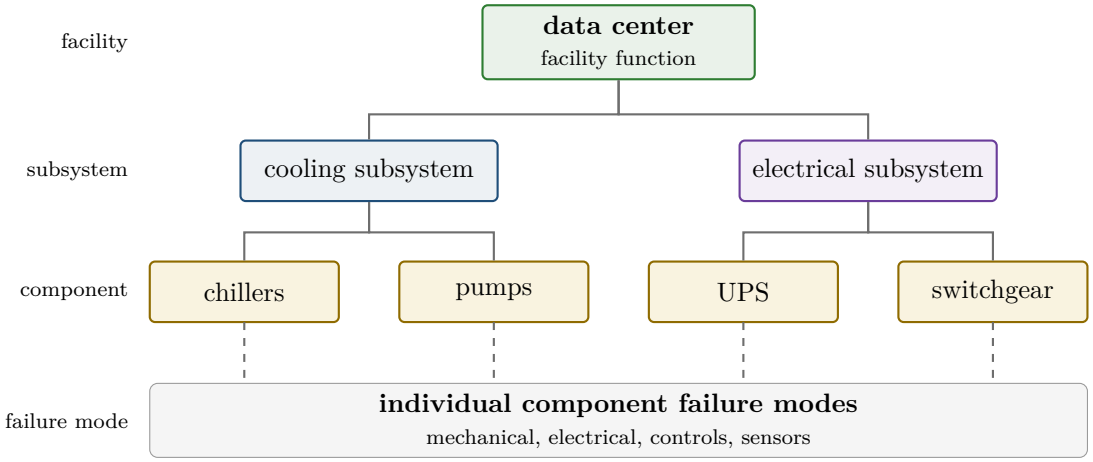


Figure 2.3: A simplified data center as a hierarchy: facility, subsystems, components and the failure modes of the components. Solid lines are composition; dashed lines link each component to its failure modes. A failure at one level need not propagate to the level above.

A failure at one level does not necessarily propagate to the next. Table 2.1 follows one event through the three upper levels of Figure 2.3. Each row is a level; the columns give the

event observed at that level and whether that level has failed. Chiller A stops: the component has failed. The two remaining chillers meet the demand: the cooling subsystem has not failed, because its capacity remains sufficient. The IT inlet temperatures remain acceptable: the facility has not failed, because its function continues.

Table 2.1: One chiller stop followed through three levels of the hierarchy of Figure 2.3, for a plant of three chillers of which two carry the demand.

Level	Event	Failed?
Component	Chiller A stops operating	Yes: chiller A
Subsystem	The two remaining chillers meet the demand	No: cooling capacity remains sufficient
Facility	IT inlet temperatures remain acceptable	No: the facility function continues

2.2.5 Component failure and system failure

The table describes three different reliability events. Let

$$F_A = \{\text{chiller A fails}\},$$
$$F_C = \{\text{the cooling system loses its required capacity}\},$$
$$F_D = \{\text{the data center loses its required service}\}.$$

The three events are not equivalent, and their relations are fixed by the system model, not by their names. In a model in which each chiller is either available at full capacity or unavailable, and two of the three carry the demand, F_C occurs exactly when at least two chillers are unavailable. Then F_A is neither sufficient for F_C, since chiller A can fail alone, nor necessary, since chillers B and C can fail together. A chiller failure can thus occur without either F_C or F_D.

Between the cooling system and the facility the relation is again a matter of the model. A cooling failure contributes to a data center outage, but it may not be sufficient for one if thermal storage, a backup cooling path or the thermal inertia of the room sustains the facility until the capacity is restored; and F_D can occur without F_C, through a failure of the electrical subsystem. Only in a model without storage or backup, in which cooling is required at every instant, does F_C imply F_D, that is, $F_C \subseteq F_D$. The logical and probabilistic relations between events at different levels are what fault trees construct, in a later chapter.

> **Principle 2.2.** *Failure events are defined at the level of the function being assessed. A component failure, a subsystem failure and a system failure are different events, and the relation between them is a property of the system model, to be derived rather than assumed.*

2.3 Multiple requirements and time horizons

Real systems have more than one requirement, and every requirement must hold over a stated period. This section combines several criteria into one system failure event, computes the probability of that event, and shows that the treatment of time in the criterion changes the verdict for one and the same history.

2.3.1 Several requirements at once

A transmission network must satisfy several requirements at every moment. With $I_\ell(t)$ the current through line ℓ, $I_{\ell,\max}(t)$ its allowable current, $V_i(t)$ the voltage magnitude at bus i, $V_{i,\min}$ and $V_{i,\max}$ its allowed range, and $P_{\text{unserved}}(t)$ the unserved demand, the requirements are

$$\begin{aligned} |I_\ell(t)| &\le I_{\ell,\max}(t) \quad \text{for every line } \ell, \\ V_{i,\min} \le V_i(t) &\le V_{i,\max} \quad \text{for every bus } i, \\ P_{\text{unserved}}(t) &= 0. \end{aligned} \tag{2.5}$$

The allowable current carries a time argument because, for an overhead line, it depends on the weather that cools the conductor; this is the weather-dependent line rating of the transmission running case. The acceptable operating region is the intersection of the sets allowed by the individual requirements: an operating state is acceptable only if it satisfies all of them.

Figure 2.4 draws that intersection for one line and one bus. The horizontal axis is the bus voltage in per unit (p.u.), that is, as a fraction of its nominal value; the vertical axis is the line current as a fraction of its allowable current. The dashed lines are the three limits, a voltage range of 0.95 to 1.05 p.u. and a current limit of 1, all illustrative. The shaded rectangle is the acceptable region, the intersection of the three half-planes. The three labeled points are operating states examined in the next subsection: P_1 lies inside, P_2 violates the lower voltage limit and P_3 the current limit. A state outside the region violates at least one requirement, and which one is read off from the side through which it has left.

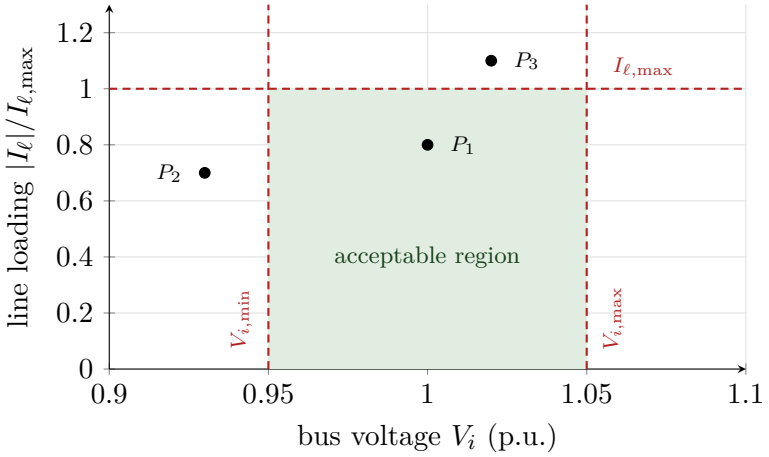


Figure 2.4: Acceptable operating region of one line and one bus as the intersection of three requirements: $V_{i,\min} = 0.95$ p.u., $V_{i,\max} = 1.05$ p.u. and $|I_\ell| \le I_{\ell,\max}$ (illustrative limits). P_1 is acceptable; P_2 violates the lower voltage limit; P_3 violates the current limit.

2.3.2 The combined margin

Each requirement has its own limit-state function. For one line and one bus,

$$g_1 = I_{\max} - |I|, \qquad g_2 = V - V_{\min}, \qquad g_3 = V_{\max} - V, \tag{2.6}$$

with the line and bus indices dropped; the unserved-demand requirement can be written in the same form as $g_4 = -P_{\text{unserved}}$, which is non-negative only when nothing is unserved. If all requirements are mandatory, the system succeeds when $g_1 \ge 0$, $g_2 \ge 0$ and $g_3 \ge 0$ hold together. One combined limit-state function is the smallest of the individual margins,

$$g_{\text{system}} = \min(g_1, g_2, g_3), \tag{2.7}$$

and $g_{\text{system}} \geq 0$ holds exactly when every requirement is satisfied, because the minimum of a set of numbers is non-negative if and only if each of them is. For a network the minimum runs over every line and every bus.

The margins of equation (2.6) have different units, amperes and volts, and different scales. The success test $g_{\text{system}} \geq 0$ is unaffected by this, since dividing each margin by a positive reference value does not change its sign. The *value* of g_{system} is affected: the minimum of an unnormalized current margin and voltage margin compares amperes with volts and has no meaning as a distance to failure. When the value is used, to rank operating states, to measure how close the system is to its limit, or to compute the sensitivity of the margin to a parameter, the margins must first be normalized. Dividing by the allowable current and by the nominal voltage gives, with V in per unit,

$$g_1 = 1 - |I|/I_{\max}, \qquad g_2 = V - 0.95, \qquad g_3 = 1.05 - V.$$

Table 2.2 evaluates them at the three points of Figure 2.4. At P_1 the smallest margin is the voltage margin, 0.05 p.u., and the state is acceptable. At P_2 the lower voltage margin is negative, -0.02, so $g_{\text{system}} < 0$ although the line is loaded to only 70 %. At P_3 the current margin is -0.10: the line is 10 % overloaded although both voltage margins are positive.

Table 2.2: Normalized margins at the three operating points of Figure 2.4 (illustrative limits 0.95 to 1.05 p.u. and $|I| \leq I_{\max}$). The binding requirement is the one with the smallest margin.

Point	V (p.u.)	$\lvert I\rvert/I_{\max}$	g_1	g_2	g_3	g_{system}	Verdict
P_1	1.00	0.80	0.20	0.05	0.05	0.05	acceptable
P_2	0.93	0.70	0.30	−0.02	0.12	−0.02	undervoltage
P_3	1.02	1.10	−0.10	0.07	0.03	−0.10	line overload

The combined margin settles whether every listed limit holds; it does not settle which violations count as *system failure*. Exceeding an operating limit, such as a short overload that a line tolerates thermally, may be classified differently from a loss of delivered electrical service, in which $P_{\text{unserved}} > 0$. Whether either constitutes system failure depends on the success criterion the study specifies, and that choice must be stated with the boundary.

2.3.3 Probability of the system failure event

Let F_1, F_2 and F_3 be the failure events of three mandatory requirements, $F_j = \{g_j < 0\}$. The system fails if any of them occurs, so the system failure event is their union,

$$F_{\text{system}} = F_1 \cup F_2 \cup F_3, \qquad \Pr(F_{\text{system}}) = \Pr(F_1 \cup F_2 \cup F_3). \tag{2.8}$$

This probability is not, in general, the sum of the individual failure probabilities. For two events, write $F_1 \cup F_2$ as the disjoint union of F_1 and of the part of F_2 outside F_1; the probability of that part is $\Pr(F_2) - \Pr(F_1 \cap F_2)$, so

$$\Pr(F_1 \cup F_2) = \Pr(F_1) + \Pr(F_2) - \Pr(F_1 \cap F_2). \tag{2.9}$$

Applying the same step twice gives, for three events,

$$\Pr(F_1 \cup F_2 \cup F_3) = \sum_j \Pr(F_j) - \sum_{j<k} \Pr(F_j \cap F_k) + \Pr(F_1 \cap F_2 \cap F_3),$$

the inclusion–exclusion formula. Without any knowledge of the intersections, the union is bounded by $\max_j \Pr(F_j) \leq \Pr(F_{\text{system}}) \leq \sum_j \Pr(F_j)$: the lower bound is reached when every

failure event is contained in the most probable one, the upper bound when the events cannot occur together.

The intersection terms matter because failures occur together. Extreme heat, for example, raises the demand on a cooling system and lowers the capacity of its equipment at the same time; in a transmission network, it raises the load and lowers the allowable current of the lines. That physical dependence must be represented in the probability model, and the direct way is to condition on the common driver. An illustrative model with two weather states shows the effect. Let a hot period occur with probability 0.1 and a normal one with probability 0.9. Let the current limit be violated (F_1) with probability 0.155 in hot weather and 0.005 in normal weather, and the voltage limit (F_2) with probability 0.21 and 0.01; within each weather state the two violations are taken as independent. Then

$$\Pr(F_1) = 0.1 \times 0.155 + 0.9 \times 0.005 = 0.02, \qquad \Pr(F_2) = 0.1 \times 0.21 + 0.9 \times 0.01 = 0.03,$$

$$\Pr(F_1 \cap F_2) = 0.1 \times 0.155 \times 0.21 + 0.9 \times 0.005 \times 0.01 = 0.003255 + 0.000045 = 0.0033,$$

five and a half times the value $0.02 \times 0.03 = 0.0006$ that an unconditional independence assumption would give. By equation (2.9), $\Pr(F_{\text{system}}) = 0.02 + 0.03 - 0.0033 = 0.0467$, against 0.0494 under independence and the upper bound 0.05. For rare events the union is close to the sum, so the error in the union is small here; the error in the intersection is not, and the intersection is the quantity that decides the reliability of redundant arrangements, in which the system fails only when several items fail together.

2.3.4 The time horizon

A reliability definition also needs a time horizon, and a statement of how the criterion is applied over it. Suppose a data center must keep a critical inlet temperature below $T_{\max} = 27°\text{C}$, and a forecast of the inlet temperature over the next ten hours is available. Figure 2.5 plots it. The horizontal axis is the time in hours, the vertical axis the inlet temperature in degrees Celsius. The dots are the hourly forecast values, joined by straight lines; the dashed red line is the limit. The temperature stays near 23°C for three hours, rises to a peak of 29°C at hour 5, and returns to about 23°C by hour 10. On the straight-line interpolation between hourly values it crosses the limit upward at $t = 4 + (27 - 26.5)/(29 - 26.5) = 4.2\,\text{h}$ and downward at $t = 6 + (28.2 - 27)/(28.2 - 26) = 6.545\,\text{h}$; the shaded area between the curve and the limit is the excursion.

The temperature exceeds the limit during part of the ten-hour period, and whether the system is classified as failed depends on the requirement. The requirement has a time structure of its own: it may ask that the criterion hold at one instant, at every instant of the horizon, or on some aggregate of the history. The next subsection writes three such definitions as events and applies them to the history of Figure 2.5.

2.3.5 Three failure definitions for one history

No excursion permitted. The strictest definition requires continuous success: the limit holds at every instant of $[0, H]$. Its failure event is

$$F_{\max} = \Big\{ \max_{t \in [0,H]} T_{\text{in}}(t) > T_{\max} \Big\}. \tag{2.10}$$

Under this definition any excursion, however short, is a failure.

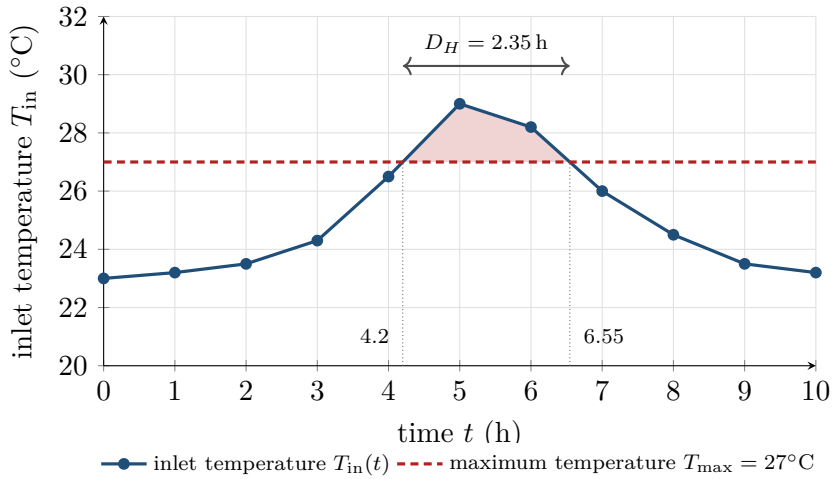


Figure 2.5: Illustrative forecast of a data center inlet temperature over $H = 10\,\mathrm{h}$ (hourly values 23, 23.2, 23.5, 24.3, 26.5, 29, 28.2, 26, 24.5, 23.5, 23.2°C), with the limit $T_{\max} = 27°\mathrm{C}$. A short excursion occurs around hour 5; with linear interpolation it lasts from 4.2 to 6.55 h, $D_H = 2.35\,\mathrm{h}$ (shaded).

Limited excursion duration permitted. Brief excursions may be acceptable provided their cumulative duration does not exceed a stated allowance $\tau_{\max}$. The total violation duration over the horizon is

$$D_H = \int_0^H \mathbf{1}\{T_{\mathrm{in}}(t) > T_{\max}\}\,\mathrm{d}t, \tag{2.11}$$

in hours, where the indicator function $\mathbf{1}\{\cdot\}$ equals 1 when its condition holds and 0 otherwise, so that the integral adds up the time spent above the limit. The failure event is $F_{\mathrm{dur}} = \{D_H > \tau_{\max}\}$. A short excursion can be acceptable under this definition and a failure under equation (2.10).

Thermal damage threshold. If equipment damage depends on accumulated exposure above a reference temperature T_{ref}, an illustrative damage index is

$$D_{\mathrm{thermal}} = \int_0^H \max\bigl(T_{\mathrm{in}}(t) - T_{\mathrm{ref}},\, 0\bigr)^n\,\mathrm{d}t, \tag{2.12}$$

in $\mathrm{K}^n\,\mathrm{h}$, with an exponent $n > 0$ and a critical value D_{crit} that must be calibrated empirically. The failure event is $F_{\mathrm{dmg}} = \{D_{\mathrm{thermal}} > D_{\mathrm{crit}}\}$. This answers a different engineering question from counting excursions, and a damage model of this form needs independent physical or empirical justification before it is used; the form here only illustrates the construction.

Application to the forecast. Table 2.3 applies the three definitions to the history of Figure 2.5, with $T_{\mathrm{ref}} = T_{\max} = 27°\mathrm{C}$ and illustrative thresholds. The rows are definitions, the columns the statistic each one computes from the history, its value, the threshold and the verdict. The statistics follow from the linear interpolation. The duration is $D_H = 6.545 - 4.2 = 2.345\,\mathrm{h}$. For the damage index the excursion is three linear pieces: the excess $T_{\mathrm{in}} - 27$ rises from 0 to 2 K over 0.8 h, falls from 2 to 1.2 K over 1 h, and falls from 1.2 to 0 over 0.545 h. The integral of a linear piece from a to b over a duration h is $h(a+b)/2$ for $n = 1$ and $h(a^2 + ab + b^2)/3$ for $n = 2$. For $n = 1$,

$$D_{\mathrm{thermal}} = \tfrac{0.8\times 2}{2} + \tfrac{1\times(2+1.2)}{2} + \tfrac{0.545\times 1.2}{2} = 0.80 + 1.60 + 0.33 = 2.73\,\mathrm{K\,h},$$

and for $n = 2$,

$$D_{\mathrm{thermal}} = \tfrac{0.8\times 4}{3} + \tfrac{4+2.4+1.44}{3} + \tfrac{0.545\times 1.44}{3} = 1.07 + 2.61 + 0.26 = 3.94\,\mathrm{K^2\,h}.$$

The same history is a failure under three of the five criteria and a success under the other two. The statistics themselves also depend on how the history is represented: counting only the hourly values above the limit, those of hours 5 and 6, would give $D_H = 2\,\mathrm{h}$ rather than $2.35\,\mathrm{h}$.

Table 2.3: Three failure definitions applied to the inlet-temperature history of Figure 2.5, with linear interpolation between hourly values and $T_{\mathrm{ref}} = T_{\max} = 27^{\circ}\mathrm{C}$. All thresholds are illustrative.

Definition	Statistic	Value	Threshold	Verdict
No excursion, eq. (2.10)	$\max_t T_{\mathrm{in}}(t)$	$29.0^{\circ}\mathrm{C}$	$T_{\max} = 27^{\circ}\mathrm{C}$	failure
Duration, eq. (2.11)	D_H	$2.35\,\mathrm{h}$	$\tau_{\max} = 1\,\mathrm{h}$	failure
Duration, eq. (2.11)	D_H	$2.35\,\mathrm{h}$	$\tau_{\max} = 3\,\mathrm{h}$	success
Damage, $n = 1$	D_{thermal}, eq. (2.12)	$2.73\,\mathrm{K\,h}$	$4\,\mathrm{K\,h}$	success
Damage, $n = 2$	D_{thermal}, eq. (2.12)	$3.94\,\mathrm{K^2\,h}$	$3\,\mathrm{K^2\,h}$	failure

Principle 2.3. *Different performance criteria define different failure events, and so produce different failure probabilities for the same physical system and the same temperature history. The time structure of the criterion, an instant, every instant, a duration or an accumulated exposure, is part of the definition of failure.*

2.4 Worked examples of boundaries and failure definitions

Two worked examples put the preceding ideas to use. The first keeps the equipment fixed and moves the system boundary; the second keeps the network fixed and changes the definition of failure.

2.4.1 Mission data and assumptions

A cooling service is evaluated over a specified mission of $H = 24\,\mathrm{h}$. The system consists of an electrical power supply, two redundant pumps, either of which is sufficient, and one heat exchanger, which is required. Table 2.4 lists the 24-hour reliability of each component, that is, the probability that the component performs its function throughout the mission. For this simplified example the success and failure events of the components are taken as independent, and no repair is possible during the mission.

Table 2.4: Mission reliabilities of the components of the cooling service, $H = 24\,\mathrm{h}$ (illustrative).

Component	Role	24-hour reliability
Electrical supply	required	0.990
Pump A	either pump sufficient	0.950
Pump B	either pump sufficient	0.950
Heat exchanger	required	0.980

Figure 2.6 draws the success logic as a reliability block diagram: the service succeeds if a path of working blocks connects the left terminal to the right one. The electrical supply and the heat exchanger are in series with everything else; the two pumps are in parallel with each other. The three nested dashed boxes are the boundaries of the three studies that follow, from the pump subsystem alone (innermost) to the complete cooling service (outermost); the reliability of each study is written in its label.

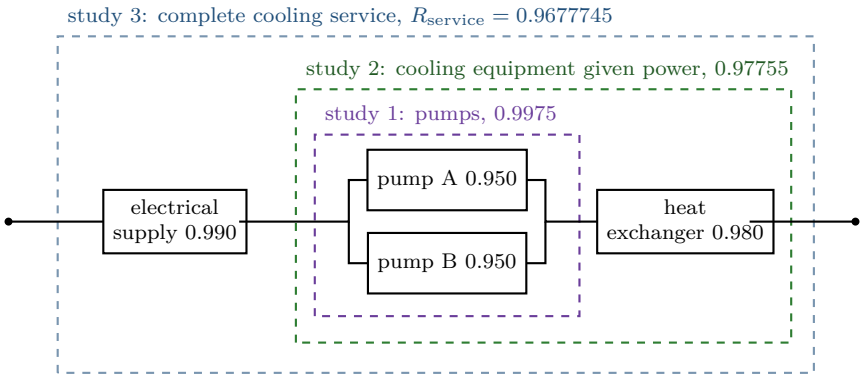


Figure 2.6: Reliability block diagram of the cooling service over a 24-hour mission, with the component reliabilities of Table 2.4. Nested dashed boxes mark the boundaries of the three studies: the pump subsystem, the cooling equipment given that power is available, and the complete cooling service including power.

2.4.2 Three boundaries, three answers

Study 1: pump subsystem only. The subsystem succeeds if at least one pump survives the mission, so it fails only if both pumps fail. With independent failures the probability that both fail is the product of the individual failure probabilities, and

$$R_{\text{pumps}} = 1 - (1 - R_A)(1 - R_B) = 1 - (0.05)^2 = 1 - 0.0025 = 0.9975.$$

Study 2: cooling equipment, given that power is available. Both the pump subsystem and the heat exchanger are required, so they are in series and their reliabilities multiply,

$$R_{\text{cooling|power}} = R_{\text{pumps}}\, R_{\text{HX}} = 0.9975 \times 0.98 = 0.97755.$$

Study 3: complete cooling service, including power. The electrical supply now enters the success logic in series,

$$R_{\text{service}} = R_{\text{power}}\, R_{\text{pumps}}\, R_{\text{HX}} = 0.99 \times 0.9975 \times 0.98 = 0.9677745.$$

Table 2.5 collects the three results. Each row is one study, with its boundary, its success logic, its reliability and its failure probability over the same 24-hour mission under the same independence and no-repair assumptions. The pump subsystem has a reliability of 99.75 %, the complete cooling service about 96.78 %. Its failure probability, 0.0322, is thirteen times that of the pump subsystem. Neither calculation is incorrect; they are answers to different system-level questions, and a report that quoted 99.75 % as the reliability of the cooling service would answer a question nobody asked.

2.4.3 Redundancy within the functional architecture

The pump subsystem has two pumps, but that alone does not make the cooling service redundant. The service requires the one heat exchanger, whose failure defeats it whatever the pumps do, and the one electrical supply, a dependency that both pumps share. The failure probability of the service can be apportioned approximately among its series elements: for small failure probabilities, $1 - \prod_i R_i \approx \sum_i (1 - R_i)$, which here is $0.01 + 0.0025 + 0.02 = 0.0325$ against the exact 0.0322. The heat exchanger contributes 0.02, about three-fifths of the total; the power supply 0.01, about three-tenths; and the redundant pumps 0.0025, under a tenth. The

Table 2.5: Reliability of the same 24-hour mission under three system boundaries, with independent component failures and no repair.

Study (boundary)	Success logic	Reliability	Failure probability
1. Pump subsystem	at least one of two pumps	0.9975	0.0025
2. Cooling equipment given power	pumps and heat exchanger	0.97755	0.02245
3. Complete cooling service	power, pumps and heat exchanger	0.9677745	0.0322255

redundant part of the system is its most reliable part, and the unreliability of the service sits in the elements that have no redundancy.

The conclusion holds generally. Redundancy at one level of a system protects the function only to the extent that every other element on the path is at least as reliable, and a shared element, here the electrical supply, sits in series with all the redundant units that depend on it. Whether redundancy is effective is a property of the complete functional architecture, not of the count of parallel units.

Principle 2.4. *Redundancy must be analyzed in the context of the complete functional architecture. A single required element or a shared dependency in series with a redundant group caps the reliability of the function, whatever the redundancy of the group.*

2.4.4 A two-path transmission network

Figure 2.7 draws a simple electrical network in which a load receives power from a source over two alternative transmission paths, line A and line B. The circles are the source and load buses; the boxes are the lines, each with its rating. Each line can carry 100 MW, and the load requires 80 MW, so either line alone can serve the load. The lines are taken as independent, each unavailable during the specified mission with probability 0.05, and no repair is modeled.

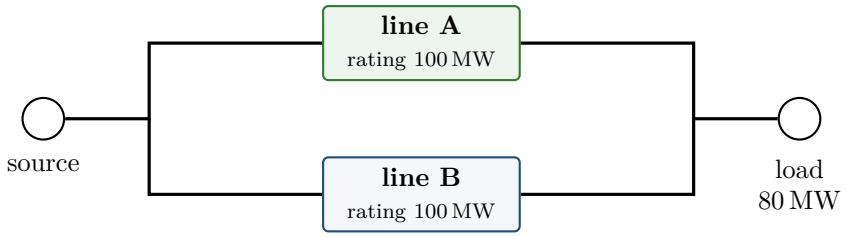


Figure 2.7: A load served over two alternative transmission paths. Each line is rated 100 MW and the load is 80 MW, so either line alone serves it (illustrative values).

Several definitions of failure are possible for this one network, and they produce different events. The first two are the following; to keep the events distinct from the chiller events of Section 2.2.5, the unavailability of line A is written F_{LA} and that of line B F_{LB}.

Definition 1: individual equipment failure. A line fails if it becomes unavailable, $F_{\mathrm{LA}} = \{\text{line A unavailable}\}$, with $\Pr(F_{\mathrm{LA}}) = 0.05$. This is a component-level event.

Definition 2: loss of the load-serving function. The network fails if neither line can deliver the required 80 MW. Since either line suffices, failure requires both lines to be unavailable, and

$$F_{\text{load}} = F_{\text{LA}} \cap F_{\text{LB}}.$$

This event identity follows from the success logic alone. Independence is needed only for the next step, the probability, which is then the product of the two unavailabilities,

$$\Pr(F_{\text{load}}) = 0.05^2 = 0.0025, \qquad R_{\text{load}} = 1 - 0.0025 = 0.9975.$$

The identity also depends on the absence of repair: with no repair, a line that becomes unavailable stays unavailable until the end of the mission, so both lines unavailable at some time in the mission is the same as both unavailable at its end. If repair were modeled, the two lines could be out at different, non-overlapping times, and the load would be lost only during an overlap; the event would then have to be defined on simultaneous unavailability.

2.4.5 Load-serving adequacy and N-1 security

A power system is commonly required not only to serve its load but to remain able to serve it after any single outage. This is the $N-1$ security criterion: a system with N elements must still meet its requirements after the loss of any one of them. Power-system reliability distinguishes *adequacy*, the existence of enough facilities to serve the load in the present state, from *security*, the ability to withstand sudden disturbances such as the outage of an element.

Definition 3: violation of the security requirement. If both lines are available, the network meets this simplified $N-1$ criterion, since either remaining line can carry 80 MW. If one line is already unavailable, the network still serves the load, but it no longer satisfies the criterion: the loss of the remaining line would interrupt service. With no repair, the criterion holds throughout the mission only if both lines survive it, so its failure event is $F_{\text{sec}} = F_{\text{LA}} \cup F_{\text{LB}}$ and

$$\Pr(F_{\text{sec}}) = 1 - (1 - 0.05)^2 = 1 - 0.9025 = 0.0975.$$

Table 2.6 sets the three definitions side by side. The rows are the definitions; the columns give the level of the event, the event, its probability over the mission and the corresponding reliability. The same network, the same lines and the same unavailabilities give a reliability of 0.95 for one line, 0.9975 for load delivery and 0.9025 for the security criterion. The probability of failing the security criterion is thirty-nine times that of losing the load. Load loss is contained in the security failure, $F_{\text{load}} \subseteq F_{\text{sec}}$, because a network that has lost both lines has also lost the ability to withstand one more outage.

Table 2.6: Three failure definitions for the network of Figure 2.7: independent lines, each unavailable with probability 0.05 over the mission, no repair. The probabilities of definitions 2 and 3 are computed from the same data.

Definition	Level	Failure event	Probability	Reliability
1. Line A unavailable	component	F_{LA}	0.05	0.95
2. Load not served	system, adequacy	$F_{\text{LA}} \cap F_{\text{LB}}$	0.0025	0.9975
3. $N-1$ criterion violated	system, security	$F_{\text{LA}} \cup F_{\text{LB}}$	0.0975	0.9025

Principle 2.5. *A system can satisfy its immediate load-serving requirement while lacking the margin or redundancy that its security criterion requires. Adequacy in the present state, security against contingencies, and the probability of service over a mission are different quantities, and a later chapter on power-system adequacy treats them separately.*

2.5 Functional decomposition and failure-mode analysis

The required function can be broken down into the functions that support it, and each of these into the ways it can fail. This section introduces that breakdown, which is the starting point of failure mode and effects analysis (FMEA), its criticality extension (FMECA) and fault tree analysis, and fixes the distinction between failure cause, failure mode and failure effect.

2.5.1 Functional decomposition

Functional decomposition starts from a high-level required function and progressively breaks it into the subfunctions that must be performed for it to be achieved; the result is often called a function tree. Figure 2.8 shows an illustrative decomposition for the data center cooling plant. The top box is the required function, maintaining IT cooling within the required temperature limits. The middle row holds three subfunctions (also called supporting functions): generate cooling, circulate the coolant, and distribute cool air. The bottom row holds, under each subfunction, the class of component fault that defeats it and the functional effect: chiller faults cause a loss of capacity, pump faults a reduction of flow, and fan and airflow faults a loss of distribution. The labels on the left name the three rows.

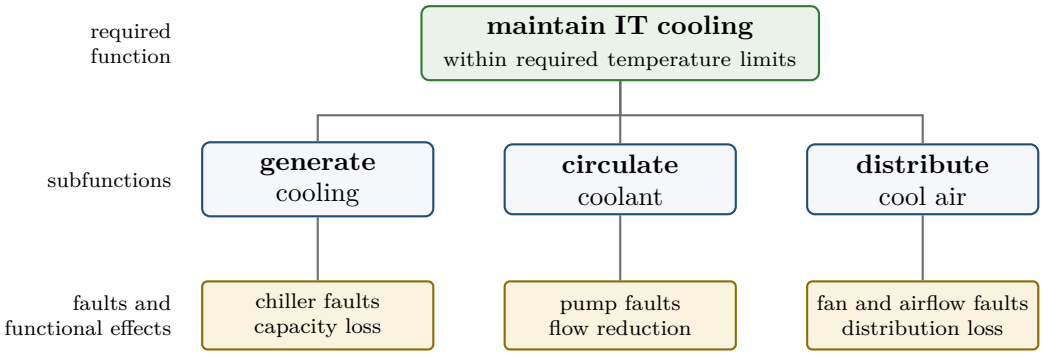


Figure 2.8: Illustrative functional decomposition of the data center cooling function into three subfunctions and the faults that defeat each. The diagram is not a complete fault tree, nor a proof that each subfunction is indispensable under all operating conditions.

The decomposition is a tool for organizing the analysis, not a proof of system logic. Each link in Figure 2.8 asserts that the subfunction is needed for the function above it, and that assertion may hold only under some operating conditions: at a low IT load and a cool outdoor temperature, for example, the room may stay within its limits for some time with no cooling generation at all. The logical relations between the events, with the conditions under which each holds, are what a fault tree states explicitly, in a later chapter.

2.5.2 From required function to failure cause

The analysis is organized in four distinct levels, which Table 2.7 lists with an example from the cooling plant; a fifth row adds the failure effect. The *required function* is what the system must achieve: maintain acceptable IT inlet temperatures. A *subfunction* is what must happen

to achieve it: circulate the coolant. A *failure mode* is the manner in which a subfunction becomes inadequate, as observed: insufficient coolant flow. A *failure cause* is the circumstance that produces the failure mode: pump seizure, a blocked strainer, a closed valve. The *failure effect* is what the failure mode does at the next level of the system: insufficient coolant flow reduces heat removal, raises the chilled-water temperatures and, if it persists, drives the IT inlet temperatures above their limits.

Table 2.7: The levels of a functional failure analysis, with examples from the data center cooling plant.

Level	Description	Example
Required function	What the system must achieve	Maintain acceptable IT inlet temperatures
Subfunction	What must happen to achieve it	Circulate the coolant
Failure mode	How a subfunction becomes inadequate	Insufficient coolant flow
Failure cause	Why the failure mode occurs	Pump seizure, blocked strainer, closed valve
Failure effect	What the failure mode does at the next level	Reduced heat removal, rising chilled-water temperature, IT inlet temperatures above their limits

The failure mechanism, the physical process such as bearing wear or corrosion that turns a cause into a failure, completes the vocabulary; it is the subject of Chapter 4. The four levels and the effect are kept separate because each answers a different question and each points to a different remedy: a cause is removed by maintenance or design, a mode is detected by monitoring, and an effect is limited by redundancy or protection.

2.5.3 Cause, mode and effect

The distinction among cause, mode and effect is the foundation of FMEA. At the level of the pump, a bearing failure is a possible cause, insufficient flow is a functional failure mode, and the loss of facility cooling is a possible system-level effect. The three descriptions refer to one chain of events seen at three positions along it.

The same event changes its role with the level of the analysis. In an FMEA of the pump itself, the bearing failure is the failure mode of the bearing, and its cause is, for example, contamination of the lubricant. In an FMEA of the cooling plant, insufficient flow is the failure mode of the circulation subfunction, and the bearing failure is one of its causes. In an FMEA of the facility, the loss of cooling is a failure mode of the cooling subsystem, and insufficient flow is one of its causes. The effect at one level is the failure mode at the next level up. The descriptions must therefore be kept consistent with the level at which the analysis is performed, and a worksheet that mixes levels in one column counts the same event twice or not at all.

2.6 Conditional models and the probabilistic formulation

Every number a reliability study reports rests on its definitions and assumptions. This section makes that dependence explicit in the notation, turns it into a checklist for reading a reliability result, and collects the ideas of the chapter into one probabilistic formulation, of which the classical lifetime models are a special case.

2.6.1 Reliability conditional on its assumptions

With the elements of the preceding sections, the reliability over a horizon H is

$$R(H) = \Pr(\text{the required function is satisfied over } [0, H]). \tag{2.13}$$

A more informative notation records what the number depends on,

$$R(H \mid \mathcal{C}, \mathcal{B}, \mathcal{M}), \tag{2.14}$$

where $\mathcal{C}$ denotes the operating conditions, including the environmental and loading assumptions; $\mathcal{B}$ the system boundary and the functional definition of success; and $\mathcal{M}$ the adopted reliability and physical model assumptions. The conditioning bar is not a probability calculation; it is a statement that the result is defined only once these three are fixed.

A cooling-system calculation might, for example, report $R(24\,\text{h}) = 0.98$. That probability might rest on the following assumptions: a maximum IT load of 800 kW; an ambient temperature below 35°C; at least one electrical feeder available; no maintenance during the 24-hour mission; and independent failures of the redundant pumps. A different set of assumptions, a higher load, a hotter day, power inside the boundary, a pump out for maintenance, or a common cause shared by the pumps, can produce a different result. Two analysts can therefore report different reliability numbers for the same facility without either making an arithmetic mistake. Their disagreement lies in the system boundary, the success criterion or the probability model, and it is resolved by comparing $\mathcal{C}$, $\mathcal{B}$ and $\mathcal{M}$, not by checking the arithmetic.

2.6.2 A checklist for reliability statements

The conditions above can be turned into a list of questions to put to any reliability result before accepting it. Table 2.8 lists seven of them. Each row is one condition the analysis should establish; the second column says what the condition fixes, and the third the subsection of this chapter that treats it.

Table 2.8: Checklist for a reliability model definition: what an analysis should establish before its result is accepted.

	Condition	What it fixes	Treated in
1	The required system function is explicitly defined	what success means	2.1.2
2	Acceptable performance limits are measurable	the limit-state function g	2.1.3, 2.1.4
3	The physical and functional system boundaries are identified	$\mathcal{B}$	2.2.1, 2.2.2
4	External dependencies are accounted for or explicitly conditioned on	what the result is conditional on	2.2.3
5	The mission duration and the initial operating condition are specified	H and the state at $t = 0$	2.3.4, 2.4.5
6	Failure events are consistently defined across component and system levels	the event whose probability is computed	2.2.5, 2.5.3
7	Important modeling assumptions and limitations are documented	$\mathcal{C}$ and $\mathcal{M}$	2.6.1

The checklist is not a complete model validation procedure. A model that passes all seven conditions can still be wrong in its structure or its parameters; the checklist only prevents the

fundamental formulation errors, in which a correct calculation answers a question different from the one posed. The initial operating condition in item 5 matters as much as the horizon: the $N-1$ example of Section 2.4.5 meets its criterion if it starts with both lines available and fails it if it starts with one.

2.6.3 The probabilistic formulation

The ideas of the chapter combine into one mathematical statement. Let $X(t)$ represent the uncertain physical state of the system at time t: temperatures, flows, voltages, stored energy. Let $S(t)$ be the vector of discrete equipment states, whose element $S_i(t)$ is 1 when unit i, a pump, a chiller or a transmission line, is available and 0 when it is not. Let Θ represent the uncertain model parameters, such as efficiencies, heat-transfer coefficients and line ratings. Suppose that m performance requirements are specified, each with a limit-state function g_j, $j = 1, \ldots, m$. The mission reliability of the system is

$$R(H) = \Pr\Big(g_j\big(X(t), S(t), \Theta\big) \geq 0 \ \text{ for all } j = 1, \ldots, m \text{ and all } t \in [0, H]\Big). \tag{2.15}$$

With a single combined margin $g = \min_j g_j$, as in equation (2.7), and the equipment states and parameters absorbed into the state, the same statement is written compactly as

$$R(H) = \Pr\big(g(X(t)) \geq 0 \ \forall t \in [0, H]\big). \tag{2.16}$$

The expression is deliberately general. It represents component lifetimes and random outages through $S(t)$; the physical behavior of the system through $X(t)$; uncertain equipment characteristics through Θ; several engineering requirements through the functions g_j; and reliability over a specified mission through H. The quantifier "for all t" makes it a statement of continuous success, the definition of equation (2.10); a duration or damage criterion replaces the quantifier by the corresponding functional of the history. The probability in equation (2.15) is never larger than the probability that the requirements hold at the single instant H, because the event that they hold at every instant is contained in the event that they hold at the last one. The two coincide only in special cases, such as the one in the next subsection.

2.6.4 Classical reliability as a special case

The formulation does not imply that every reliability problem requires a detailed physics-based model. For many systems a binary-state or lifetime model is sufficient, and equation (2.15) contains it. Take the three-chiller plant with each chiller of capacity C_i, a constant demand L, and two chillers able to carry it. The physical state is replaced by the available capacity, and the single limit-state function is

$$g(t) = C_{\text{available}}(t) - L = \sum_{i=1}^{3} S_i(t)\, C_i - L.$$

Without repair, each chiller is available until its time to failure T_i, so $S_i(t) = \mathbf{1}\{T_i > t\}$. Each status can only fall from 1 to 0, so $C_{\text{available}}(t)$ never increases, and $g(t) \geq 0$ for all $t \in [0, H]$ holds exactly when $g(H) \geq 0$. With two chillers sufficient, that is the event that at least two of the three lifetimes exceed H. For identical, independent chillers of mission reliability $R = \Pr(T_i > H)$, counting the outcomes with exactly two or exactly three survivors gives

$$R(H) = 3R^2(1 - R) + R^3 = 3R^2 - 2R^3,$$

the 2-out-of-3 reliability; for $R = 0.95$ it is 0.99275. The classical result is thus equation (2.15) with binary states, fixed capacity, fixed demand and no repair.

Each of those simplifications can be removed within the same formulation. A demand $L(t)$ that rises on a hot afternoon, a capacity $C_i(t)$ that falls with the outdoor temperature, or a repair that returns a chiller to service makes $g(t)$ non-monotone, and the event over $[0, H]$ then differs from the event at H. The general formulation is what relates classical reliability theory to physical performance analysis: the former is the latter with the physics held fixed. The next chapter, Chapter 3, separates reliability from the related quantities that a repairable system requires, availability, maintainability and risk.

2.7 What to retain

Definition 2.1 (Defining failure before estimating its probability)**.** The engineering function and the failure event are defined before any probability is estimated. Reliability,

$$R(H \mid \mathcal{C}, \mathcal{B}, \mathcal{M}) = \Pr(g_j \geq 0 \text{ for all } j \text{ and all } t \in [0, H]),$$

is a number that is meaningful only once the required function, its measurable criteria, the system boundary, the horizon and the level of the hierarchy are fixed.

Table 2.9 lists the terms of the chapter, with their meaning, their units and the subsection that defines each.

Table 2.9: Terms of the chapter and where each is defined.

Term	Meaning	Units	Defined in
Required function	what the system must accomplish; reliability is defined relative to it	–	2.1.2
Operational status	whether a component is physically operating; not the same as functional performance	–	2.1.2
Limit-state function g	performance margin, $g \geq 0$ acceptable, $g < 0$ requirement violated	units of the criterion	2.1.3
Performance criterion	measurable condition $g \geq 0$ that defines success	–	2.1.3
System boundary	items and interfaces inside the model; different boundaries ask different questions	–	2.2.1
Conditional reliability	result valid given an excluded dependency, $R_{C\|E}$	–	2.2.3
Functional hierarchy	facility, subsystem, component, failure mode; component failure need not be system failure	–	2.2.4, 2.2.5
Combined margin	$g_{\text{system}} = \min_j g_j$, normalized when its value is used	normalized	2.3.2
System failure event	union of the failure events of mandatory requirements, by inclusion–exclusion	–	2.3.3
Mission time H	horizon over which the function must hold	h	2.3.4
Violation duration D_H	time above the limit within $[0, H]$, eq. (2.11)	h	2.3.5
Damage index D_{thermal}	accumulated exposure above T_{ref}, eq. (2.12)	K^n h	2.3.5
Adequacy, $N-1$ security	serving the load now; withstanding the loss of any one element	–	2.4.5
Subfunction	function that must be performed for the required function to be achieved	–	2.5.1
Failure mode, cause, effect	how a function becomes inadequate; why; what it does at the next level	–	2.5.2, 2.5.3
$R(H \mid \mathcal{C}, \mathcal{B}, \mathcal{M})$	reliability conditional on conditions, boundary and model	–	2.6.1
Probabilistic formulation	$\Pr(g_j(X(t), S(t), \Theta) \geq 0 \ \forall j, t)$, eq. (2.15)	–	2.6.3

Further reading. Rausand, Barros and Høyland, *System Reliability Theory*, third edition, for system structure, components, reliability terminology and functional failure. O'Connor and Kleyner, *Practical Reliability Engineering*, for practical failure definitions, system requirements and reliability engineering. Modarres, Kaminskiy and Krivtsov, *Reliability Engineering and Risk Analysis*, for system reliability concepts and the formulation of risk.

2.8 Exercises

Exercise 2.1. Answer each part in one or two sentences, referring to the definitions of the chapter. (a) A redundant chiller fails, but the cooling demand is still met. What has failed? (The individual chiller, but not necessarily the cooling system: failure is evaluated at the specified system level.) (b) A model excludes external electrical power failure. What does its reliability estimate represent? (A result conditional on the stated power assumption; a narrower model is valid if its boundary and assumptions are reported.) (c) Two mandatory requirements have failure events F_1 and F_2. When does the system fail? (When either occurs: the system failure event is the union $F_1 \cup F_2$.) (d) Two reliability estimates for the same equipment differ. Can both be correct? (Yes, when the mission, the conditions, the initial state or the success

criteria differ, since reliability is defined relative to them.)

Exercise 2.2. A cooling service over a 24-hour mission consists of an electrical supply of reliability 0.98, two pumps of reliability 0.90 each, of which either is sufficient, and a single heat exchanger of reliability 0.97, all independent and without repair. (a) Compute the reliability of the pump subsystem. (b) Compute the reliability of the cooling equipment given that power is available. (c) Compute the reliability of the complete cooling service. (d) Apportion the failure probability of (c) approximately among the series elements and identify the largest contributor. (e) Compare two improvements: a second, independent electrical supply of reliability 0.98 in parallel with the first, or a second heat exchanger of reliability 0.97 in parallel with the first. Which raises the service reliability more, and by how much?

Exercise 2.3. An inlet temperature takes the hourly values 24, 26, 28, 28, 26 and 24°C at hours 0 to 5, with linear interpolation between them, against a limit $T_{\max} = 27$°C over $H = 5$ h. (a) Decide whether the system fails under the no-excursion definition, equation (2.10). (b) Compute the crossing times and the violation duration D_H of equation (2.11), and decide under the duration definition with $\tau_{\max} = 1.5$ h and with $\tau_{\max} = 2.5$ h. (c) Compute the damage index of equation (2.12) with $T_{\mathrm{ref}} = 27$°C for $n = 1$ and $n = 2$. (d) Compare each result with the corresponding value in Table 2.3 and explain why one history can rank above another under one definition and below it under another.

Exercise 2.4. The two lines of Figure 2.7 are replaced by three lines of 50 MW each, serving the same 80 MW load. Each line is unavailable during the mission with probability 0.05, independently, and there is no repair. (a) Enumerate the line states and compute the probability that the load is served throughout the mission. (b) Find the states in which the network meets the $N-1$ security criterion, and compute the probability that it meets the criterion throughout the mission. (c) Compare both results with Table 2.6, and state which design is better under each definition of failure. (d) Show that $F_{\mathrm{load}} \subseteq F_{\mathrm{sec}}$ for the three-line network, and use equation (2.9) to compute $\Pr(F_{\mathrm{load}} \cup F_{\mathrm{sec}})$.

Chapter 3

Reliability, Availability, Maintainability, and Risk

3.1 Four measures of dependability and the questions they answer

Chapter 2 established that reliability is defined only relative to a required function, a system boundary, a set of operating conditions and a mission duration. Once those are fixed, an item can still be judged by several different measures, and the engineering literature keeps four of them apart: whether the item continues to function without failure, whether it is operating when it is needed, how quickly it is restored after a failure, and what the failures cost. The question of this chapter is

how the probability of surviving without failure, the probability of being operational, the ability to restore a failed item, and the consequences of failure are distinguished from one another, and which of them a given engineering decision requires.

The four measures are closely related, and they are often reported together, but they answer different questions. A system can have low reliability and high availability, or high reliability and unacceptable consequences when it does fail. The chapter defines each measure in turn with one running comparison of two cooling pumps, then derives the relations among them for the simplest repairable model, and ends with their use in electric power systems and in the broader RAMS framework.

Jargon: reliability, mission reliability, availability, maintainability, risk, repairable item, MTTF, MTBF, MTTR, RAMS

Reliability is the probability that an item performs its required function, under specified conditions, without failure over a specified period; over a stated mission of length H without credit for repair it is the *mission reliability* $R(H)$. A *repairable item* is one that is restored to operation after a failure. *Availability* is the probability that an item is in an operating state at a given time, $A(t)$, or the long-run fraction of time it operates, A, restoration included. *Maintainability* is the probability that a failed item is restored within a specified time under specified procedures and resources. *Risk* is the combination of the probabilities of undesirable outcomes with their consequences. *MTTF* (mean time to failure) is the mean operating time until a failure; *MTBF* (mean time between failures) is the mean time from one failure to the next; *MTTR* (mean time to repair, or to restore) is the mean duration of a restoration. All three are in hours. *RAMS* stands for reliability, availability, maintainability and safety, considered together.

Key idea

Reliability, availability, maintainability and risk are related but not substitutable. Each answers a different probability question, and two items that are equal on one of them can differ by a factor of ten on another.

This section states the comparison that motivates the distinctions and names the four questions the rest of the chapter makes precise.

3.1.1 Two competing cooling pumps

A data center operator selects between two industrial cooling pumps that deliver the same flow. Each is a repairable item: when it fails it is restored and returned to service. Historical operating records, which are the evidence available for the choice, give the figures of Table 3.1. Pump A operates on average 500 h between failures and is restored in 1 h; Pump B operates on average 5000 h between failures and is restored in 10 h. Pump A therefore fails about ten times as often as Pump B, and each of its failures is repaired ten times faster. The numbers are illustrative.

Table 3.1: Operating records of the two cooling pumps (illustrative). The mean operating time between failures is the mean up time, written MTTF in this book; the mean restoration time is written MTTR.

Parameter	Pump A	Pump B
Mean operating time between failures, MTTF	500 h	5000 h
Mean restoration time, MTTR	1 h	10 h
Failure frequency during operation	higher	lower

Which pump is better depends on what matters to the facility. If every failure is disruptive, even when the repair is quick, Pump B is preferable, because it fails less often. If the priority is to minimize the total time without pumping, the two pumps perform almost identically, as Section 3.3.2 shows. If a single failure can interrupt critical cooling long enough to damage equipment, the consequence of each individual failure must be weighed as well, which is the subject of Section 3.5. The same records therefore support three different rankings, and each ranking corresponds to one of the measures defined in this chapter.

3.1.2 Four questions

The four measures answer four distinct engineering questions about the same item:

1. *Reliability*: will the item continue to function without failure over the mission?
2. *Availability*: is the item operational at the moment it is needed, or for what fraction of the time is it operational?
3. *Maintainability*: how quickly can a failed item be restored?
4. *Risk*: which undesirable outcomes can occur, how probable is each, and how severe are their consequences?

Each question has its own random variable. Reliability concerns the time to the first failure, T. Availability concerns the state of the item at a time t, which alternates between operating and failed. Maintainability concerns the duration of a restoration. Risk concerns a consequence, measured in money, energy not served, outage hours or another defined unit. Keeping the four random variables apart is what keeps the four measures apart.

The distinction between the physical system, the model and the evidence applies to each measure. The pumps fail and are repaired; the model represents their up and down durations by

probability distributions; the operating records are a finite sample from which the parameters of those distributions are estimated. The numbers computed in this chapter are properties of the model, conditional on the records.

3.2 Reliability as survival without failure

This section defines the reliability function and the cumulative failure probability, derives the exponential model from a constant hazard, and compares the mission reliability of the two pumps.

3.2.1 The reliability function

Let T be the random time to the first failure of an item that starts in a specified operating condition at $t = 0$. Its reliability function, or survivor function, is the probability that the first failure occurs after t,

$$R(t) = \Pr(T > t), \tag{3.1}$$

and its cumulative failure probability is the complement,

$$F(t) = \Pr(T \le t) = 1 - R(t). \tag{3.2}$$

Both are dimensionless probabilities and functions of the duration t in operating hours. $R(0) = 1$, R does not increase with t, and $R(t) \to 0$ as $t \to \infty$ for any item that eventually fails. Reliability defined this way is a statement about the *first* failure: an item that fails at hour 100 and is repaired at hour 101 has not survived a 200 h mission, whatever its state at hour 200.

The failure density is $f(t) = \mathrm{d}F/\mathrm{d}t = -\mathrm{d}R/\mathrm{d}t$, and the hazard function is the failure density conditional on survival to t,

$$\lambda(t) = \frac{f(t)}{R(t)} = -\frac{1}{R(t)}\frac{\mathrm{d}R}{\mathrm{d}t}, \tag{3.3}$$

in units of h^{-1}. For a small interval Δt, $\lambda(t)\,\Delta t$ is approximately the probability that an item which has survived to t fails in $(t, t + \Delta t]$.

3.2.2 The constant-hazard model and the MTTF

The simplest model assumes a hazard that does not depend on age, $\lambda(t) = \lambda$ for all t. Equation (3.3) then reads $\mathrm{d}R/\mathrm{d}t = -\lambda R$, and with $R(0) = 1$ its solution is

$$R(t) = \mathrm{e}^{-\lambda t}. \tag{3.4}$$

This is the exponential model, the standard textbook model of a constant failure rate. Its defining property is that it has no memory: the probability of surviving a further s hours, given survival to t, is

$$\Pr(T > t + s \mid T > t) = \frac{R(t+s)}{R(t)} = \frac{\mathrm{e}^{-\lambda(t+s)}}{\mathrm{e}^{-\lambda t}} = \mathrm{e}^{-\lambda s},$$

the same as for a new item. The instantaneous propensity to fail does not depend on how long the item has already operated. The assumption is not appropriate for every component, in particular where aging, wear or degradation matter; Chapter 5 develops hazards that increase or decrease with age.

The mean time to failure is the expected value of T. Integrating by parts, with $f = -\mathrm{d}R/\mathrm{d}t$ and $t\,R(t) \to 0$ as $t \to \infty$,

$$\mathrm{MTTF} = \mathbb{E}[T] = \int_0^\infty t\, f(t)\,\mathrm{d}t = \Big[-t\,R(t)\Big]_0^\infty + \int_0^\infty R(t)\,\mathrm{d}t = \int_0^\infty R(t)\,\mathrm{d}t, \tag{3.5}$$

and for the exponential model MTTF $= \int_0^\infty \mathrm{e}^{-\lambda t}\,\mathrm{d}t = 1/\lambda$. The failure rate and the MTTF are then reciprocal, and either one fixes the whole reliability function. Both are estimated from operating records: under the exponential model, the estimate of λ is the number of failures observed divided by the total operating time accumulated, and the estimate of the MTTF is its reciprocal, the total operating time per failure. Estimation with incomplete records is the subject of a later chapter on estimation and censoring.

3.2.3 Mission reliability of the two pumps

Treating the mean operating times of Table 3.1 as MTTFs of exponential models gives the failure rates

$$\lambda_A = \frac{1}{500\,\mathrm{h}} = 0.002\,\mathrm{h}^{-1}, \qquad \lambda_B = \frac{1}{5000\,\mathrm{h}} = 0.0002\,\mathrm{h}^{-1}.$$

The probability of operating 100 h without failure is, by equation (3.4),

$$R_A(100\,\mathrm{h}) = \mathrm{e}^{-0.002\times 100} = \mathrm{e}^{-0.2} = 0.8187, \qquad R_B(100\,\mathrm{h}) = \mathrm{e}^{-0.0002\times 100} = \mathrm{e}^{-0.02} = 0.9802.$$

Pump A fails during a 100 h mission with probability $1 - 0.8187 = 0.1813$, Pump B with probability $1 - 0.9802 = 0.0198$. For a 100 h mission Pump B is considerably more reliable: its probability of failure is about nine times smaller.

Figure 3.1 extends the comparison to longer missions. The horizontal axis is the mission duration in operating hours, from 0 to 2000 h; the vertical axis is the mission reliability. The solid red curve is Pump A and the dashed blue curve is Pump B; the dots mark the values at 0, 100, 250, 500, 750, 1000, 1500 and 2000 h, and the dotted vertical line marks the 100 h mission just computed. The gap between the curves grows with the mission. At 500 h, one MTTF of Pump A, its reliability has fallen to $\mathrm{e}^{-1} = 0.3679$ against 0.9048 for Pump B; at 1000 h it is 0.1353 against 0.8187; at 2000 h it is 0.0183 against 0.6703. Pump A almost certainly fails at least once in 2000 h, while Pump B survives that mission two times out of three.

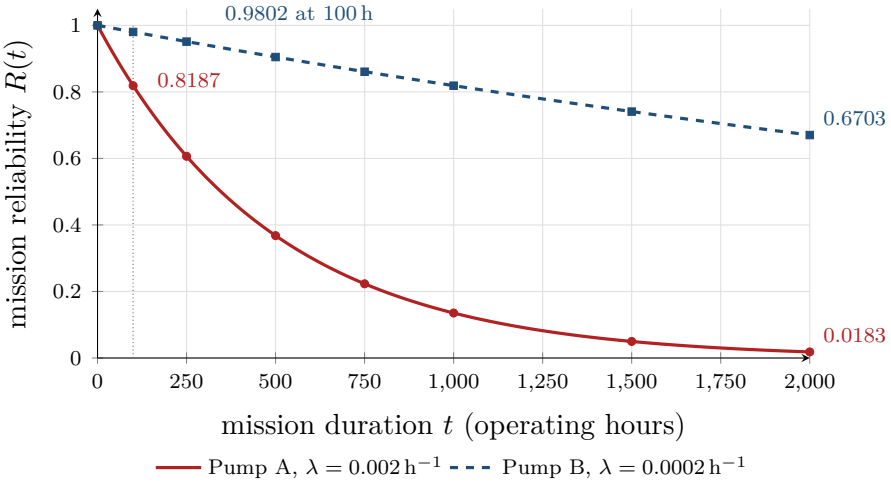


Figure 3.1: Mission reliability of the two pumps, $R(t) = \mathrm{e}^{-\lambda t}$ (equation (3.4)) with constant hazards $\lambda_A = 0.002\,\mathrm{h}^{-1}$ (MTTF = 500 h) and $\lambda_B = 0.0002\,\mathrm{h}^{-1}$ (MTTF = 5000 h); illustrative values. Repairs are not credited: each curve is the probability that no failure occurs before t. The difference between the pumps grows with the mission duration.

This calculation deliberately ignores repairs. Each curve is the probability that no failure occurs at all before t, which is the relevant quantity when any interruption ends the mission. It is not the fraction of time a repairable pump is available, which also depends on how quickly it is restored. The next section introduces that second quantity.

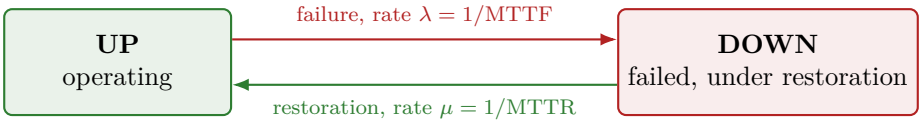
3.3 Availability of a repairable item

A repairable item alternates between operating and unavailable states, and availability measures how much of the time it spends operating. This section states the two-state model, derives the steady-state and the instantaneous availability, and shows that availability and reliability answer different probability questions.

3.3.1 The two-state model

The simplest model of a repairable item has two states: *up*, in which it performs its required function, and *down*, in which it does not because it has failed and is being restored. From the up state the item fails at a constant rate λ (in h^{-1}); from the down state it is restored at a constant rate μ (in h^{-1}). Constant rates mean that the up durations are exponential with mean $\mathrm{MTTF} = 1/\lambda$ and the down durations are exponential with mean $\mathrm{MTTR} = 1/\mu$, and that the item is as good as new after each restoration. A model of this kind, in which the future depends on the present state only and the transitions occur at constant rates, is a continuous-time Markov model.

Figure 3.2 draws it. The green box on the left is the up state and the red box on the right the down state. The upper arrow is the failure transition at rate λ; the lower arrow is the restoration transition at rate μ. The item moves back and forth between the two boxes indefinitely, which is what distinguishes it from the nonrepairable item of Section 3.2, whose failure is final. A later chapter on Markov models derives this model systematically and extends it to several units.



continuous-time Markov model of one repairable item

Figure 3.2: Two-state model of a repairable item. The item fails at the constant rate λ and is restored at the constant rate μ, so up durations are exponential with mean MTTF and down durations exponential with mean MTTR.

3.3.2 Steady-state availability

Let $A(t)$ be the probability that the item is up at time t. Over a long time the influence of the starting state fades, and $A(t)$ settles at a constant value A, the steady-state availability. It is obtained from a balance of probability flows: in steady state the probability flow out of the up state, λA, equals the flow into it, $\mu(1 - A)$. Solving $\lambda A = \mu(1 - A)$ gives

$$A = \frac{\mu}{\lambda + \mu}. \tag{3.6}$$

Substituting $\lambda = 1/\mathrm{MTTF}$ and $\mu = 1/\mathrm{MTTR}$ and multiplying numerator and denominator by $\mathrm{MTTF} \cdot \mathrm{MTTR}$,

$$A = \frac{1/\mathrm{MTTR}}{1/\mathrm{MTTF} + 1/\mathrm{MTTR}} = \frac{\mathrm{MTTF}}{\mathrm{MTTF} + \mathrm{MTTR}}. \tag{3.7}$$

The second form holds well beyond the exponential model. If the up durations $U_1, U_2, \ldots$ and the down durations $D_1, D_2, \ldots$ alternate, each sequence independent and identically distributed with finite means $\mathbb{E}[U] = \mathrm{MTTF}$ and $\mathbb{E}[D] = \mathrm{MTTR}$, the item follows an alternating renewal process. Over n cycles the fraction of time up is $\sum_k U_k / \sum_k (U_k + D_k)$; dividing numerator and

denominator by n and applying the law of large numbers, it tends to $\mathbb{E}[U]/(\mathbb{E}[U]+\mathbb{E}[D])$, which is equation (3.7). This is the standard result for the long-run availability of an alternating renewal process, and it requires only the two means, not exponential distributions. The steady-state availability is estimated from records as the total up time divided by the total observed time. Its complement, the unavailability,

$$1 - A = \frac{\text{MTTR}}{\text{MTTF} + \text{MTTR}} \approx \frac{\text{MTTR}}{\text{MTTF}} \quad \text{when MTTR} \ll \text{MTTF},$$

is the long-run fraction of time down.

3.3.3 Equal availability, unequal reliability

For the two pumps of Table 3.1,

$$A_A = \frac{500}{500+1} = \frac{500}{501} = 0.998004, \qquad A_B = \frac{5000}{5000+10} = \frac{5000}{5010} = 0.998004.$$

The two availabilities are identical, 99.8004 %, although the failure rates and the restoration durations differ by a factor of ten. The reason is that equation (3.7) depends on the ratio MTTR/MTTF only, and the ratio is 1/500 for both pumps. Pump A fails about ten times as often during operation and is repaired ten times faster; Pump B fails less often and each failure costs ten times more downtime. The ratio of the mean operating time to the mean cycle duration is the same.

> **Principle 3.1.** *Two items can have identical availability and very different reliability. Availability depends on the ratio of restoration time to operating time; reliability depends on the failure rate alone.*

3.3.4 Reliability and availability as different probability questions

For a repairable item that is up at $t = 0$, the two measures are the probabilities of two different events,

$$R(t) = \Pr(\text{no failure occurs in } [0,t]), \qquad A(t) = \Pr(\text{the item is up at time } t). \tag{3.8}$$

The second event admits histories in which the item failed before t and was restored in time; the first does not. Every history without a failure in $[0,t]$ ends in the up state, so the first event is contained in the second, and

$$A(t) \geq R(t) \quad \text{for every } t. \tag{3.9}$$

The difference $A(t) - R(t)$ is the probability of the histories in which at least one failure occurred and the item has been restored by t. For a long mission it is large: Pump A is up at hour 2000 with probability close to 0.998, while it survives the same 2000 h without a failure with probability 0.0183. Availability can therefore remain high when the probability of surviving a long mission without failure is low. Which of the two is the relevant measure is decided by the required function: if any interruption defeats the mission, reliability is; if only the state at the moment of demand matters, availability is.

Common mistake

Reading a high availability as a high reliability. A pump that is up 99.8 % of the time can still fail several times a month; a no-failure requirement over a mission is violated at the first failure, regardless of how quickly the item is restored.

3.3.5 Instantaneous availability

For the two-state model, $A(t)$ follows from a balance over a short interval Δt. The item is up at $t + \Delta t$ either because it was up at t and did not fail, with probability $A(t)(1 - \lambda\Delta t)$, or because it was down at t and was restored, with probability $[1 - A(t)]\,\mu\Delta t$, to first order in Δt. Subtracting $A(t)$, dividing by Δt and letting $\Delta t \to 0$,

$$\frac{\mathrm{d}A(t)}{\mathrm{d}t} = -\lambda\, A(t) + \mu\,[1 - A(t)]. \tag{3.10}$$

The first term is the probability flow from up to down, the second the flow from down to up. Rearranged, $\mathrm{d}A/\mathrm{d}t + (\lambda + \mu)A = \mu$, a linear equation whose constant particular solution is $\mu/(\lambda + \mu)$ and whose homogeneous solution decays as $\mathrm{e}^{-(\lambda+\mu)t}$. With the item up at the start, $A(0) = 1$, the constant of the homogeneous part is $1 - \mu/(\lambda + \mu) = \lambda/(\lambda + \mu)$, and

$$A(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu}\,\mathrm{e}^{-(\lambda+\mu)t}. \tag{3.11}$$

As $t \to \infty$ the exponential term vanishes and $A(t) \to A$ of equation (3.6). The approach is governed by the time constant $1/(\lambda + \mu)$, which is close to the MTTR when repairs are short compared with the MTTF: $1/(0.002 + 1) = 0.998\,\mathrm{h}$ for Pump A and $1/(0.0002 + 0.1) = 9.98\,\mathrm{h}$ for Pump B.

Figure 3.3 plots equation (3.11) for both pumps over the first 60 h. The horizontal axis is the time since the pump was put into service in the up state; the vertical axis is the probability of being up, on an axis that spans only 0.9975 to 1. Both curves start at 1 and fall toward the common steady-state value 0.998004, drawn dashed. Pump A reaches it within a few hours, Pump B only after several tens of hours; the Pump B curve is the Pump A curve stretched ten times along the time axis, because both have the same ratio $\lambda/(\lambda + \mu)$ and Pump B has a ten times smaller $\lambda + \mu$. The steady-state value is a long-run result: the availability during a short mission can depend strongly on the starting state and on the restoration process, and for a mission shorter than a few MTTRs it is closer to 1 than A suggests.

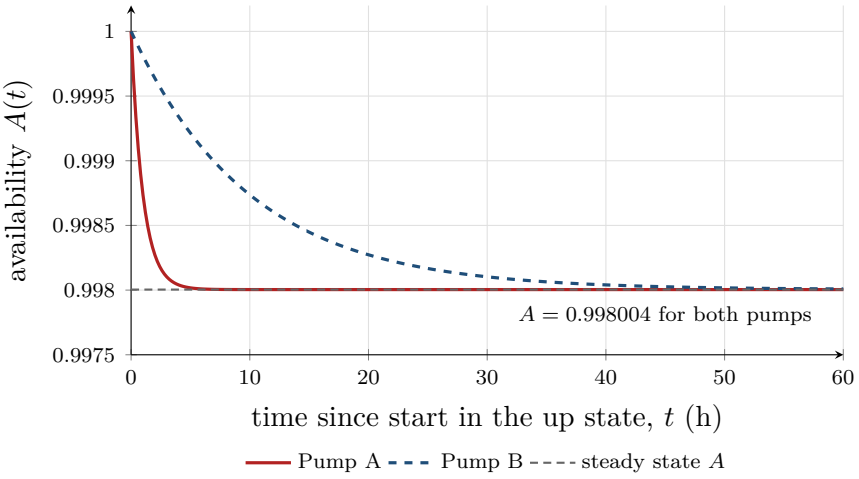


Figure 3.3: Instantaneous availability of the two pumps, $A(t) = \mu/(\lambda+\mu) + [\lambda/(\lambda+\mu)]\,\mathrm{e}^{-(\lambda+\mu)t}$ (equation (3.11)), for Pump A ($\lambda = 0.002\,\mathrm{h}^{-1}$, $\mu = 1\,\mathrm{h}^{-1}$) and Pump B ($\lambda = 0.0002\,\mathrm{h}^{-1}$, $\mu = 0.1\,\mathrm{h}^{-1}$), both up at $t = 0$; illustrative values. Both curves approach the same steady-state availability, with time constants of about 1 h and 10 h.

3.4 Mean times and maintainability

The quantities of the previous sections were written with three mean times whose definitions vary between sources. This section fixes their meaning, shows why a restoration time must be defined before it is used, introduces the maintainability function, and closes with a worked table that compares reliability and availability over a range of parameters.

3.4.1 MTTF, MTBF, MTTR and related means

Table 3.2 lists the mean times used in specifications. The MTTF is the mean time to failure: for a nonrepairable item the mean of its lifetime, and for a repairable item the mean operating duration from a restoration to the next failure. The MTBF is the mean time between failures of a repairable item. The MTTR is the mean time to repair or to restore. Two further means are common in the standards: the mean up time (MUT), the mean operating duration in a repair cycle, and the mean down time (MDT), the mean duration of a down period including every defined source of downtime, which can extend beyond active repair.

Table 3.2: Mean-time measures. The last column gives the convention used in this book.

Measure	Meaning	Typical application	Convention in this book
MTTF	mean time to (first) failure	nonrepairable items; first-failure analysis	mean operating (up) duration; $1/\lambda$ for the exponential model
MTBF	mean time between failures	repairable systems	elapsed time from failure to failure, MTTF + MTTR
MTTR	mean time to repair or restore	maintainability and availability	full restoration time unless stated otherwise; $1/\mu$
MDT	mean down time	operational availability	all defined downtime, beyond active repair
MUT	mean up time	repairable cycles	equal to the MTTF of a repairable item

The MTBF is the term whose meaning varies most. For a stationary alternating renewal process, many sources use MTBF for the mean operating time between failures, which is the MUT. Others define it as the mean elapsed interval between successive failures, downtime included, which is MUT + MDT. This book uses the second meaning and always states whether a time measure includes the repair periods:

$$\mathrm{MTBF} = \mathrm{MTTF} + \mathrm{MTTR}. \tag{3.12}$$

For Pump A the MTBF is 501 h and for Pump B 5010 h. With this convention the steady-state availability of equation (3.7) is also $A = \mathrm{MTTF}/\mathrm{MTBF}$. A source that uses MTBF for the up time writes the same result as $\mathrm{MTBF}/(\mathrm{MTBF} + \mathrm{MTTR})$; the two expressions agree only when the reader knows which convention each source uses.

Figure 3.4 shows the definitions on a timeline. The horizontal axis is calendar time. Green bars are up intervals U_1, U_2, U_3 and red bars are down intervals D_1, D_2; a red triangle marks each failure and a green dot each completed restoration. The brace above U_2 is one up duration, whose mean is the MTTF; the brace above D_2 is one down duration, whose mean is the MTTR; the brace below the axis runs from the first failure to the second and spans one down and one up interval, whose mean is the MTBF of equation (3.12). The intervals are drawn schematically and not to scale: for the pumps, an up interval is 500 or 5000 times longer than a down interval.

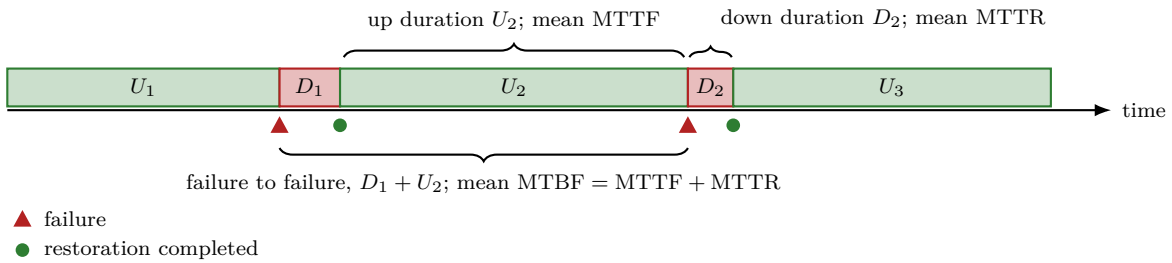


Figure 3.4: Timeline of operating and repair intervals of a repairable item, schematic and not to scale. Up durations have mean MTTF, down durations mean MTTR, and the time from one failure to the next has mean MTBF = MTTF + MTTR in this book's convention.

3.4.2 The definition of restoration time

The MTTR alone can mislead, because sources disagree on what it includes. Consider a pump that fails at 08:00 and returns to service at 15:00. Figure 3.5 decomposes the outage. The horizontal axis is clock time; the bar is divided into the four phases of the restoration, each labeled with its duration: one hour of detection and diagnosis, three hours waiting for a technician, two hours of active repair, and one hour of testing and restart. The upper brace marks the active repair alone; the lower brace marks the whole outage. The phases are illustrative.

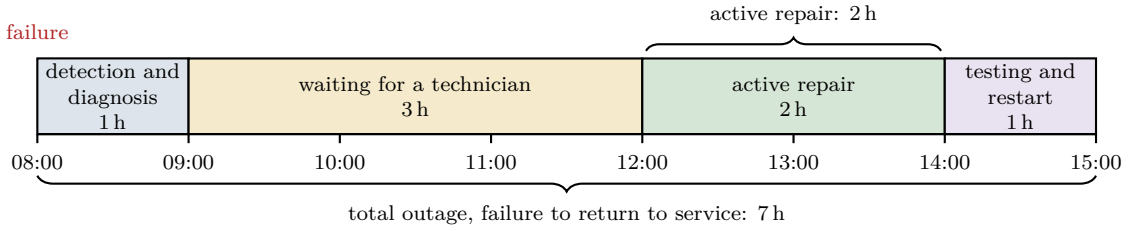


Figure 3.5: Phases of one illustrative restoration of a pump that fails at 08:00. Active repair is 2 h; the outage, from failure to return to service, is 7 h. A restoration-time measure must state which phases it includes.

A report that defines the MTTR as hands-on repair time records 2 h for this event; a report that defines it as full restoration time records 7 h. The two values cannot be substituted into the same availability formula without checking the definitions: inserted into equation (3.7) for a pump with MTTF = 500 h, the first gives an unavailability of $2/502 = 0.0040$ and the second $7/507 = 0.0138$, a factor of 3.5. For operational availability, the relevant downtime includes detection, logistics, repair and restart delays whenever they prevent the required function, which is the MDT of Table 3.2. In this book the MTTR in an availability formula always means that full restoration time.

Common mistake

Assuming that the MTTR always includes all downtime. Active repair time and total restoration time can differ by a large factor; only the total restoration time belongs in an availability calculation.

3.4.3 The maintainability function

Maintainability is the ability of a failed item to be restored to a state in which it performs its required function, under stated maintenance procedures and resources. It is quantified by a

probability distribution. Let T_{R} be the random restoration duration of a failed item, in hours. The maintainability function is its cumulative distribution,

$$M(t) = \Pr(T_{\mathrm{R}} \le t), \tag{3.13}$$

the probability that a restoration is completed within t hours of the failure. It plays for restoration the role that $F(t)$ plays for failure: $M(0) = 0$, M does not decrease, and $M(t) \to 1$. Its mean is the MTTR, by the same integration by parts as equation (3.5),

$$\mathrm{MTTR} = \mathbb{E}[T_{\mathrm{R}}] = \int_0^\infty [1 - M(t)]\,\mathrm{d}t.$$

Maintainability is therefore a distribution, not merely a mean repair time, and two items can have the same MTTR and very different probabilities of a long restoration. It is estimated from maintenance records as the fraction of recorded restorations completed within t, with each restoration measured according to a stated definition of Section 3.4.2. It depends on the design (access, modularity, diagnostics) and equally on the organization (spare parts, staffing, procedures).

3.4.4 Exponential restoration time and its tail

If the restoration duration is exponential with rate μ, the model of Section 3.3.1, then

$$M(t) = 1 - \mathrm{e}^{-\mu t}, \qquad \mathrm{MTTR} = \frac{1}{\mu}. \tag{3.14}$$

For an illustrative MTTR $= 4\,\mathrm{h}$, the restoration rate is $\mu = 1/4 = 0.25\,\mathrm{h}^{-1}$ and the probability of restoration within eight hours is

$$M(8\,\mathrm{h}) = 1 - \mathrm{e}^{-0.25\times 8} = 1 - \mathrm{e}^{-2} = 0.8647,$$

about 86.5 %. The time by which a fraction p of restorations is complete, the p-percentile, solves $M(t_p) = p$, so $t_p = -\ln(1-p)/\mu$; the 90 % percentile is $4 \ln 10 = 9.21\,\mathrm{h}$, more than twice the mean. This result depends on the exponential model. Real restoration durations can have long tails, several modes and a dependence on spare-part availability or on the severity of the failure.

The tail matters even at a fixed mean. Consider two restoration strategies with the same mean downtime of four hours. Under strategy A every restoration takes exactly 4 h. Under strategy B most restorations are quick but a small fraction take much longer; as an illustrative case, 90 % are exponential with mean 2 h and 10 %, those waiting for a spare part, exponential with mean 22 h, so that the mean is $0.9 \times 2 + 0.1 \times 22 = 4\,\mathrm{h}$. Figure 3.6 compares the two with the exponential model of the same mean. The horizontal axis is the time since the failure, in hours; the vertical axis is the maintainability $M(t)$. The gray step is strategy A, the solid blue curve the exponential model, and the dashed red curve strategy B; the dot marks $M(8\,\mathrm{h}) = 0.8647$ of the worked example. Strategy B completes more restorations early, 79 % within 4 h against 63 % for the exponential model, but it is the slowest to approach 1: the probability that a restoration lasts more than a day is $\mathrm{e}^{-6} = 0.0025$ under the exponential model, $0.1\,\mathrm{e}^{-24/22} + 0.9\,\mathrm{e}^{-12} = 0.034$ under strategy B, and zero under strategy A. For a facility that cannot tolerate an extended outage, strategy B is the least attractive of the three, although all three have MTTR $= 4\,\mathrm{h}$. This is why restoration times are specified by percentiles and distributions, not by the MTTR alone.

3.4.5 Reliability, steady-state availability and point availability over a mission

The three quantities derived so far can be computed side by side for any choice of MTTF, MTTR and mission duration H. The model is the two-state model of Section 3.3.1: exponential

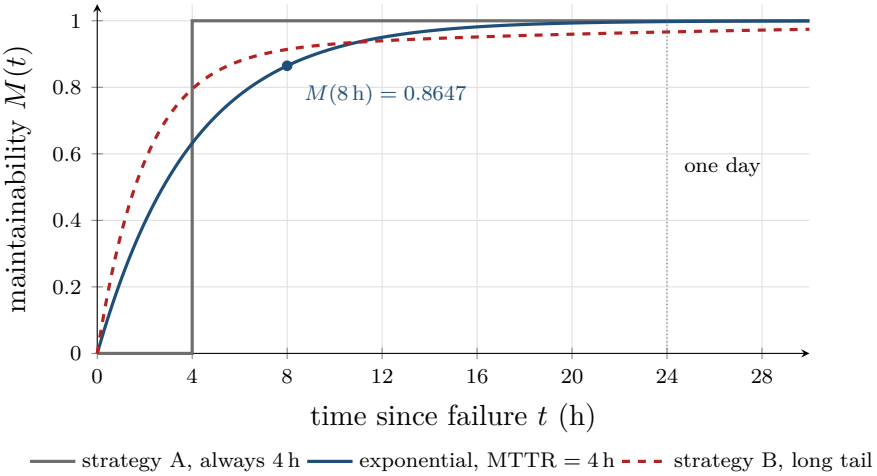


Figure 3.6: Maintainability functions of three restoration processes with the same mean of 4 h: a fixed duration (strategy A); an exponential duration, $M(t) = 1 - \mathrm{e}^{-t/4}$ (equation (3.14)); and an illustrative mixture with a long tail, $M(t) = 1 - 0.9\,\mathrm{e}^{-t/2} - 0.1\,\mathrm{e}^{-t/22}$ (strategy B). Equal means do not imply equal probabilities of a long outage.

up and down durations, with the item up at $t = 0$. With $\lambda = 1/\mathrm{MTTF}$ and $\mu = 1/\mathrm{MTTR}$, the three quantities are

$$R(H) = \mathrm{e}^{-H/\mathrm{MTTF}}, \tag{3.15}$$

$$A = \frac{\mathrm{MTTF}}{\mathrm{MTTF} + \mathrm{MTTR}}, \tag{3.16}$$

$$A_{\mathrm{mission}} = A(H) = A + (1 - A)\,\mathrm{e}^{-H(1/\mathrm{MTTF}+1/\mathrm{MTTR})}. \tag{3.17}$$

The first is the mission reliability, the probability of no failure in $[0, H]$. The second is the steady-state availability. The third is the instantaneous availability of the two-state model at the end of the mission, equation (3.11) evaluated at $t = H$: since $1 - A = \lambda/(\lambda + \mu)$ and $1/\mathrm{MTTF} + 1/\mathrm{MTTR} = \lambda + \mu$, equation (3.17) is the same expression written with the mean times. This book writes it A_{mission} and calls it the point availability at the end of the mission: the probability that the item is up at time H, allowing failures and restorations before then. It is neither the mission reliability nor the average fraction of the mission spent up; the name mission availability is used in the literature for more than one quantity, and the formula is what defines it here.

Table 3.3 evaluates the three quantities. The first five rows are the parameter sets of the comparison; the two rows below the rule are Pump B, over the same 100 h mission as Pump A and over a mission of 10 h, added to show the transient of $A(t)$. The column $(\lambda + \mu)H$ is the mission duration in units of the time constant of equation (3.11).

The second row shows the arithmetic. With MTTF = 500 h, MTTR = 24 h and $H = 100$ h, the rates are $\lambda = 1/500 = 0.002\,\mathrm{h}^{-1}$ and $\mu = 1/24 = 0.041667\,\mathrm{h}^{-1}$, so $\lambda + \mu = 0.043667\,\mathrm{h}^{-1}$ and $(\lambda + \mu)H = 4.3667$. Then

$$\begin{aligned} R(100\,\mathrm{h}) &= \mathrm{e}^{-100/500} = \mathrm{e}^{-0.2} = 0.818731, \\ A &= \frac{500}{500 + 24} = \frac{500}{524} = 0.954198, \\ A_{\mathrm{mission}} &= 0.954198 + 0.045802 \times \mathrm{e}^{-4.3667} = 0.954198 + 0.045802 \times 0.012693 \\ &= 0.954198 + 0.000581 = 0.954780. \end{aligned}$$

The item is up at hour 100 with probability 0.954780, slightly more than the long-run 0.954198, because it started up and 100 h is only about four time constants. It survives the 100 h without

Table 3.3: Mission reliability, steady-state availability and point availability at the end of the mission for the two-state model with exponential up and down durations, item up at $t = 0$ (equations (3.15)–(3.17)); illustrative parameter values.

MTTF (h)	MTTR (h)	H (h)	$(\lambda+\mu)H$	$R(H)$	A	A_{mission}
500	1	100	100.2	0.818731	0.998004	0.998004
500	24	100	4.367	0.818731	0.954198	0.954780
2000	1	100	100.1	0.951229	0.999500	0.999500
2000	24	500	21.08	0.778801	0.988142	0.988142
10 000	8	1000	125.1	0.904837	0.999201	0.999201
5000	10	100	10.02	0.980199	0.998004	0.998004
5000	10	10	1.002	0.998002	0.998004	0.998737

any failure with probability 0.818731, the same as Pump A in the first row, since the MTTF is the same and reliability does not depend on the MTTR.

Three trends follow from the table. First, reliability depends on the MTTF and the mission only, and falls as the mission lengthens; the MTTR does not enter it, which is why the first two rows share $R(H)$. Second, the steady-state availability depends on the ratio MTTR/MTTF only: Pump A (first row) and Pump B (sixth row) share $A = 0.998004$ while their mission reliabilities are 0.818731 and 0.980199; and raising the MTTR from 1 to 24 h at a fixed MTTF lowers A from 0.998004 to 0.954198. Third, the point availability equals the steady-state value to six decimals whenever the mission exceeds about ten time constants, which for these parameters means a mission longer than about ten MTTRs; it differs only when the mission is short compared with the restoration, as in the second and the last rows. In the last row, Pump B over 10 h, the point availability 0.998737 exceeds the steady-state value because the item started up and 10 h is only one time constant, so the probability of being down has not yet risen to its long-run value; it also exceeds the mission reliability 0.998002, as inequality (3.9) requires. In every row $A_{\text{mission}} \geq R(H)$.

Remark 3.1. The average availability over the mission, the expected fraction of $[0, H]$ spent up, is a fourth quantity. Integrating equation (3.11),

$$\bar{A}(H) = \frac{1}{H}\int_0^H A(t)\,\mathrm{d}t = \frac{\mu}{\lambda+\mu} + \frac{\lambda}{(\lambda+\mu)^2\,H}\left[1 - \mathrm{e}^{-(\lambda+\mu)H}\right].$$

For the second row of Table 3.3 it is $0.954198 + 0.045802 \times (1 - 0.012693)/4.3667 = 0.964554$, higher than both A and $A(H)$ because the item is almost certainly up during the first hours of the mission. Some texts call this average the mission availability. Whichever name is used, the quantity must be stated by its formula.

Principle 3.2. *For a repairable item that starts up, $A(t) \geq R(t)$ at every t; the point availability approaches the steady-state value within a few restoration times, while the mission reliability keeps falling with the mission duration and does not depend on the restoration at all.*

The distinction matters most when an equipment failure triggers a consequence that restoring the equipment does not reverse. The next section attaches consequences to failures.

3.5 Risk and the frequency of interruptions

Reliability measures the probability that a required function is maintained; risk asks which undesirable outcomes can occur and how consequential they are. This section defines risk as an

expected consequence, computes it for the two pumps, and separates expected downtime from interruption frequency.

3.5.1 Risk as probability combined with consequences

Let C be a random consequence or loss of operating an item over a defined period, measured in a defined unit: money, energy not served, outage duration, lost production. One common quantitative measure of risk is the expected loss,

$$\mathcal{R} = \mathbb{E}[C]. \tag{3.18}$$

When the possible outcomes form a discrete collection of mutually exclusive scenarios $S_1, S_2, \ldots$, each with probability $\Pr(S_i)$ and consequence $C(S_i)$, the expectation is the sum

$$\mathcal{R} = \sum_i \Pr(S_i)\, C(S_i). \tag{3.19}$$

When the consequence is a monetary loss over one year, $\mathcal{R}$ is the expected annual loss, in R\$/yr. The definition makes risk a property of the pair (probability, consequence): a failure probability alone is not a risk, and a consequence alone is not a risk.

A single expected value does not describe every aspect of risk. Two portfolios of scenarios can have the same $\mathcal{R}$, one made of frequent small losses and the other of a rare catastrophic one; the expected value treats them alike, while an operator, an insurer or a regulator usually does not. Rare outcomes with severe consequences, above all harm to people, may need separate treatment: a bound on their probability, a separate safety requirement, or the full distribution of C rather than its mean.

3.5.2 Expected annual loss of the two pumps

Suppose that every loss of the pumping function triggers a costly process interruption, whatever the duration of the repair, and that each interruption costs R\$ 50 000 (illustrative). The consequence is then proportional to the number of failures, not to the downtime. For a repairable item modeled as an alternating renewal process, the long-run number of failures per unit calendar time tends to 1/MTBF, one failure per mean cycle; this is the elementary renewal theorem, a standard result. For a pump in continuous service over a year of 8760 h, the expected number of failures per year is therefore approximately

$$\mathbb{E}[N_{\mathrm{yr}}] \approx \frac{8760\,\mathrm{h}}{\mathrm{MUT} + \mathrm{MDT}} = \frac{8760\,\mathrm{h}}{\mathrm{MTBF}}, \tag{3.20}$$

with the MTBF of equation (3.12). For the two pumps,

$$\mathbb{E}[N_A] \approx \frac{8760}{501} = 17.49\,\mathrm{yr}^{-1}, \qquad \mathbb{E}[N_B] \approx \frac{8760}{5010} = 1.75\,\mathrm{yr}^{-1},$$

and by equation (3.19), with one scenario per failure, the expected annual losses are

$$\mathcal{R}_A \approx 17.485 \times \mathrm{R\$}\,50\,000 = \mathrm{R\$}\,874\,251\,\mathrm{yr}^{-1}, \qquad \mathcal{R}_B \approx 1.7485 \times \mathrm{R\$}\,50\,000 = \mathrm{R\$}\,87\,425\,\mathrm{yr}^{-1}.$$

Table 3.4 collects the comparison. The rows are the steady-state availability, the expected failures per year, the expected downtime per year (Section 3.5.3) and the expected annual loss. The two pumps are equal on the first and third rows and differ by a factor of ten on the second and fourth. These are long-run renewal averages, not exact predictions for a particular first calendar year: a pump that starts in the up state has, in its first year, very slightly more expected failures than the long-run rate gives, because its availability starts at 1 rather than at A.

Table 3.4: Comparison of the two pumps for a consequence of R$ 50 000 per interruption (illustrative), long-run averages over a year of 8760 h.

Measure	Pump A	Pump B
Steady-state availability A	0.998004	0.998004
Expected failures per year	17.49	1.75
Expected downtime per year	17.49 h	17.49 h
Expected annual loss $\mathcal{R}$ (per year)	R$ 874 251	R$ 87 425

Principle 3.3. *Availability alone is insufficient for selecting equipment. Equal availabilities can hide a tenfold difference in interruption frequency, and when consequences are incurred per interruption, a tenfold difference in risk.*

3.5.3 Expected downtime and interruption frequency

For a facility that operates continuously, the expected downtime per year follows from the steady-state availability,

$$\mathbb{E}[D_{\mathrm{yr}}] \approx 8760\,\mathrm{h} \times (1 - A). \tag{3.21}$$

For both pumps, $\mathbb{E}[D_{\mathrm{yr}}] \approx 8760 \times (1 - 0.998004) = 8760 \times 0.001996 = 17.49$ h per year. The same number results from counting failures: the expected number of failures per year times the mean restoration time is 17.485×1 h for Pump A and 1.7485×10 h for Pump B, both 17.49 h. The two calculations agree in general, since $(8760/\mathrm{MTBF}) \cdot \mathrm{MTTR} = 8760\,\mathrm{MTTR}/(\mathrm{MTTF} + \mathrm{MTTR}) = 8760\,(1 - A)$.

The two pumps therefore have the same long-run annual downtime, but its structure differs. Figure 3.7 shows one illustrative year of each. Each horizontal bar is one year of calendar time; red marks are outages. Pump A has eighteen outages of one hour, Pump B two outages of ten hours, close to the expected 17.5 and 1.75. The widths of the outages are exaggerated fifty times so that they are visible; at true scale a one-hour outage is 0.0014 cm wide on this axis. The total red length is about the same in both bars, which is the equal downtime; the number of marks differs tenfold, which is the different interruption frequency.

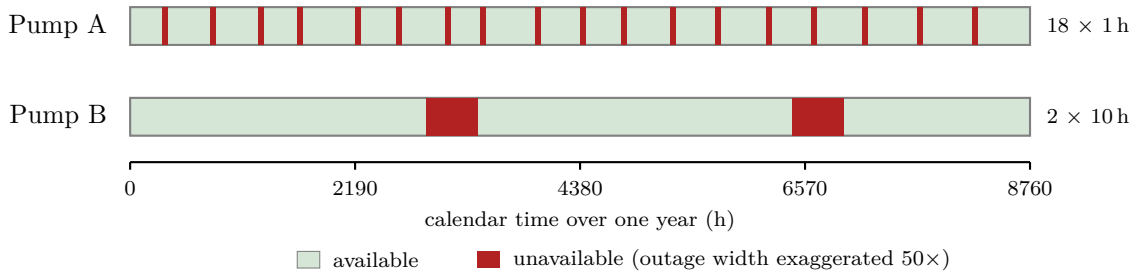


Figure 3.7: One illustrative year of interruptions of the two pumps. Both are down about 17.5 h in expectation (equation (3.21)); Pump A in many short interruptions, Pump B in a few long ones. Outage positions are illustrative and widths exaggerated.

For some systems a short interruption is tolerable and only the total downtime matters. For others even a millisecond-long interruption has major consequences, because the process it supplies cannot ride through it: a semiconductor fabrication line loses the wafers in process, a data center loses the computation in memory, and an electrical network can trip protection and spread the disturbance. For such systems the number of interruptions, not their total duration, sets the consequence. A reliability assessment may therefore need to quantify, besides

the expected downtime, the interruption frequency, the distribution of outage durations, and the severity of each interruption.

3.6 Power-system reliability, the RAMS framework and common mistakes

Electric power systems have their own vocabulary for the distinctions of this chapter, and engineering standards group them into the RAMS framework. This section states the power-system measures, places the chapter's four measures in RAMS, and lists the common mistakes in their use.

3.6.1 Adequacy and security

Power-system reliability is divided into two aspects. *Adequacy* is the existence of sufficient facilities within the system, generation and transmission capacity, to satisfy the demand, accounting for planned and unplanned outages of those facilities. *Security* is the ability of the system to withstand sudden disturbances, such as the loss of a line or a generator or a short circuit, and to remain in acceptable operation through and after them. Adequacy is a static question about capacity against demand over a period; security is a dynamic question about the response to an event.

The two are related but different. A system with ample installed capacity can still be insecure, if a single disturbance triggers a cascade of trips before that capacity can be used; a system that would withstand every credible disturbance can still be inadequate, if maintenance and forced outages leave too little capacity during the peak. The first is closer to the availability question of this chapter, the probability that enough capacity is in the up state at a given time; the second is closer to the reliability question, whether the system continues to function through an event without interruption.

3.6.2 Generation adequacy at one instant

Take a demand of $L = 180\,\mathrm{MW}$ supplied by two generators of capacity $C_1 = C_2 = 100\,\mathrm{MW}$ each. If both are available, the available capacity is $C_{\text{available}} = 200\,\mathrm{MW} \geq L$. If either is unavailable, the remaining capacity is 100 MW, which cannot meet the demand. Assume that each generator is unavailable at a randomly selected time with probability 0.02, its unavailability $1 - A$, and that the two states are independent. The available capacity takes three values:

$$\begin{aligned} \Pr(C_{\text{available}} = 200\,\mathrm{MW}) &= 0.98^2 = 0.9604, \\ \Pr(C_{\text{available}} = 100\,\mathrm{MW}) &= 2 \times 0.98 \times 0.02 = 0.0392, \\ \Pr(C_{\text{available}} = 0\,\mathrm{MW}) &= 0.02^2 = 0.0004. \end{aligned}$$

The probability of insufficient generation at that instant is

$$\Pr(C_{\text{available}} < L) = 1 - 0.9604 = 0.0396 = 3.96\,\%. \tag{3.22}$$

Figure 3.8 draws this distribution, which power engineers call a capacity outage probability table. The horizontal axis is the available capacity in megawatts; the vertical axis, on a logarithmic scale, is the probability of each capacity state. Red bars are states below the demand, the green bar the state that meets it, and the dashed line marks the demand of 180 MW. The probability of a shortfall is the total height of the red bars, $0.0392 + 0.0004$.

This number is a snapshot loss-of-load probability (LOLP) under the stated model: the probability that available capacity falls short of demand at one randomly chosen instant. It is not a mission failure probability, which would ask whether a generator fails during an interval,

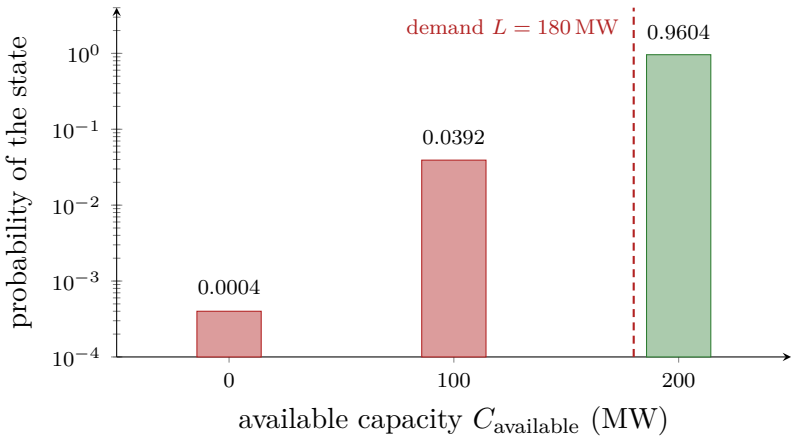


Figure 3.8: Capacity states of two independent 100 MW generators, each unavailable with probability 0.02, against a demand of 180 MW (illustrative). States below the demand (red) have total probability 0.0396, the loss-of-load probability of equation (3.22).

and it is not an annual loss-of-load expectation, which sums such probabilities over the hours of a year.

3.6.3 Loss-of-load expectation and expected unserved energy

Over a study period divided into N intervals of length Δt, with demand $L(t)$ and available capacity $C_{\text{available}}(t)$ in interval t, the loss-of-load expectation (LOLE) in its simplest time-discretized form is

$$\text{LOLE} = \sum_{t=1}^{N} \Pr\big(C_{\text{available}}(t) < L(t)\big)\,\Delta t. \tag{3.23}$$

With $\Delta t = 1$ h it is the expected number of hours with a shortfall over the study period, in hours per year for an annual study. It is an expected value: the sum of the indicator of a shortfall over the intervals has expectation equal to the sum of the probabilities, whatever the correlation between intervals. The expected unserved energy (EUE), also called the expected energy not supplied (EENS), weights each shortfall by its size,

$$\text{EUE} = \sum_{t=1}^{N} \mathbb{E}\Big[\big(L(t) - C_{\text{available}}(t)\big)_{+}\Big]\,\Delta t, \qquad (x)_{+} = \max(x, 0), \tag{3.24}$$

in megawatt-hours over the period. The function $(x)_+$ keeps only the shortfalls: a surplus of capacity contributes nothing. LOLE and EUE are the fundamental quantities of power-system adequacy studies; a later part of the book develops them in full.

For the two generators, if the demand stayed at 180 MW for all 8760 h of a year and the hourly states followed the same probabilities (an illustrative simplification), equation (3.23) would give LOLE $= 0.0396 \times 8760 = 346.9$ h per year. The expected shortfall in one hour is $0.0392 \times (180 - 100) + 0.0004 \times (180 - 0) = 3.136 + 0.072 = 3.208$ MW, and equation (3.24) would give EUE $= 3.208$ MW $\times$ 8760 h $= 28\,102$ MWh per year.

3.6.4 Equal LOLE, unequal EUE

Two generation portfolios can have the same expected shortage hours and very different expected unserved energy. System A experiences frequent small shortfalls; System B rare large ones. Table 3.5 gives an illustrative pair. System A has two one-hour shortfalls of 5 MW per year; System B has one eight-hour shortfall of 500 MW every four years. Both have LOLE $= 2$ h per

year. Their EUEs are $2 \times 1\,\text{h} \times 5\,\text{MW} = 10\,\text{MWh}$ and $0.25 \times 8\,\text{h} \times 500\,\text{MW} = 1000\,\text{MWh}$ per year, a hundredfold difference. A time-based measure alone does not capture the difference in their adequacy outcomes, in the same way that the availability of Section 3.3.3 does not capture the difference in failure frequency of the two pumps.

Table 3.5: Two illustrative generation portfolios with equal LOLE and unequal EUE.

System	Shortfalls per year	Duration (h)	Shortfall (MW)	LOLE (h/yr)	EUE (MWh/yr)
A	2	1	5	2	10
B	0.25	8	500	2	1000

3.6.5 The RAMS framework

The measures of this chapter belong to a broader engineering framework called RAMS: reliability, availability, maintainability and safety. Table 3.6 states each dimension with its meaning and an example question for the cooling and pumping systems of this book. The S in RAMS stands for safety, not for risk. Safety is freedom from unacceptable risk of harm, to people above all; risk analysis is closely related to safety engineering, but it also extends to financial, environmental, operational and other consequences.

Table 3.6: The four dimensions of RAMS, with example questions.

Dimension	Meaning	Example question
Reliability	continued performance without failure	Will the chiller survive the next 1000 h?
Availability	probability of being operational	Will cooling be available at 14:00 tomorrow?
Maintainability	ability to restore performance	Can the pump be restored within four hours?
Safety	freedom from unacceptable risk of harm	Could a cooling-system failure expose people to dangerous conditions?

Reliability and safety are related but not interchangeable. A data center outage may carry a large financial risk with little direct threat to human safety. An industrial pressure-vessel failure may carry a serious safety risk even if it is extremely unlikely. High reliability does not by itself establish acceptable safety: when the consequences of a rare failure include loss of life, the risk and the safety requirements must be evaluated as well.

3.6.6 Common mistakes

Table 3.7 lists the misreadings of the four measures that occur most often, each with the correct interpretation. Most of them substitute one measure for another: availability for reliability, one definition of a mean time for another, a probability for a risk. The last row anticipates Chapter 5: the constant failure rate used throughout this chapter is a modeling choice, not a property of every component.

The measures of this chapter treat a failure as an event with a rate; Chapter 4 turns to the physical mechanisms, causes, modes and effects that produce those failures.

Table 3.7: Common mistakes in the use of reliability, availability, maintainability and risk measures.

Mistake	Correct interpretation
High availability implies high reliability.	Fast restoration can produce high availability despite frequent failures (Section 3.3.3).
The MTBF is the same calendar time between outages in every source.	Definitions vary; operating and repair intervals must be distinguished (Section 3.4.1).
The MTTR always includes all downtime.	Active repair time and total restoration time can differ (Section 3.4.2).
Equal availability means equivalent operational performance.	Interruption frequency and the distribution of outage durations can differ (Section 3.5.3).
Failure probability is risk.	Risk also involves the consequences (Section 3.5.1).
Expected loss completely characterizes risk.	Tail events and unacceptable outcomes may need separate analysis (Section 3.5.1).
A constant failure rate is always reasonable.	Aging and degradation can require time-varying hazard models (Section 3.2.2).

3.7 What to retain

Definition 3.1 (Reliability, availability, maintainability and risk)**.** Reliability is the probability of no failure over a mission, $R(t) = \Pr(T > t)$; availability is the probability of being operational at a time, $A(t)$, or the long-run fraction of time operational, $A = \mathrm{MTTF}/(\mathrm{MTTF} + \mathrm{MTTR})$; maintainability is the probability of restoration within a time, $M(t) = \Pr(T_{\mathrm{R}} \leq t)$; and risk is the expected consequence, $\mathcal{R} = \mathbb{E}[C]$, for a defined loss measure. The four are related but not substitutable, and a sound engineering decision chooses the measures appropriate to the system's required function and to the consequences of its failure.

A system can have high availability and frequent failures; high reliability and very long restorations when it does fail; a low failure probability and severe consequences; and reliable components with poor system-level performance because of shared dependencies. Table 3.8 lists the chapter's terms, their meaning and units, and the subsection that defines each.

Table 3.8: The terms of this chapter and where each is defined.

Term	Meaning	Expression	Units	Defined in
Reliability	no failure in $[0,t]$	$R(t)=\Pr(T>t)$	–	3.2.1
Cumulative failure probability	failure by t	$F(t)=1-R(t)$	–	3.2.1
Constant failure rate	age-independent hazard	$R(t)=\mathrm{e}^{-\lambda t}$	h^{-1}	3.2.2
MTTF	mean operating time to failure	$\int_0^\infty R\,\mathrm{d}t=1/\lambda$	h	3.2.2
Two-state model	up and down states, rates λ, μ	Figure 3.2	–	3.3.1
Steady-state availability	long-run fraction of time up	MTTF/(MTTF + MTTR)	–	3.3.2
Instantaneous availability	up at time t	equation (3.11)	–	3.3.5
MTBF	mean time from failure to failure	MTTF + MTTR	h	3.4.1
MTTR, MDT	mean restoration, mean down time	$1/\mu$	h	3.4.1
Maintainability function	restored within t	$M(t)=\Pr(T_\mathrm{R}\le t)$	–	3.4.3
Point availability at mission end	up at $t=H$	$A_\mathrm{mission}=A(H)$	–	3.4.5
Risk	expected consequence	$\mathcal{R}=\sum_i \Pr(S_i)\,C(S_i)$	R$/yr	3.5.1
Failures per year	long-run interruption frequency	8760/MTBF	yr^{-1}	3.5.2
Expected downtime	hours down per year	$8760\,(1-A)$	h/yr	3.5.3
Adequacy, security	enough capacity; withstanding disturbances	–	–	3.6.1
LOLP	shortfall at one instant	$\Pr(C_\mathrm{available}<L)$	–	3.6.2
LOLE	expected shortfall hours	equation (3.23)	h/yr	3.6.3
EUE	expected unserved energy	equation (3.24)	MWh/yr	3.6.3
RAMS	reliability, availability, maintainability, safety	Table 3.6	–	3.6.5

Further reading. Rausand, Barros and Høyland, *System Reliability Theory*, third edition, for reliability, availability, repairable systems and the fundamental terminology; O'Connor and Kleyner, *Practical Reliability Engineering*, fifth edition, for RAM concepts and practical reliability measures; Billinton and Allan, *Reliability Evaluation of Power Systems*, second edition, for adequacy, unavailability, LOLE and expected energy not supplied; Kumamoto and Henley, *Probabilistic Risk Assessment and Management for Engineers and Scientists*, for the connections between reliability, consequences and risk.

3.8 Exercises

Exercise 3.1. (a) Two pumps have identical steady-state availability, but Pump A fails ten times as often as Pump B. State what can be concluded about their reliability over a fixed uninterrupted mission and about their interruption frequency and restoration times. (b) A component fails after 100 h and is repaired successfully. State whether it satisfied a 200 h no-failure mission. (c) For an item with MTTF = 500 h and MTTR = 1 h, compute the mission reliability over 200 h, the probability of at least one failure in that time, and the point availability at $t = 200$ h, and explain why the last exceeds the first. (Answers: (a) only that MTTR/MTTF is the same; availability does not determine failure frequency, repair duration or mission reliability, which can differ greatly. (b) No; a no-failure requirement is violated at the first failure, regardless of later restoration. (c) $\mathrm{e}^{-0.4} = 0.6703$, 0.3297, and $A(200\,\mathrm{h}) = 0.998004$, because availability credits restoration after a failure.)

Exercise 3.2. The restoration time of a chiller is exponential with MTTR = 6 h. (a) Compute the restoration rate μ. (b) Compute $M(4\,\mathrm{h})$ and $M(12\,\mathrm{h})$. (c) Compute the time by which 90 % of restorations are complete. (d) Compare with a crew that always restores in exactly 6 h. (e) State which quantity gives the probability that a failed item is restored within a specified duration. (Answers: (a) $0.1667\,\mathrm{h}^{-1}$; (b) 0.4866 and 0.8647; (c) $6 \ln 10 = 13.82$ h; (d) the fixed crew gives $M(4\,\mathrm{h}) = 0$ and $M(12\,\mathrm{h}) = 1$, so it is slower early but never leaves a long tail; (e) the maintainability function $M(t) = \Pr(T_\mathrm{R} \le t)$, which describes the restoration-time distribution.)

Exercise 3.3. The demand of Section 3.6.2, $L = 180$ MW, is supplied by three independent 100 MW generators, each unavailable with probability 0.02. (a) List the capacity states and their probabilities. (b) Compute the loss-of-load probability. (c) Assuming the demand stays at 180 MW for all 8760 h of a year, compute the LOLE and the EUE, and compare them with the two-generator values of Section 3.6.3. (d) State which power-system measure quantifies the expected amount of demand not served. (Answers: (a) 300 MW with 0.941192, 200 MW with 0.057624, 100 MW with 0.001176, 0 MW with 0.000008; (b) 0.001184; (c) 10.37 h per year and $0.09552\,\mathrm{MW} \times 8760\,\mathrm{h} = 836.8$ MWh per year, against 346.9 h and 28 102 MWh; (d) the EUE, which accounts for the size and the duration of each shortfall.)

Exercise 3.4. A cooling unit X has MTTF = 2000 h and MTTR = 24 h; an alternative unit Y has MTTF = 10 000 h and MTTR = 8 h. Each interruption costs R$ 20 000 and each hour of downtime a further R$ 1000 (illustrative). (a) For each unit, compute the steady-state availability, the expected number of failures per year and the expected downtime per year. (b) Compute the expected annual loss of each. (c) A rare failure mode of unit Y can release refrigerant into an occupied room and cause a loss of life. State whether the high reliability of Y establishes that it is acceptably safe. (Answers: (a) X: 0.988142, $4.328\,\mathrm{yr}^{-1}$, 103.9 h; Y: 0.999201, $0.8753\,\mathrm{yr}^{-1}$, 7.00 h; (b) X: R$ 86 561 + R$ 103 874 = R$ 190 435 per year; Y: R$ 17 506 + R$ 7002 = R$ 24 508 per year; (c) no; a low failure probability does not by itself establish acceptable safety when the consequences may be severe, and the risk and the safety requirements must be evaluated as well.)

Chapter 4

Failure Mechanisms and Failure Mode Analysis

4.1 Failure cause, mechanism, mode and effect

Chapter 3 computed reliability, availability and risk from a failure rate λ that was taken as a given parameter of each component. A failure rate summarizes the outcome of physical processes: material is removed, cracks grow, insulation ages, a seal hardens, a sensor drifts. Where those processes run faster or slower, because the load, the temperature or the operating regime changes, the rate changes with them. This chapter asks

why engineered systems fail, how physical processes lead to failures, and how failures can be identified and evaluated systematically before they occur.

The answer has two halves. The first is physical: a failure is the end of a chain that starts with an initiating condition and runs through a physical, chemical or electrical process to an observable loss of function, and that process can often be modeled. The second is systematic: the ways in which every item of a system can fail are listed, traced to their effects on the required function and ranked, which is failure mode and effects analysis. The chapter first separates the four concepts that describe a failure; it then surveys the physical failure mechanisms, turns them into mathematical failure models, works mechanical fatigue in detail as an example of physics-based prediction, builds an FMEA of the cooling-water pump system, and ends with criticality, quantitative risk and the relation between FMEA and fault tree analysis.

Jargon: failure cause, failure mechanism, failure mode, failure effect, physics of failure, FMEA, FMECA, functional failure, severity, occurrence, detection, RPN

A *failure cause* is the condition or event that initiates a failure, such as inadequate lubrication or an overload. A *failure mechanism* is the physical, chemical or electrical process that leads from the cause to the failure, such as fatigue, wear, corrosion or dielectric breakdown. A *failure mode* is the manner in which an item fails to perform its required function, as it is observed: a bearing seizes, a valve sticks closed, a sensor reads high. A *failure effect* is the consequence of the failure mode at the component, subsystem or system level. *Physics of failure* is the practice of modeling failure mechanisms quantitatively, so that the time or probability of failure follows from the stresses an item experiences. *FMEA* (failure mode and effects analysis) is a structured, bottom-up analysis that lists the potential failure modes of the elements of a system and traces their effects on the required system function; *FMECA* (failure modes, effects and criticality analysis) adds an evaluation of the relative criticality of each mode. A *functional failure* is a description of inadequate performance of a function, without the reason for it. In criticality screening, *severity* S, *occurrence* O and *detection* D are ordinal ratings, usually on a scale of 1 to 10, of how serious the consequence is, how likely the mode is, and how difficult the mode is to detect before its effect occurs; the *risk priority number* (RPN) is their product.

Key idea

A failure is a chain from an initiating cause through a physical mechanism to an observable mode and its effect on the required function. Modeling the mechanism turns a failure probability from a fixed property of a component into a function of stress, temperature and time; FMEA organizes the modes and their effects so that they can be quantified, and ranks them by their contribution to system loss, not by how often they occur.

4.1.1 The four concepts

The same event is often described in different words at different levels of a system, and the terminology of failure analysis is confusing for that reason. Four concepts are kept apart in this book: the initiating condition (the cause), the physical process (the mechanism), the observable failure (the mode) and the resulting consequence (the effect). Table 4.1 defines each and illustrates it with a bearing failure in the cooling-water pump. The rows read in the order in which the events occur: inadequate lubrication starts the sequence, contact fatigue and surface damage are what the lubrication deficit does to the rolling surfaces, the seizure of the bearing is what an observer sees, and the stopped pump with its lost cooling-water flow is what the plant experiences. The categories can overlap in practice, and standards differ in the details of their wording; the four-way distinction above is the one in common use in failure analysis and the one used throughout this book.

Figure 4.1 draws the sequence as a chain. The upper row is the analysis at the level of the bearing: its four boxes are the cause, the mechanism, the failure mode and the effect, joined by arrows in the direction of time. The dashed loop above the mechanism box marks a feedback that the linear chain hides, discussed in the next subsection. The lower row is the analysis of the same event one level up, at the level of the pump, and the dashed arrow between the rows shows that the failure mode of the bearing reappears there as a cause.

4.1.2 Feedback and the level of analysis

The chain of Figure 4.1 is useful, but real failure mechanisms are not always linear. Excessive vibration can damage a bearing, and a damaged bearing generates more vibration; the mechanism

Table 4.1: The four concepts that describe a failure, with the pump bearing as the example.

Concept	Meaning	Pump example
Failure cause	The condition or event that initiates failure	Inadequate lubrication
Failure mechanism	The physical or chemical process leading to failure	Contact fatigue and surface damage
Failure mode	The manner in which the item fails to perform its function	Bearing seizure
Failure effect	The consequence at the component, subsystem or system level	Pump stops, causing loss of cooling-water flow

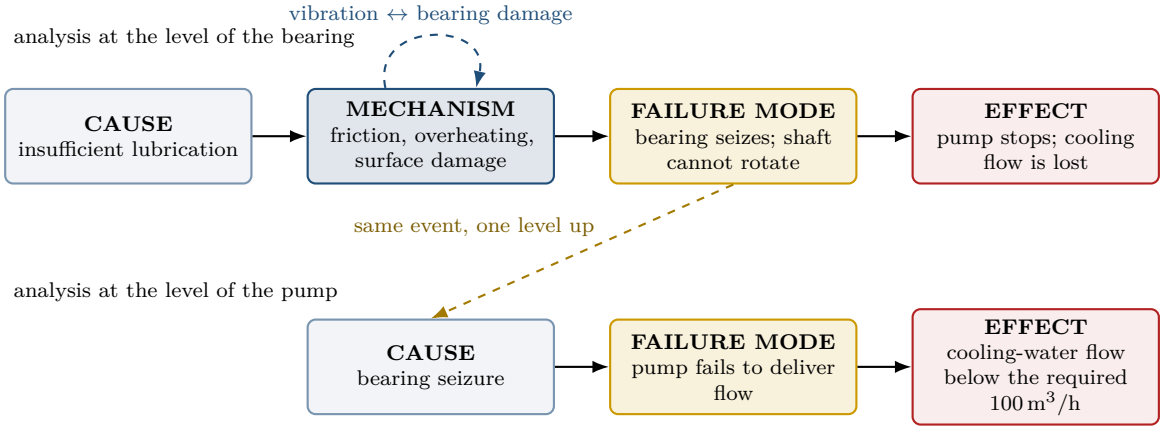


Figure 4.1: An illustrative causal sequence for a pump bearing. Upper row: cause, mechanism, failure mode and effect at the level of the bearing, with the feedback between vibration and bearing damage (dashed loop). Lower row: at the level of the pump, the seizure of the bearing is a cause of the pump's failure mode "fails to deliver flow".

then feeds itself, and the damage grows faster the further it has progressed. A thermal runaway, in which a hot component loses efficiency and generates more heat, and a corrosion pit that concentrates stress and starts a fatigue crack, are feedback loops of the same kind. A feedback of this kind is one reason why many degradation processes accelerate toward the end of life, which the fatigue crack-growth model of Section 4.4 shows quantitatively.

The second complication is the level of analysis. Bearing seizure is a failure mode when the item under analysis is the bearing; it is a cause when the item is the pump, whose failure mode is "fails to deliver flow"; and the pump's failure mode is in turn a cause of the failure of the cooling system to supply its required $100\,\mathrm{m}^3/\mathrm{h}$. The level of the system hierarchy at which an item is analyzed is called its indenture level. A statement of the form "X is the failure mode" is therefore incomplete until the item and the function under analysis are specified, which is the discipline of system boundaries and required functions of Chapter 2.

Principle 4.1. *Cause, mechanism, mode and effect are defined relative to a stated item and a stated function. The failure mode of an item at one level of the system hierarchy is a cause of the failure mode of the item at the next level up.*

4.2 Physical failure mechanisms

Most physical equipment failures arise through identifiable mechanisms. This section groups the common ones into mechanical, thermal, chemical, electrical and environmental processes, describes for each what drives it and what failure modes it produces, and adds the failures of software, control and operation, which remove a function without any physical degradation.

4.2.1 Mechanical fatigue

Repeated mechanical loading can initiate and grow cracks even when every individual load is below the static strength of the material. A crack initiates at a stress concentration, such as a surface scratch, a weld toe, a keyway or an inclusion; it then grows a small distance with each load cycle; and the part fractures when the remaining cross-section can no longer carry the peak load. Rotating shafts, turbine blades, aircraft structures and bolted or welded connections exposed to repeated stress cycles fail in this way. The damage depends on

- the stress range and the mean stress;
- the number and the sequence of loading cycles;
- the material properties and the surface condition;
- existing defects and the crack geometry;
- the temperature and the environmental exposure.

Fatigue is the standard example of a cumulative mechanism: each cycle adds a small, invisible increment of damage, and the failure appears only when the accumulated damage reaches a critical value. Because the damage per cycle can be related to the stress range by laboratory tests, fatigue is also the mechanism for which physics-based life prediction is most developed; Section 4.4 works it in detail.

4.2.2 Wear and friction

Mechanical contact between moving surfaces gradually removes or redistributes material. The common forms are adhesive wear, in which microscopic contact points weld and tear; abrasive wear, in which hard particles or asperities cut the softer surface; erosive wear, in which a fluid carrying particles or cavitation bubbles removes material; and fretting wear, in which surfaces in nominally fixed contact rub through small oscillating displacements. The standard engineering relation for sliding wear makes the worn volume proportional to the normal load and to the

sliding distance, and inversely proportional to the hardness of the softer surface, so that wear accumulates steadily with operating time at a rate set by the load.

Wear changes performance before it causes breakage. In the cooling-water pump, worn wear rings increase the internal clearance between the impeller and the casing, more water recirculates inside the pump, and the head and flow at a given speed fall. The pump keeps running; it fails in the sense of Chapter 2 when its flow drops below the required $100\,\mathrm{m}^3/\mathrm{h}$. Wear also reduces efficiency and generates heat, and the heat can activate a second mechanism, such as the loss of lubricant viscosity that leads to the bearing seizure of Figure 4.1.

4.2.3 Corrosion

Corrosion is the degradation of a material by chemical or electrochemical reaction with its environment. In the electrochemical case, metal dissolves at anodic sites and a reduction reaction, usually of dissolved oxygen, takes place at cathodic sites, with water or a moist deposit as the electrolyte between them. Its forms include uniform wall loss of a pipeline; pitting corrosion, in which the attack concentrates in small deep pits; galvanic corrosion, in which the less noble of two metals in electrical contact corrodes preferentially; and corrosion-assisted cracking, in which a corrosive environment and a tensile stress together grow cracks, as in stress corrosion cracking and corrosion fatigue.

Corrosion reduces wall thickness, compromises electrical connections by raising their contact resistance, and increases the susceptibility of a part to fatigue or fracture, because a pit is a stress concentration from which a fatigue crack can start. In the language of Section 4.3, corrosion is a mechanism that lowers the strength of a part over time while the load stays the same: the pipe wall that held the operating pressure with a wide margin when new holds it with a shrinking margin as it thins.

4.2.4 Thermal degradation

Elevated temperature accelerates chemical reactions, the aging of insulation, diffusion and other forms of material degradation. The rate of a thermally activated process increases exponentially with temperature, so that a moderate increase in operating temperature can shorten the life of an insulation system or a lubricant several fold; Section 4.3.4 quantifies this with the Arrhenius model.

Temperature can also produce immediate failure, through overheating beyond a material limit, through thermal expansion that closes a clearance or opens a joint, or through the thermal stress of a temperature gradient. Repeated temperature changes add thermal fatigue, the fatigue of Section 4.2.1 driven by thermal rather than mechanical stress. A component can therefore be affected both by long-term thermal aging and by short-term temperature excursions, and the two act on different time scales: aging consumes life slowly and invisibly, an excursion can end it at once.

4.2.5 Electrical degradation and failure

Electrical failure mechanisms include insulation breakdown, dielectric aging, electrical arcing and semiconductor damage. An insulation breaks down when the electric field across it exceeds its dielectric strength; dielectric aging, driven by heat, partial discharges and moisture, lowers that strength over time; an arc forms across a gap whose insulation has failed and erodes the contacts it bridges; and semiconductor devices are damaged by overvoltage, by electrostatic discharge, or by the slow migration of metal atoms in their interconnects under high current density (electromigration).

Some electrical failures occur almost instantaneously, others develop through cumulative degradation. The distinction is general and has standard names: an *overstress* mechanism

produces failure when a single load exceeds the strength, as in an insulation puncture by a voltage surge; a *wear-out* mechanism produces failure when accumulated damage reaches a critical level, as in dielectric aging or electromigration. The two kinds are modeled differently, by the stress-strength interference model and by the accumulated degradation model of Section 4.3, and they shape the hazard function differently, as a later chapter on lifetime distributions shows.

4.2.6 Software, control and operational failures

Not every system failure results from a degrading physical component. A function can be lost through

- incorrect control logic;
- erroneous sensor measurements;
- software configuration errors;
- incorrect protection settings;
- operator actions;
- unanticipated combinations of operating conditions.

These events remove an engineering function even though the hardware remains physically intact. A relay with an incorrect setting trips a healthy transmission line; a biased flow sensor leads a controller to deliver less water than the plant needs (Section 4.5.7); a chiller sequencing program that never anticipated two simultaneous alarms leaves the plant with one chiller running. None of these has a failure mechanism in the physical sense, but each has a failure mode and an effect, and each belongs in a failure mode analysis.

The distinction becomes especially important in automated power networks and data centers, where much of the required function is delivered by control and protection systems acting on measurements, and where a measurement error can disable a function as completely as a broken pump. Table 4.2 collects the classes of this section. For each it lists representative mechanisms, the stresses that drive them, the typical failure modes they produce, whether they act by overstress or by wear-out, and the model of this chapter that represents them.

4.3 From physical mechanisms to mathematical failure models

A mechanism becomes useful for reliability once it is written as a mathematical model that relates the stresses an item experiences to the time or the probability of its failure. This section develops three common models: stress-strength interference for overstress mechanisms, accumulated degradation for wear-out mechanisms, and the Arrhenius model for the effect of temperature on the rate of degradation.

4.3.1 Stress-strength interference

Let an item have a strength C, its resistance to a particular kind of load, and let the applied load or stress be L, both in the same units (kilonewtons for a structural member, megapascals for a pressure boundary, kilovolts for an insulation). The strength is written C because it is the item's capacity in the sense of Chapter 2; it is often written S in the stress-strength literature, but S denotes the severity rating in this book. The item fails when the load exceeds the strength,

$$L > C. \tag{4.1}$$

The performance margin is

$$g = C - L, \tag{4.2}$$

success requires $g \geq 0$ and failure occurs when $g < 0$.

Table 4.2: Classes of failure mechanisms, their driving stresses, typical failure modes, time behavior and the model of this chapter that represents them.

Class	Mechanisms	Driving stresses	Typical failure modes	Time behavior	Model
Mechanical fatigue	crack initiation and growth	stress range, mean stress, number of cycles	fracture of shaft, blade, weld	wear-out	S-N curve and Miner rule; Paris law
Wear	adhesive, abrasive, erosive, fretting	contact load, sliding distance, particles	reduced flow and efficiency, seizure	wear-out	accumulated degradation
Corrosion	uniform, pitting, galvanic, assisted cracking	moisture, chemistry, tensile stress	leak, wall rupture, high-resistance joint	wear-out	degradation; strength falling over time
Thermal	aging, diffusion, expansion, thermal stress	temperature, temperature cycles	insulation failure, seizure, cracking	both	Arrhenius; stress-strength
Electrical	breakdown, dielectric aging, arcing, electromigration	voltage, current density, temperature	short circuit, open circuit, trip	both	stress-strength; Arrhenius
Software, control, operation	logic, data, settings, actions	operating conditions	wrong command, spurious trip, function not demanded	either	FMEA of sensing and control

If both strength and load are uncertain, they are random variables, and the probability of failure, written p_f, is

$$p_f = \Pr(L > C) = \Pr(g < 0). \tag{4.3}$$

This is the stress-strength interference model. The name comes from the picture of two probability densities, one for the load and one for the strength, that interfere where they overlap. When C and L are independent with densities f_C and f_L and cumulative distribution functions F_C and F_L, conditioning on the strength gives the general form: an item of strength c fails if the load exceeds c, which has probability $1 - F_L(c)$, and averaging over the strength,

$$p_f = \int_{-\infty}^{\infty} f_C(c)\,[1 - F_L(c)]\,\mathrm{d}c. \tag{4.4}$$

The model applies wherever a demand is compared with a capacity: structural loading, pressure vessels, the current a transmission line may carry against its thermal limit, the heat a cooling plant must remove against its available capacity, and many other engineering requirements.

4.3.2 A worked stress-strength example

Let the strength and the maximum load of a component be independent normal random variables,

$$C \sim \mathcal{N}(120,\, 10^2)\,\mathrm{kN}, \qquad L \sim \mathcal{N}(90,\, 12^2)\,\mathrm{kN},$$

that is, a mean strength of 120 kN with a standard deviation of 10 kN, and a mean load of 90 kN with a standard deviation of 12 kN (illustrative values). The difference of two independent normal variables is normal, with mean equal to the difference of the means and variance equal to the sum of the variances, so the margin is $g \sim \mathcal{N}(\mu_g, \sigma_g^2)$ with

$$\mu_g = \mu_C - \mu_L = 120 - 90 = 30\,\mathrm{kN},$$
$$\sigma_g = \sqrt{\sigma_C^2 + \sigma_L^2} = \sqrt{10^2 + 12^2} = \sqrt{244} = 15.62\,\mathrm{kN}.$$

Standardizing the margin, $(g - \mu_g)/\sigma_g$ is a standard normal variable, and

$$p_f = \Pr(g < 0) = \Pr\left(\frac{g - \mu_g}{\sigma_g} < -\frac{\mu_g}{\sigma_g}\right) = \Phi\left(-\frac{\mu_g}{\sigma_g}\right) = \Phi\left(-\frac{30}{15.62}\right) = \Phi(-1.921) = 0.0274, \tag{4.5}$$

where Φ is the standard normal cumulative distribution function. The component has a probability of about 2.74 % of meeting a load that exceeds its strength under this random loading model. The ratio $\mu_g/\sigma_g = 1.921$, the mean margin measured in standard deviations of the margin, is called the reliability index.

Figure 4.2 shows the two views of this example. In panel (a) the horizontal axis is load or strength in kilonewtons and the vertical axis is probability density; the gray curve is the density of the load, centered at 90 kN, the blue curve the density of the strength, centered at 120 kN, and the shaded region is the area under both curves. In panel (b) the horizontal axis is the margin g and the vertical axis its density; the red area to the left of zero is the probability of failure. The two areas are different numbers: the overlap area in panel (a) is 0.172, more than six times the probability of failure. The overlap shows only that some loads exceed some strengths. The probability of failure is the probability that a paired realization, one load meeting one strength, has $L > C$, and only the distribution of the margin in panel (b), or the integral of equation (4.4), gives it.

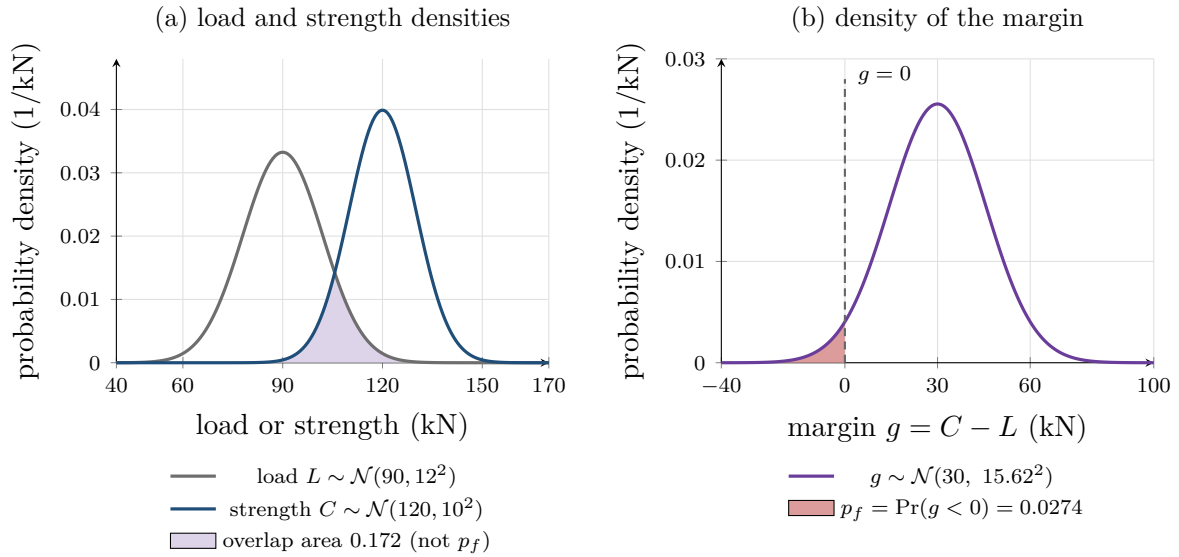


Figure 4.2: Stress-strength interference, illustrative. (a) Normal densities of the load, $\mathcal{N}(90, 12^2)$ kN, and of the strength, $\mathcal{N}(120, 10^2)$ kN; the shaded overlap has area 0.172. (b) Density of the margin $g = C - L \sim \mathcal{N}(30, 244)$ kN; the shaded tail is the probability of failure, $\Phi(-30/15.62) = 0.0274$. The probability of failure is the probability that a paired realization has $L > C$, not the overlap area of the two densities.

Common mistake

Reading the probability of failure as the area where the load and strength densities overlap. The overlap area of Figure 4.2(a) is 0.172; the probability of failure is 0.0274. The correct quantity is $\Pr(C - L < 0)$, computed from the distribution of the margin or from equation (4.4).

Repeated loading. The model as stated covers one load application. For a repeated-loading problem, the stress history, a strength that may evolve over time, and the probability that the

limit is crossed at any point during the mission must be accounted for. Even the simplest case shows that the extension is not obvious. Let the item meet n independent loads with the same distribution, while its strength, unknown but fixed, does not change. An item of strength c survives all n loads if the largest of them is below c, which has probability $F_L(c)^n$, so

$$p_f(n) = 1 - \int_{-\infty}^{\infty} f_C(c)\, F_L(c)^n \,\mathrm{d}c. \tag{4.6}$$

For the example, numerical integration gives $p_f(10) = 0.172$ and $p_f(100) = 0.499$. Treating the ten applications as independent trials with probability 0.0274 each would give $1-(1-0.0274)^{10} = 0.243$, and for a hundred, 0.938. The independent-trial formula is wrong here because the strength is the same in every trial: a strong item survives repeated loads that would fail a weak one, and the failures concentrate on the weak tail of the strength distribution.

4.3.3 Accumulated degradation

Wear-out mechanisms fail an item because damage accumulates over time. Let $D(t)$ be the accumulated damage at time t, a dimensionless quantity normalized so that failure occurs when it reaches a critical threshold,

$$D(t) \geq D_{\text{crit}}. \tag{4.7}$$

(D for damage is the standard symbol; it is unrelated to the detection rating D of Section 4.6.) A simple model of the damage rate is

$$\frac{\mathrm{d}D}{\mathrm{d}t} = k\, s(t)^n, \tag{4.8}$$

where k is a degradation coefficient, $s(t) \geq 0$ is a measure of the intensity of the stress, and n is the stress sensitivity exponent: doubling the stress multiplies the rate of damage by 2^n. Integrating from an initial damage $D_0 = D(0)$,

$$D(t) = D_0 + \int_0^t k\, s(u)^n \,\mathrm{d}u, \tag{4.9}$$

and under a constant stress s the damage grows linearly, $D(t) = D_0 + ks^n t$. Setting $D(T) = D_{\text{crit}}$ and solving for the time to failure, assuming $D_0 < D_{\text{crit}}$,

$$T = \frac{D_{\text{crit}} - D_0}{k\, s^n}. \tag{4.10}$$

For a normalized damage model with $D_0 = 0.10$, $D_{\text{crit}} = 1$ and, under the specified operating condition, $ks^n = 0.001\,\mathrm{h}^{-1}$ (illustrative values),

$$T = \frac{1 - 0.10}{0.001\,\mathrm{h}^{-1}} = 900\,\mathrm{h}.$$

This is a deterministic prediction under the assumed parameters and loading. It is not yet a reliability prediction. A probability of failure requires uncertainty in the degradation parameters, the future loading, the initial damage or the failure threshold. The cumulative failure probability over a horizon H is

$$F(H) = \Pr(T \leq H) = \Pr\big(D(H) \geq D_{\text{crit}}\big), \tag{4.11}$$

and the second equality holds because the damage of equation (4.8) never decreases: the threshold has been reached by H if and only if the damage at H is at or above it. Suppose, for illustration, that the rate $r = ks^n$ varies from unit to unit, lognormally with a median

of $0.001\,\mathrm{h}^{-1}$ and a standard deviation of $\ln r$ of 0.2. Then $\ln T = \ln 0.9 - \ln r$ is normal with median T of 900 h and the same standard deviation, and

$$F(H) = \Phi\left(\frac{\ln(H/900\,\mathrm{h})}{0.2}\right),$$

which gives $F(750\,\mathrm{h}) = 0.181$, $F(900\,\mathrm{h}) = 0.500$, $F(1000\,\mathrm{h}) = 0.701$ and $F(1125\,\mathrm{h}) = 0.868$.

Figure 4.3 draws the model. The horizontal axis is operating time in hours and the vertical axis the normalized damage. Each line is one unit with a constant damage rate: the median unit at $0.001\,\mathrm{h}^{-1}$, and two units at 0.0012 and $0.0008\,\mathrm{h}^{-1}$, the rates one standard deviation of $\ln r$ above and below the median (approximately $e^{\pm 0.2} = 1.22$ and 0.82 times the median; the figure rounds them to 1.2 and 0.8). The dashed red line is the threshold $D_{\mathrm{crit}} = 1$. All three lines start at $D_0 = 0.10$ and cross the threshold at 750, 900 and 1125 h: the spread of the rate becomes a spread of failure times, which is the distribution F above. The distinction between predicting a failure time and predicting a probability distribution of failure times is central to the later chapters on degradation and prognostics.

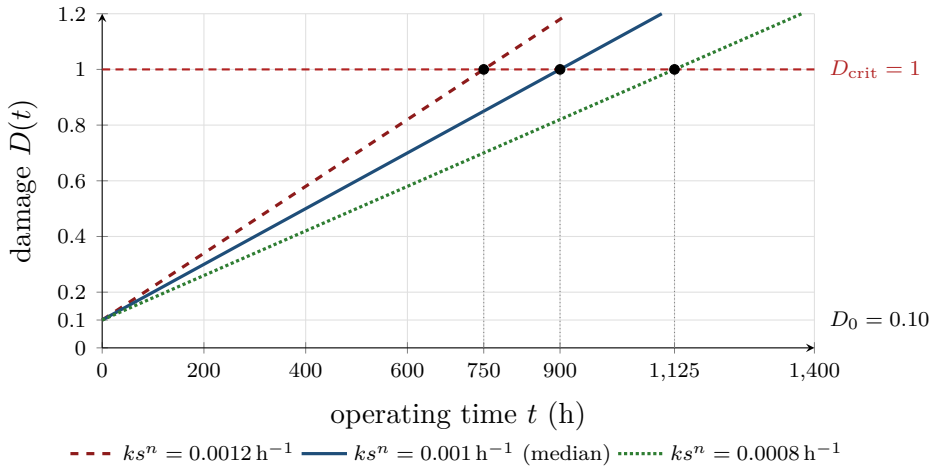


Figure 4.3: Accumulated degradation under constant stress, $D(t) = D_0 + ks^n t$ with $D_0 = 0.10$ and three illustrative damage rates. Each line reaches the threshold $D_{\mathrm{crit}} = 1$ at $T = (D_{\mathrm{crit}} - D_0)/(ks^n)$: 750, 900 and 1125 h. Uncertainty in the rate becomes a distribution of the time to failure.

Principle 4.2. *A deterministic time to failure, computed from assumed parameters and loads, is not a reliability prediction. It becomes one when the uncertainty in the parameters, the loads, the initial damage and the threshold is carried through to a distribution of the time to failure.*

4.3.4 Temperature-accelerated degradation

For thermally activated mechanisms, such as chemical reactions, diffusion and material migration, the rate of degradation depends on temperature through the Arrhenius model, a standard result of physical chemistry,

$$k(T) = k_0 \exp\left(-\frac{E_a}{k_B T}\right), \tag{4.12}$$

where $k(T)$ is the degradation or reaction rate, k_0 the pre-exponential factor (in the units of the rate), E_a the activation energy of the mechanism in electronvolts, $k_B = 8.61733 \times 10^{-5}\,\mathrm{eV/K}$ the Boltzmann constant, and T the absolute temperature in kelvin. In this subsection T denotes

temperature, not the time to failure. The pre-exponential factor is rarely known, but it cancels in the ratio of the rates at two temperatures T_1 and T_2:

$$\frac{k(T_2)}{k(T_1)} = \frac{\exp(-E_a/(k_B T_2))}{\exp(-E_a/(k_B T_1))} = \exp\left[\frac{E_a}{k_B}\left(\frac{1}{T_1} - \frac{1}{T_2}\right)\right]. \tag{4.13}$$

The ratio is the acceleration factor, written AF, of the higher temperature relative to the lower; it is the basis of accelerated life testing, in which items are aged at high temperature and the result is translated to the service temperature.

Insulation aging. Take an electronic insulation with $E_a = 0.70\,\mathrm{eV}$ (illustrative) whose operating temperature rises from 40°C to 60°C, that is, from $T_1 = 313.15\,\mathrm{K}$ to $T_2 = 333.15\,\mathrm{K}$. Then

$$\begin{aligned} \frac{E_a}{k_B} &= \frac{0.70\,\mathrm{eV}}{8.61733 \times 10^{-5}\,\mathrm{eV/K}} = 8123\,\mathrm{K}, \\ \frac{1}{T_1} - \frac{1}{T_2} &= \frac{1}{313.15} - \frac{1}{333.15} = \frac{20}{313.15 \times 333.15} = 1.9171 \times 10^{-4}\,\mathrm{K^{-1}}, \\ \mathrm{AF} &= \exp(8123 \times 1.9171 \times 10^{-4}) = \exp(1.5573) = 4.75. \end{aligned}$$

Under this model the insulation degrades 4.75 times faster at 60°C than at 40°C. If the degradation of Section 4.3.3 were this mechanism, with its 900 h computed at 40°C, the same damage would be reached in $900/4.75 = 190\,\mathrm{h}$ at 60°C.

That does not mean that the item's overall failure probability is 4.75 times greater. The relation between the degradation rate and the probability of failure also depends on the damage threshold, the exposure duration, competing failure mechanisms and the uncertainties of Section 4.3.3. The Arrhenius model is an important accelerated-life model for diffusion, chemical reactions and material migration, but it should not be applied indiscriminately: a mechanism driven by mechanical cycles, by humidity or by voltage needs a model of its own stress.

Figure 4.4 shows how strongly the result depends on the activation energy. The horizontal axis is the operating temperature in degrees Celsius and the vertical axis, on a logarithmic scale, the acceleration factor relative to 40°C, from equation (4.13). The three curves are activation energies of 0.5, 0.7 and 1.0 eV, and all pass through one at 40°C by construction. At 60°C they give 3.04, 4.75 and 9.25; at 80°C, 8.16, 18.9 and 66.5. For 0.7 eV near 40°C, each 10 K of temperature roughly doubles the rate (2.23 at 50°C), the origin of the rule of thumb that a 10 K rise halves the life of an insulation; the rule holds only for activation energies near that value.

4.4 Mechanical fatigue as physics-based failure prediction

Mechanical fatigue demonstrates how a failure mechanism can be modeled before a component actually fails. This section states the two standard models of fatigue life, the stress-life curve with the Miner rule for damage accumulation and the Paris law for crack growth, works each with units, and shows how each becomes a probability of failure.

4.4.1 The stress-life curve and the Miner rule

The stress-life or S-N curve of a material or a structural detail gives the number of constant-amplitude cycles N to failure as a function of the stress range $\Delta\sigma$, the difference between the largest and the smallest stress of a cycle. It is measured by testing specimens to failure at several stress ranges, and over the range of interest it is usually a straight line on logarithmic axes,

$$N(\Delta\sigma) = K\,\Delta\sigma^{-b} \qquad \Longleftrightarrow \qquad \log N = \log K - b \log \Delta\sigma, \tag{4.14}$$

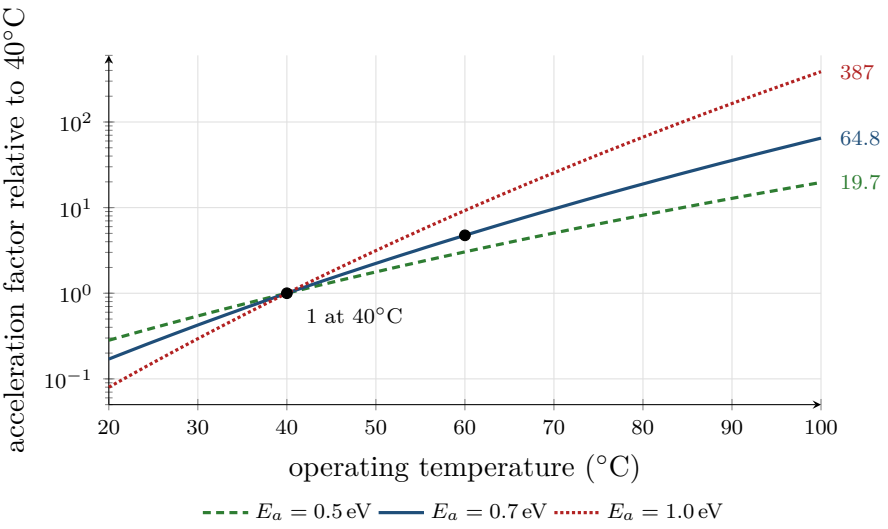


Figure 4.4: Arrhenius acceleration factor relative to 40°C, $\exp[(E_a/k_B)(1/313.15 - 1/T)]$ with T in kelvin, for three illustrative activation energies. The marked points are the reference, 1 at 40°C, and the worked example, 4.75 at 60°C for $E_a = 0.70$ eV. The values at 100°C are printed at the right edge.

where b is the S-N exponent, the inverse slope of the line, and K is a constant with units of MPa^b times cycles. A detail with $b = 3$ and $K = 2 \times 10^{12}\,\mathrm{MPa}^3$, illustrative values of the order used for welded steel, lasts

$$N(100\,\mathrm{MPa}) = \frac{2 \times 10^{12}}{100^3} = 2.0 \times 10^6 \text{ cycles},$$

3.91×10^6 cycles at 80 MPa and 5.93×10^5 cycles at 150 MPa. With $b = 3$, a 50 % increase of the stress range divides the life by $1.5^3 = 3.4$.

Real loads are not of constant amplitude. The Miner rule, the standard linear damage accumulation rule, assigns to each cycle at stress range $\Delta\sigma_i$ the damage $1/N(\Delta\sigma_i)$, the fraction of the life at that range that one cycle consumes. After n_i cycles at each of several ranges, the accumulated damage is

$$D = \sum_i \frac{n_i}{N(\Delta\sigma_i)}, \tag{4.15}$$

and failure is predicted when D reaches 1. The Miner rule is the accumulated degradation model of Section 4.3.3 with cycles in place of hours: per cycle, $\mathrm{d}D/\mathrm{d}N = \Delta\sigma^b/K$, which is equation (4.8) with $s = \Delta\sigma$, $n = b$, $k = 1/K$, $D_0 = 0$ and $D_{\mathrm{crit}} = 1$.

Consider a welded joint on the discharge piping of the cooling-water pump, loaded by the pressure changes of normal operation. Its illustrative load spectrum per year has three blocks: 2×10^5 cycles at 80 MPa from flow-control movements, 5×10^4 cycles at 100 MPa from pressure surges, and 2000 cycles at 150 MPa from pump starts and stops (about five a day). With the S-N curve above, taken without a fatigue limit, the damage per year is

$$\begin{aligned} D_{\mathrm{yr}} &= \frac{2 \times 10^5}{3.906 \times 10^6} + \frac{5 \times 10^4}{2.000 \times 10^6} + \frac{2000}{5.926 \times 10^5} \\ &= 0.0512 + 0.0250 + 0.0034 = 0.0796 \text{ per year}, \end{aligned}$$

and the predicted life is $1/0.0796 = 12.6$ yr. The 2000 start-stop cycles, at the highest stress, contribute only 4.2 % of the damage; the 2×10^5 small cycles contribute 64.3 %. The rule is linear and ignores the sequence of the cycles, although the sequence matters in practice (a

few large cycles early can accelerate the damage of later small ones), which is one reason why measured damage sums at failure scatter widely around one.

Figure 4.5 draws the example. Both axes are logarithmic; the horizontal axis is the number of cycles and the vertical axis the stress range in megapascals. The straight line is the S-N curve of equation (4.14). The filled dots on it are the lives $N(\Delta\sigma_i)$ at the three stress ranges of the spectrum; the open circles on the same rows are the cycles applied per year, n_i. On a logarithmic axis the horizontal distance between an open circle and its filled dot is $\log(N_i/n_i)$, so the closer the two marks, the larger the share of life consumed each year; the 80 MPa block is the closest, at a factor of 19.5.

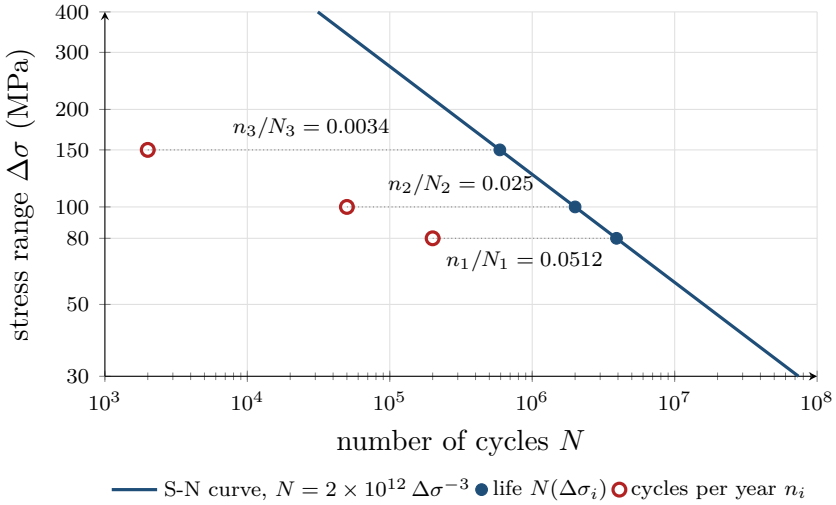


Figure 4.5: Illustrative S-N curve $N = K\Delta\sigma^{-b}$ with $K = 2 \times 10^{12}\,\mathrm{MPa}^3$ and $b = 3$, and the yearly load spectrum of the pump's discharge weld. Filled dots: life at each stress range; open circles: cycles applied per year. The Miner sum is $0.0512 + 0.0250 + 0.0034 = 0.0796$ per year, a predicted life of 12.6 years.

4.4.2 Crack growth and the Paris law

The S-N curve treats the whole life as a black box. Fracture mechanics opens it by following a crack that already exists. The classic crack-growth model is the Paris law,

$$\frac{\mathrm{d}a}{\mathrm{d}N} = C_{\mathrm{P}}\,(\Delta K)^m, \tag{4.16}$$

where a is the crack length, N the number of load cycles, $\mathrm{d}a/\mathrm{d}N$ the crack growth per cycle, C_{P} and m are material and environmental parameters determined by experiment, and ΔK is the range of the stress intensity factor, the quantity of fracture mechanics that measures the intensity of the stress field at the crack tip. For a simplified geometry,

$$\Delta K = Y\,\Delta\sigma\sqrt{\pi a}, \tag{4.17}$$

with Y a dimensionless geometry correction factor and $\Delta\sigma$ the cyclic stress range; ΔK is in $\mathrm{MPa}\sqrt{\mathrm{m}}$ when $\Delta\sigma$ is in megapascals and a in meters. Substituting,

$$\frac{\mathrm{d}a}{\mathrm{d}N} = C_{\mathrm{P}}\big(Y\,\Delta\sigma\sqrt{\pi a}\big)^m, \tag{4.18}$$

which connects the mechanical loading to the evolution of the crack length. The growth rate increases with the crack length as $a^{m/2}$, so the crack accelerates as it grows.

If the crack starts at a_0 and failure occurs at a critical length a_c, separating the variables, $\mathrm{d}N = \mathrm{d}a/[C_{\mathrm{P}}(Y\Delta\sigma\sqrt{\pi a})^m]$, and integrating gives the number of cycles to failure,

$$N_f = \int_{a_0}^{a_c} \frac{\mathrm{d}a}{C_{\mathrm{P}}\left(Y\Delta\sigma\sqrt{\pi a}\right)^m}. \tag{4.19}$$

With Y constant the integral has a closed form. The integrand is $a^{-m/2}$ times a constant, whose antiderivative for $m \neq 2$ is $a^{1-m/2}/(1-m/2)$, so

$$N_f = \frac{a_0^{\,1-m/2} - a_c^{\,1-m/2}}{(m/2-1)\, C_{\mathrm{P}}\left(Y\Delta\sigma\sqrt{\pi}\right)^m}, \qquad m \neq 2. \tag{4.20}$$

The result estimates the number of cycles for the crack to grow from a_0 to a_c, under the assumptions of the model. The Paris law is intended for the intermediate regime of fatigue crack growth; growth near the threshold below which cracks do not propagate, and the rapid fracture when the crack approaches its critical size, require additional treatment.

Worked example. Take illustrative values of the order of those for a structural steel: $C_{\mathrm{P}} = 10^{-11}$ meters per cycle per $(\mathrm{MPa}\sqrt{\mathrm{m}})^3$, $m = 3$, $Y = 1.12$, $\Delta\sigma = 100\,\mathrm{MPa}$, an initial crack $a_0 = 1\,\mathrm{mm}$ and a critical length $a_c = 10\,\mathrm{mm}$. Then

$$\begin{aligned}
Y\Delta\sigma\sqrt{\pi} &= 1.12 \times 100 \times 1.7725 = 198.5\,\mathrm{MPa},\\
C_{\mathrm{P}}\left(Y\Delta\sigma\sqrt{\pi}\right)^3 &= 10^{-11} \times 7.823 \times 10^6 = 7.823 \times 10^{-5}\,\mathrm{m}^{-1/2} \text{ per cycle},\\
a_0^{-1/2} - a_c^{-1/2} &= (0.001\,\mathrm{m})^{-1/2} - (0.010\,\mathrm{m})^{-1/2} = 31.62 - 10.00 = 21.62\,\mathrm{m}^{-1/2},\\
N_f &= \frac{21.62}{0.5 \times 7.823 \times 10^{-5}} = 5.53 \times 10^5 \text{ cycles}.
\end{aligned}$$

The stress intensity range grows from $198.5\sqrt{0.001} = 6.3\,\mathrm{MPa}\sqrt{\mathrm{m}}$ at a_0 to $19.9\,\mathrm{MPa}\sqrt{\mathrm{m}}$ at a_c, and the growth per cycle from $2.5 \times 10^{-9}\,\mathrm{m}$ to $7.8 \times 10^{-8}\,\mathrm{m}$, $10^{3/2} = 31.6$ times faster. Three sensitivities follow from equation (4.20). The life scales as $\Delta\sigma^{-m}$: 1.08×10^6 cycles at 80 MPa, 3.20×10^5 at 120 MPa, the same inverse-cube law as the S-N curve of the previous subsection, which is why welded details, whose life is mostly crack growth, have S-N exponents near three. Halving the initial crack to 0.5 mm raises the life by 61 %, to 8.88×10^5 cycles, because most of the cycles are spent while the crack is small. Doubling the critical length to 20 mm raises it by only 13.5 %, to 6.28×10^5 cycles.

Figure 4.6 plots the crack length against the number of cycles for three stress ranges, from the integrated form $a(N) = \left[a_0^{-1/2} - \tfrac{1}{2}C_{\mathrm{P}}(Y\Delta\sigma\sqrt{\pi})^3 N\right]^{-2}$. The horizontal axis is cycles in units of 10^5 and the vertical axis the crack length in millimeters; the dashed line is the critical length. Each curve stays nearly flat for most of its life and turns sharply upward at the end, and the three curves differ only by a stretching of the horizontal axis in proportion to $\Delta\sigma^{-3}$. At 100 MPa the crack takes 2.37×10^5 cycles, 43 % of the life, to grow from 1 to 2 mm, and 81 % of the life to reach 5 mm, half the critical length; the last 5 mm take only 19 %. An inspection interval must therefore be short compared with the life, or a crack found small at one inspection can be critical before the next.

4.4.3 From a deterministic life to a probability of failure

The physics-based model matters for reliability engineering because of what it adds to statistics. Suppose two otherwise identical components operate under different cyclic loading. Traditional empirical reliability analysis would estimate a separate lifetime distribution for each from its own history of failures, and would need failures at both loadings to do so. A physics-based approach connects the difference in loading to a difference in predicted crack-growth rates: by

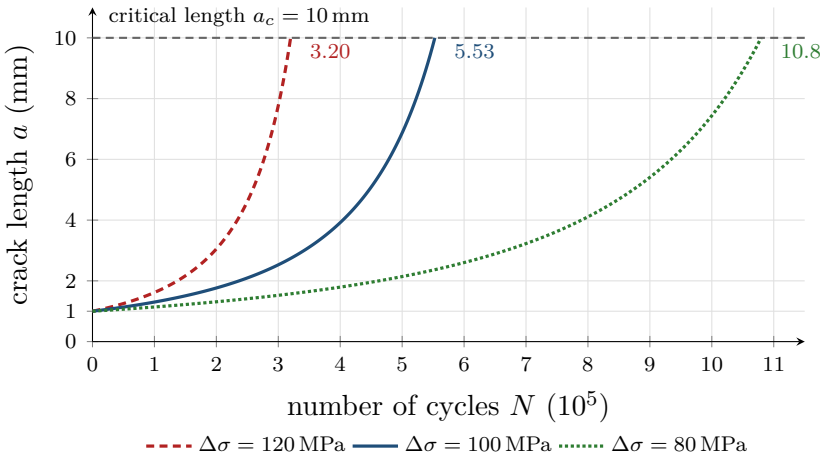


Figure 4.6: Fatigue crack growth by the Paris law with illustrative steel parameters $C_\text{P} = 10^{-11}$, $m = 3$, $Y = 1.12$, from $a_0 = 1\,\text{mm}$ to $a_c = 10\,\text{mm}$. The lives (3.20, 5.53 and 10.8×10^5 cycles, printed at the end of each curve) scale as $\Delta\sigma^{-3}$; most of the life is spent while the crack is small.

equation (4.20), a component loaded at 120 instead of 100 MPa lives $(100/120)^3 = 0.58$ times as many cycles, whether or not either has yet failed in service.

To make the approach probabilistic, the parameters C_P, m, a_0 and $\Delta\sigma$ are represented as uncertain quantities, with distributions estimated from material tests, inspection records of initial defects and measured loads. The number of cycles to failure N_f computed from equation (4.19) is then a random variable, and the quantity of interest is its cumulative distribution,

$$\Pr(N_f \le N), \tag{4.21}$$

the probability that the crack reaches its critical length within N cycles, the fatigue form of equation (4.11). The sensitivities of the worked example show where the uncertainty matters most: the initial crack size and the stress range, rather than the critical length. This creates a direct connection between physical degradation models and probabilistic reliability, and it is the foundation of many methods for estimating remaining useful life, in which the distribution of N_f is updated as inspections measure the current crack length.

> **Principle 4.3.** *A physics-of-failure model makes the probability of failure a function of the stresses an item experiences. Uncertain parameters, initial defects and loads turn the model's deterministic life into a distribution of the time or the number of cycles to failure.*

4.5 Failure mode and effects analysis

The analysis now moves from predicting particular physical mechanisms to identifying systematically how a complete system can fail. This section defines FMEA, applies its four steps to the cooling-water pump system, records the result in an FMEA worksheet, and closes with the failure mode of a sensor, which shows why sensing and control belong in the analysis.

4.5.1 Purpose and system definition

FMEA is a structured, bottom-up analysis that examines the potential failure modes of the elements of a system and traces their effects on the required system function. It starts from

the system functions, identifies the functional failures, the dominant failure modes that can produce them, their causes and their consequences, and records each link of that chain. It is inductive: it reasons from a supposed failure of a part to its consequences.

The example is the industrial cooling-water system that recurs in this book. It contains

- one centrifugal pump;
- one electric motor;
- a flow-control valve;
- supply and return piping;
- a flow sensor;
- a controller.

Its required function is to deliver at least $100\,\mathrm{m}^3/\mathrm{h}$ of cooling water while maintaining the specified operating pressure. The system is functionally failed when it cannot satisfy those requirements, whatever the state of its individual parts.

4.5.2 Step 1: functional failures

The first step lists the ways in which the required function can be lost or degraded. Table 4.3 gives five for the cooling-water system, each with an identifier for reference in the later steps. They range from the complete loss of the function (no flow), through its partial loss (insufficient flow) and the violation of a secondary requirement (excessive discharge pressure), to the loss of control of the function (unstable flow) and an unwanted side effect (external leakage beyond the allowable limit).

Table 4.3: Functional failures of the cooling-water system.

ID	Functional failure
FF-01	No cooling-water flow
FF-02	Insufficient cooling-water flow
FF-03	Excessive discharge pressure
FF-04	Uncontrolled or unstable flow
FF-05	External water leakage beyond allowable limits

These are descriptions of inadequate performance, not explanations of why it happened. FF-02, for example, is stated against the requirement of $100\,\mathrm{m}^3/\mathrm{h}$ and is the same functional failure whether the flow is low because the impeller is worn, the valve is stuck or the controller asks for too little. Keeping the functional failures free of causes is what allows the next steps to find every cause of each.

4.5.3 Step 2: component failure modes

The second step identifies how each item could fail. Each failure mode is stated as what is observed at the item, and each can be traced to one or more of the functional failures of Table 4.3. Table 4.4 lists eight modes for the six items, with an illustrative tracing to the functional failures each can produce. A valve stuck closed stops the flow, and with the pump still running against it the discharge pressure rises to the pump's shutoff head, so it can produce FF-01 and FF-03 together; a flow sensor that reads high leads the controller to reduce the flow it delivers, producing FF-02 or, as the controller hunts, FF-04.

The list is deliberately not restricted to equipment that can break. Two of its eight modes, the sensor and the controller, are failures of measurement and logic, the class of Section 4.2.6, and two of them produce several functional failures at once.

Table 4.4: Component failure modes of the cooling-water system and the functional failures of Table 4.3 each can produce (the tracing is illustrative).

Component	Failure mode	Functional failures
Pump	Impeller cannot rotate	FF-01
Pump	Reduced hydraulic performance	FF-02
Motor	Fails to start	FF-01
Valve	Stuck closed	FF-01, FF-03
Valve	Stuck partially open	FF-02, FF-04
Flow sensor	Reports an erroneously high flow	FF-02, FF-04
Controller	Commands an incorrect operating point	FF-02, FF-03, FF-04
Piping	Ruptures or becomes obstructed	FF-01, FF-02, FF-05

4.5.4 Step 3: causes and physical mechanisms

The third step attaches causes and mechanisms to each mode. Consider the failure mode "pump cannot rotate". Its potential causes include bearing seizure, a jammed impeller and shaft damage. Bearing seizure itself may be associated with inadequate lubrication, contamination of the lubricant, overheating or fatigue damage of the rolling surfaces, and each of those is a mechanism of Section 4.2 with a model of Section 4.3 behind it. The step therefore descends the indenture levels of Section 4.1.2: a cause of the pump's mode is a mode of the bearing, whose own causes are lubrication and contamination.

The purpose of FMEA is not necessarily to resolve all such mechanisms quantitatively. It is to identify sufficiently credible causes and trace their effects, so that the engineering controls against each, such as a lubrication schedule, a filter, a temperature alarm or a vibration monitor, can be evaluated. Where a mode turns out to dominate, the mechanism behind it is the candidate for the quantitative models of the earlier sections.

4.5.5 Step 4: local and system effects

The fourth step traces each mode upward to its effects. The standard practice distinguishes the local effect, at the item itself; the next-higher-level effect, at the subsystem; and the end effect, at the level of the system whose required function is being analyzed. Figure 4.7 traces a valve stuck partially closed. The four boxes, from left to right, are the failure mode, its local effect (increased flow resistance), its subsystem effect (reduced cooling-water flow) and its potential facility-level effect (insufficient cooling and elevated equipment temperature). The note under the last arrow lists the conditions that decide whether the last step happens.

The word *potential* matters. Reduced flow does not necessarily produce a facility-level failure. Whether it does depends on the available redundancy, the cooling demand, the thermal storage and the response of the control system: a standby pump may start, the demand may be low, the water in the loop may absorb the deficit until the valve is freed. The local and subsystem effects are properties of the item; the end effect is a property of the item in its system.

Principle 4.4. *A component failure mode must be evaluated in the context of the system. Its local effect belongs to the item; its end effect depends on redundancy, demand, storage and control.*

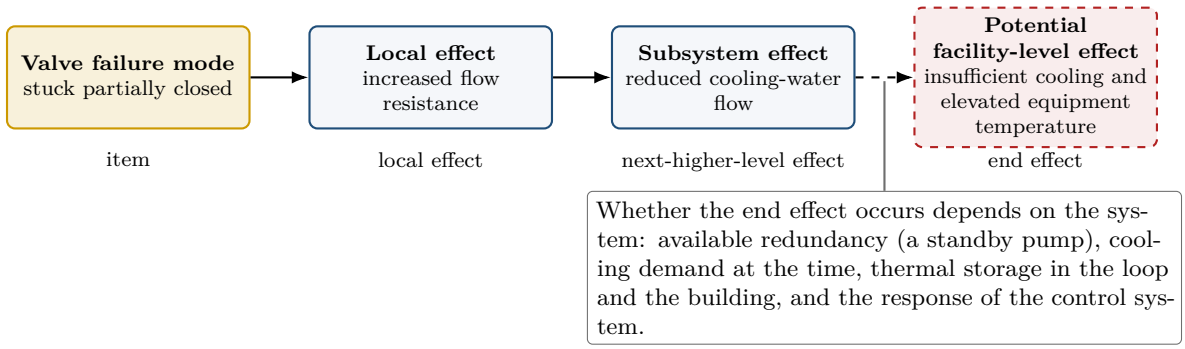


Figure 4.7: Effects of one failure mode traced through the levels of the system: a valve stuck partially closed raises the flow resistance, reduces the cooling-water flow, and can, depending on redundancy, demand, thermal storage and control, cause insufficient cooling at the facility (dashed: a potential effect).

4.5.6 The FMEA worksheet

An FMEA worksheet records the analysis in a traceable form, one row per failure mode. The exact columns vary among organizations and standards; Table 4.5 uses a useful minimum. The columns are an identifier, the item and its failure mode, a possible cause, the local effect, the system effect, and the means of detection or mitigation. Each row reads as one chain of Figure 4.1 with a response attached: the pump bearing that seizes from loss of lubrication stops the shaft, loses the cooling flow, and is met by vibration monitoring, which detects the developing mechanism, and by a standby pump, which mitigates the effect.

Table 4.5: An FMEA worksheet for the cooling-water system.

ID	Item and failure mode	Possible cause	Local effect	System effect	Detection / mitigation
F01	Pump bearing seizes	Lubrication loss, wear	Shaft stops	Cooling flow lost	Vibration monitoring, standby pump
F02	Valve stuck partially closed	Actuator malfunction	High resistance	Reduced flow	Flow monitoring, valve inspection
F03	Flow sensor reads high	Calibration drift	Incorrect flow indication	Controller may underdeliver flow	Redundant sensor, plausibility checks
F04	Motor fails to start	Electrical fault	Pump unavailable	Cooling flow lost on demand	Motor testing, electrical protection
F05	Pipe develops leak	Corrosion, fatigue	Water escapes	Reduced pressure, possible flooding	Leak detection, inspection

The worksheet does not simply list equipment that can break. It connects failure modes to functions and to engineering responses, and the last column distinguishes two kinds of response: detection, which finds the mode or its mechanism before the effect (vibration monitoring, plausibility checks, leak detection), and mitigation, which limits the effect once the mode has occurred (a standby pump, a redundant sensor, electrical protection). Row F04 records a mode that matters only on demand: a motor that cannot start is invisible while the pump is stopped, and is found by a test or by the demand itself.

4.5.7 Sensor bias

Row F03 is the most instructive. Let the required flow be $100\,\mathrm{m}^3/\mathrm{h}$, the true flow

$$Q_{\text{true}} = 85\,\mathrm{m}^3/\mathrm{h},$$

and the flow sensor report

$$Q_{\text{meas}} = 105\,\mathrm{m}^3/\mathrm{h},$$

a bias of $+20\,\mathrm{m}^3/\mathrm{h}$. If the controller relies on that sensor, it concludes that the flow requirement is satisfied with $5\,\mathrm{m}^3/\mathrm{h}$ to spare, while the plant is $15\,\mathrm{m}^3/\mathrm{h}$, or 15 %, short. The system then has two interacting problems:

1. the physical cooling requirement is violated;
2. the measurement system fails to reveal the violation.

The measurement error is not itself equivalent to insufficient flow. Rather, it may prevent the correction of an existing flow deficiency, as here, or cause an inappropriate control action, as when a controller throttles a healthy flow down to match a reading that is too high. The sensor's failure mode turns a detectable failure into a hidden one, which is why its detection rating in the next section is poor. An FMEA should therefore consider not only physical equipment but also sensing, software and controls, and the questions it raises, whether a measurement can be checked against others and whether the state of interest can be inferred from the measurements available, are those of the later chapters on observability and bad-data detection.

4.6 From FMEA to criticality and quantitative risk

FMEA identifies failure modes and their effects; deciding which deserve attention requires ranking them. This section describes the ordinal screening of FMECA and its limits, replaces the ordinal occurrence by failure probabilities computed from competing hazards, weights them by consequence, and places FMEA beside fault tree analysis.

4.6.1 FMECA and the risk priority number

FMECA extends FMEA by evaluating the relative criticality of each failure mode. A common introductory screening approach assigns each mode three ordinal scores:

- **severity** S: how serious the consequences are;
- **occurrence** O: how likely the failure mode is;
- **detection** D: how difficult the mode is to detect before the relevant effect occurs.

Each score runs from 1 to 10, with larger numbers indicating greater concern; a high detection score therefore means a mode that is hard to detect. One traditional screening index is the risk priority number,

$$\mathrm{RPN} = S \times O \times D, \tag{4.22}$$

which ranges from 1 to 1000. Table 4.6 gives illustrative ratings for four modes of the cooling-water system. Under this screening method, flow sensor bias receives the highest score, $9 \times 5 \times 8 = 360$, almost four times that of bearing seizure, because its severity is high (it hides a loss of cooling) and its detection is poor (row F03 of Table 4.5).

An RPN is not a numerical probability of failure and not an expected loss. The scores are ordinal: a severity of 8 is worse than a severity of 4 but not twice as bad, and an occurrence of 6 is not twice as likely as an occurrence of 3. Multiplying ordinal scores does not create a physically calibrated risk measure, and different combinations can give identical RPNs despite very different consequences. For serious reliability studies, a quantitative criticality analysis uses appropriately estimated event frequencies, failure probabilities, system effects and consequences, as the following subsections do.

Table 4.6: Illustrative criticality ratings and risk priority numbers for four failure modes of the cooling-water system.

Failure mode	S	O	D	RPN
Bearing seizure	8	4	3	96
Valve stuck closed	7	3	4	84
Flow sensor bias	9	5	8	360
Motor fails to start	8	3	2	48

4.6.2 Screening with the risk priority number

Table 4.7 shows how the RPN responds when the ratings change. Each row is one combination of severity, occurrence and detection, computed with equation (4.22); the comment says what the combination represents. The first row is the flow sensor bias as rated, $9 \times 5 \times 8 = 360$. A second, independent flow sensor with a plausibility check between the two improves detection to 3 and lowers the RPN to $9 \times 5 \times 3 = 135$; a shorter calibration interval lowers occurrence to 2 and the RPN to $9 \times 2 \times 8 = 144$. The fourth row is a mode of low severity that is frequent and nearly undetectable; its RPN, $4 \times 9 \times 10 = 360$, ties with the sensor bias. The fifth row is a rupture of the discharge manifold that floods the plant room, the most severe outcome on the scale, but rare and impossible to miss: $10 \times 2 \times 2 = 40$, lower than every mode of Table 4.6.

Table 4.7: Worked risk priority numbers, RPN $= S\,O\,D$, for the flow sensor bias as rated (9/5/8) and for four other combinations (illustrative).

Case	S	O	D	RPN	Comment
Flow sensor bias, as rated	9	5	8	360	baseline
Bias with a redundant sensor and plausibility check	9	5	3	135	better detection; severity unchanged
Bias with a shorter calibration interval	9	2	8	144	lower occurrence; still hard to detect
Minor, frequent, hard-to-detect mode	4	9	10	360	same RPN as the baseline, far milder consequence
Discharge manifold rupture, plant room flooded	10	2	2	40	highest severity, lowest RPN of all

The trend is the point of the table. Severity enters the product with the same weight as the two other ratings, so a catastrophic mode that is rare and obvious can rank below every nuisance on the list, and a screening rule of the form “act on every mode with an RPN above 100” would leave the manifold rupture unaddressed. Ties are common: over the 1000 combinations of three ratings from 1 to 10, the product takes only 120 distinct values, and 360 alone is produced by 15 combinations. Many organizations therefore review every mode of severity 9 or 10 regardless of its RPN, or rank by severity first. The useful outcome of the screening is not which score is largest in an illustrative example: a formal FMEA forces the analyst to document what can fail, what consequences follow, how the failure can be detected and where engineering action may be warranted, and the analysis is updated as the design, the processes, the operation and the knowledge of the system evolve.

Common mistake

Treating the RPN as a probability or an expected loss. The RPN is a product of ordinal scores; it carries no units, equal RPNs can describe very different risks, and a high-severity mode can have the lowest RPN on the list. Quantitative criticality requires probabilities and consequences.

4.6.3 Competing failure modes with constant hazards

The connection from FMEA to the mathematics of reliability is made by giving each failure mode a hazard. Let the cooling pump have three failure modes with constant cause-specific hazards (illustrative): bearing seizure, $\lambda_1 = 2 \times 10^{-4}\,\mathrm{h}^{-1}$; motor electrical failure, $\lambda_2 = 1 \times 10^{-4}\,\mathrm{h}^{-1}$; impeller failure, $\lambda_3 = 0.5 \times 10^{-4}\,\mathrm{h}^{-1}$. The modes are competing failure mechanisms: the first failure from any of them ends the pump's mission without failure. If the modes act independently, the pump survives to t only if none of them has occurred, so

$$R(t) = \prod_{i=1}^{3} e^{-\lambda_i t} = e^{-\lambda_{\mathrm{total}} t}, \qquad \lambda_{\mathrm{total}} = \lambda_1 + \lambda_2 + \lambda_3 = 3.5 \times 10^{-4}\,\mathrm{h}^{-1}, \tag{4.23}$$

the hazards of competing modes add. For a mission of 1000 h,

$$R(1000\,\mathrm{h}) = e^{-3.5\times10^{-4}\times1000} = e^{-0.35} = 0.7047, \qquad F(1000\,\mathrm{h}) = 1 - 0.7047 = 0.2953,$$

a probability of 29.53 % that the pump fails during the mission.

Which mechanism contributes most follows from the probability that the first failure occurs through mechanism i before the horizon H. The first failure is by mode i in the short interval $[t, t+\mathrm{d}t)$ if the pump has survived all modes to t, with probability $R(t)$, and mode i then occurs in the interval, with probability $\lambda_i\,\mathrm{d}t$. Integrating over the horizon,

$$F_i(H) = \int_0^H \lambda_i\, e^{-\lambda_{\mathrm{total}} t}\,\mathrm{d}t = \frac{\lambda_i}{\lambda_{\mathrm{total}}}\left(1 - e^{-\lambda_{\mathrm{total}} H}\right). \tag{4.24}$$

For constant hazards, the share of first failures due to mode i is $\lambda_i/\lambda_{\mathrm{total}}$ at every horizon, and the probabilities $F_i(H)$ add up to $F(H)$. For bearing seizure,

$$F_1(1000\,\mathrm{h}) = \frac{2}{3.5} \times 0.2953 = 0.5714 \times 0.2953 = 0.1687.$$

Table 4.8 lists the three modes. Each row gives the hazard, the share of first failures and the probability of a first failure by that mode within 1000 h; the last row is the total.

Table 4.8: Competing failure modes of the cooling pump with constant hazards (illustrative): share of first failures, $\lambda_i/\lambda_{\mathrm{total}}$, and probability of a first failure by each mode within 1000 h, equation (4.24).

Mechanism	Hazard (h^{-1})	Share of first failures	$F_i(1000\,\mathrm{h})$
Bearing seizure	2.0×10^{-4}	57.14 %	16.87 %
Motor failure	1.0×10^{-4}	28.57 %	8.44 %
Impeller failure	0.5×10^{-4}	14.29 %	4.22 %
Total	3.5×10^{-4}	100 %	29.53 %

This is considerably more informative than an ordinal occurrence score. The numbers, however, describe only how likely each component failure is; they do not yet say how much each mode contributes to system-level loss. An impeller failure may cause an immediate loss of cooling, while an early bearing defect may be detected by vibration monitoring and corrected before the required function is lost. Failure-mode analysis must be connected to the system architecture and to consequences.

4.6.4 Importance and failure frequency

Consider two failure modes. A pump bearing failure occurs frequently, but an automatic standby pump restores the flow. A failure of a common electrical controller occurs infrequently, but it disables both pumps at once. Which is more important depends on its effect on the required system function, not on its frequency. A simplified measure is the expected-loss contribution of failure event F_i,

$$\mathcal{R}_i = \Pr(F_i)\,\mathbb{E}[C \mid F_i], \tag{4.25}$$

where C is a defined loss and the events F_i are specified over a common period. If failure scenarios overlap, their contributions must be defined carefully to avoid counting the same loss twice; when the scenarios are mutually exclusive, the contributions add to the total risk $\mathcal{R} = \sum_i \mathcal{R}_i$, as in Chapter 3.

Table 4.9 takes two mutually exclusive system outcome scenarios over one year. Each row gives the annual probability, the consequence in reais and their product. The bearing-related service interruption has probability 0.10 and costs R$ 10 000, an expected loss of $0.10 \times$ R$ 10 000 = R$ 1 000 per year; the common controller outage has probability 0.01 and costs R$ 1 000 000, an expected loss of R$ 10 000 per year. The controller outage is ten times less probable and contributes ten times more expected loss: 91 % of the total of R$ 11 000 per year.

Table 4.9: Expected-loss contributions of two mutually exclusive system outcome scenarios over one year (illustrative).

Scenario	Annual probability	Consequence	Expected loss contribution
Bearing-related service interruption	0.10	R$ 10 000	R$ 1 000
Common controller outage	0.01	R$ 1 000 000	R$ 10 000
Total			R$ 11 000

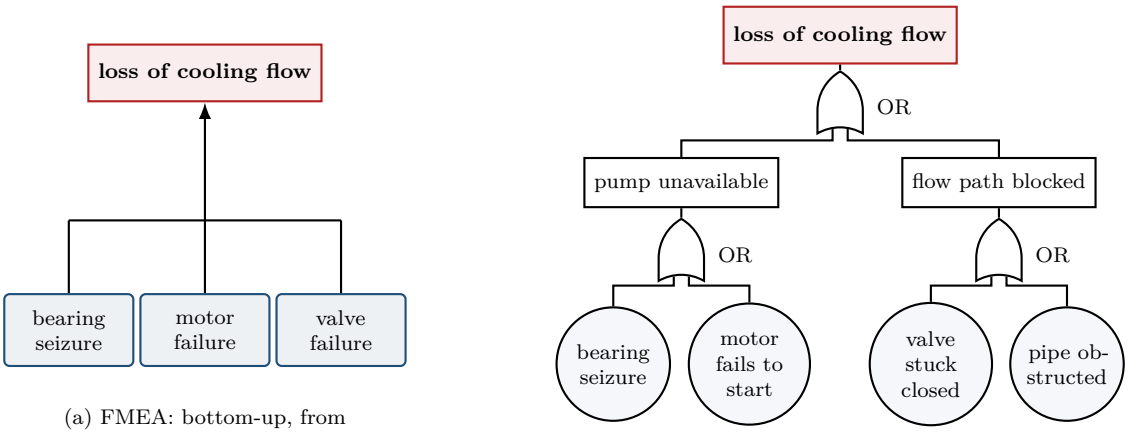
This is one reason why FMECA, fault trees, event trees and probabilistic risk assessment are complementary. FMEA identifies candidate failure pathways; quantitative system reliability analysis determines how those pathways combine, propagate and contribute to system failure. The ranking of components by their contribution to system unreliability or risk is made precise by the importance measures of a later chapter.

> **Principle 4.5.** *The most frequently failing component is not necessarily the largest contributor to system risk. A rare failure of a shared dependency that defeats redundancy can dominate the expected loss.*

4.6.5 FMEA and fault tree analysis

FMEA and fault tree analysis (FTA) approach the same engineering system from opposite directions. Figure 4.8 shows both, for the loss of cooling flow. Panel (a) is the FMEA view: three component failure modes at the bottom, bearing seizure, motor failure and valve failure, each traced upward to the system effect at the top. Panel (b) is a fault-tree fragment: the undesired top event at the top, an OR gate below it (the shield-shaped symbol, whose output occurs if any input occurs), two intermediate events, “pump unavailable” and “flow path blocked”, each the output of a further OR gate, and at the bottom the basic events (circles), the component failures at which the decomposition stops. In (a) the arrows run from the parts to the system, in the direction of the question “what happens if this item fails”; in (b) the analysis runs from the

top event down, in the direction of the question "what combinations of events can cause this". The two diagrams are simplified illustrations, not complete analyses of the same equipment configuration.



(a) FMEA: bottom-up, from component failure modes to effects

(b) fault tree: top-down, from the undesired event to its causes

Figure 4.8: The same undesired system event approached from two directions, simplified. (a) FMEA starts from component failure modes and traces their effects upward. (b) A fault-tree fragment starts from the top event and decomposes it through OR gates into intermediate events and basic events (circles).

Table 4.10 compares the two methods on five properties. FMEA is inductive and bottom-up: it starts from a component or a function and asks what happens if it fails, its emphasis is the identification of failures and their effects, and its output is the worksheet. FTA is deductive and top-down: it starts from an undesired system event and asks which combinations of events can cause it, its emphasis is the logic of failure and its probability, and its output is the fault tree and its minimal cut sets, the smallest combinations of basic events sufficient to cause the top event.

Table 4.10: FMEA and fault tree analysis compared.

Property	FMEA	Fault tree analysis
Direction	Bottom-up (inductive)	Top-down (deductive)
Starting point	Component or function	Undesired system event
Principal question	What happens if this item fails?	What combinations can cause this event?
Emphasis	Failure identification and effects	Failure logic and probability
Typical output	FMEA worksheet	Fault tree and minimal cut sets

The two are used together. The failure modes of an FMEA worksheet are the natural basic events of a fault tree, and the system effects of the worksheet suggest its top events. FMEA examines one failure at a time, so it is weak on combinations, such as the standby pump that fails to start while the duty pump is under repair, and on shared dependencies, such as the common controller of Section 4.6.4, whose single failure appears in every branch; the fault tree represents both explicitly. Fault trees and their minimal cut sets, and common-cause failure models, are developed in later chapters. The models of this chapter describe how and why items fail; the next chapter, Chapter 5, describes the probability of failure as a function of time through the lifetime density, the cumulative failure probability, the reliability function and the hazard function.

4.7 What to retain

Definition 4.1 (Failure mechanisms and failure mode analysis)**.** A failure is a chain from a cause through a physical mechanism to a failure mode and its effect, defined for a stated item and function; physics-of-failure models make its probability a function of stress, temperature and time, and FMEA identifies the modes and traces their effects, but ranking them for decisions requires probabilities and consequences at the system level, because the most frequent failure is not necessarily the most important.

Table 4.11 lists the chapter's terms with their meaning, their expression and units, and the subsection that defines each.

Table 4.11: Terms of failure mechanisms and failure mode analysis.

Term	Meaning	Expression, units	Defined in
Failure cause	what initiates a failure sequence	–	4.1.1
Failure mechanism	physical or chemical process leading to failure	–	4.1.1
Failure mode	how the item fails to perform its function	–	4.1.1
Failure effect	consequence at component, subsystem or system level	local, next-higher-level, end effect	4.1.1, 4.5.5
Indenture level	level of the hierarchy at which an item is analyzed	–	4.1.2
Overstress, wear-out	failure by one excessive load; by accumulated damage	–	4.2.5
Stress-strength interference	probability that load exceeds strength	$p_f = \Pr(C - L < 0)$	4.3.1
Reliability index	mean margin in standard deviations	μ_g/σ_g; $p_f = \Phi(-\mu_g/\sigma_g)$	4.3.2
Accumulated degradation	damage grows to a threshold	$\mathrm{d}D/\mathrm{d}t = ks^n$; $T = (D_{\mathrm{crit}} - D_0)/(ks^n)$, h	4.3.3
Arrhenius acceleration factor	rate ratio of a thermally activated mechanism	$\exp[(E_a/k_B)(1/T_1 - 1/T_2)]$	4.3.4
S-N curve, Miner rule	life against stress range; linear damage sum	$N = K\Delta\sigma^{-b}$; $D = \sum_i n_i/N_i$	4.4.1
Paris law	crack growth per cycle	$\mathrm{d}a/\mathrm{d}N = C_{\mathrm{P}}(\Delta K)^m$, m/cycle	4.4.2
Physics-of-failure model	relates stress, time and degradation to failure	$\Pr(N_f \le N)$, $F(H)$	4.4.3
Functional failure	inadequate performance of a function	FF-01 to FF-05	4.5.2
FMEA	systematic identification of failure modes and effects	worksheet	4.5.1, 4.5.6
FMECA	FMEA with an assessment of criticality	–	4.6.1
Risk priority number	ordinal screening index	$\mathrm{RPN} = S\,O\,D$, 1 to 1000	4.6.1
Competing failure modes	first failure from any mode ends the mission	$\lambda_{\mathrm{total}} = \sum_i \lambda_i$, h^{-1}	4.6.3
Cause-specific failure probability	first failure by mode i before H	$F_i(H)$, eq. (4.24)	4.6.3
Quantitative failure model	probabilities from stated assumptions and evidence	$\mathcal{R}_i = \Pr(F_i)\mathbb{E}[C \mid F_i]$, R$/yr	4.6.4
Fault tree analysis	deductive logic of a top event	gates, basic events, minimal cut sets	4.6.5

Three ideas carry forward. Failure probabilities should not always be treated as fixed properties of components: mechanical stress, temperature, degradation and operating conditions can change them substantially. FMEA provides a systematic framework for identifying failures, but it is not sufficient by itself to calculate the reliability of a system. And the most frequently failing component is not necessarily the most important contributor to system-level risk.

Further reading. Rausand, Barros and Høyland, *System Reliability Theory*, third edition (failure modes, component reliability and system reliability analysis). Modarres, Kaminskiy and Krivtsov, *Reliability Engineering and Risk Analysis* (failure mechanisms, reliability analysis and risk methods). O'Connor and Kleyner, *Practical Reliability Engineering* (failure analysis and engineering applications). NASA, *Reliability-Centered Maintenance Guide* (Section 4.2, on FMEA). NASA Goddard, *FMEA and Risk Assessment Guideline* (2024) (practical methodology for maintaining FMECA studies).

4.8 Exercises

Exercise 4.1. In a data center chilled-water system, a temperature sensor develops a positive bias of +4 K. The actual supply-water temperature is 10°C and the reported temperature is 14°C. The control system responds to the reported high temperature by increasing the cooling output, and the energy consumption rises, although the actual supply-water temperature was already acceptable. Which classification is correct? (a) The sensor bias is the cause and the excessive cooling command is a possible effect. (b) The higher energy consumption is the physical failure mechanism of the sensor. (c) The chiller has necessarily experienced a mechanical failure. Then name the failure mode of the sensor and a possible cause of it. (Answer: (a); the sensor's failure mode is an erroneously high temperature reading, possibly caused by calibration drift, and the resulting control response raises the cooling energy without any mechanical failure of the chiller, so whether it is a system-level functional failure depends on whether the requirements include energy performance as well as temperature control.)

Exercise 4.2. A system has two parallel cooling-water pumps, each able to supply the required flow. The pumps share a common electrical feeder, a common flow sensor and a common discharge manifold. Table 4.12 is a starting table for an FMEA. (a) Extend it into a worksheet with the columns of Table 4.5 for at least five credible failure modes. (b) Classify each mode as mitigated by the redundancy of the two pumps or as affecting a shared dependency. (c) Over a mission, each pump fails to deliver flow with probability 0.05, independently; the feeder loses supply with probability 0.002; the manifold ruptures with probability 0.001 (illustrative). Compute the probability of losing the flow counting the pumps only, and counting the feeder and the manifold as well, assuming all these events independent. What fraction of the second result comes from the shared items? (d) Explain why the nominal existence of two pumps does not imply highly reliable cooling service, with reference to the flow sensor.

Exercise 4.3. (a) A bracket has strength $C \sim \mathcal{N}(150, 15^2)\,\mathrm{kN}$ and meets a load $L \sim \mathcal{N}(100, 20^2)\,\mathrm{kN}$, independent. Compute μ_g, σ_g, the reliability index and p_f. With the same standard deviations, what mean strength gives $p_f = 10^{-3}$? (b) The discharge weld of Section 4.4.1, with the same S-N curve, is moved to a duty of 10^5 cycles per year at 120 MPa and 5000 cycles per year at 200 MPa. Compute the Miner damage per year and the predicted life in years. (c) For an insulation with $E_a = 0.5\,\mathrm{eV}$, compute the Arrhenius acceleration factor from 40°C to 60°C and compare it with the value for 0.70 eV of Section 4.3.4.

Exercise 4.4. A pump has four failure modes with constant hazards (illustrative): seal leakage $3 \times 10^{-4}\,\mathrm{h}^{-1}$, bearing seizure $2 \times 10^{-4}\,\mathrm{h}^{-1}$, motor failure $1 \times 10^{-4}\,\mathrm{h}^{-1}$ and impeller failure

Table 4.12: Starting table for an FMEA of two parallel pumps with shared dependencies.

Item	Candidate failure mode	Questions to investigate
Pump A	Fails to start	Can pump B take over automatically?
Pump B	Reduced hydraulic capacity	Can pump A meet demand alone?
Electrical feeder	Loss of supply	Are both pumps disabled?
Flow sensor	Erroneously high flow indication	Could control logic fail to detect insufficient flow?
Discharge manifold	Leakage or rupture	Does a single physical failure disable both pumps?

$0.5 \times 10^{-4}\,\mathrm{h}^{-1}$. (a) Compute λ_{total}, $R(2000\,\mathrm{h})$ and, for each mode, the probability $F_i(2000\,\mathrm{h})$ that it causes the first failure within 2000 h. (b) The consequence of a first failure is R\$ 2 000 for seal leakage, R\$ 10 000 for bearing seizure or motor failure, and R\$ 50 000 for impeller failure. Compute the expected-loss contribution of each mode over the mission and rank the modes. (c) The FMECA ratings (S, O, D) are $(3,8,3)$, $(7,6,5)$, $(7,4,2)$ and $(9,3,4)$ in the same order. Compute the RPNs, rank the modes by RPN, and explain the difference between the two rankings.

Part II

Statistical Reliability Models

Chapter 5

Lifetime Distributions and Hazard Functions

5.1 Lifetime as a random variable

Chapter 4 described the physical processes by which components fail: fatigue, wear, corrosion and thermal degradation, each driven by the operating conditions and by the damage the component has already accumulated. That description explains why a component fails. It does not say when, and an engineer who must plan spares, schedule maintenance or accept a design needs the answer as a probability. The question of this chapter is

how the time at which a component fails is described mathematically, how its probability of failure evolves as it ages, and how it is decided whether its propensity to fail increases with age.

The chapter answers it in five steps. It defines the four functions that describe a lifetime, the cumulative failure probability, the failure density, the reliability function and the hazard function, and derives the relations among them. It derives the exact probability that a component which has survived to a given age fails within a further interval, and compares it with the short-interval approximation. It classifies the possible shapes of the hazard function and works two examples, one with a constant and one with an increasing hazard. It derives the mean time to failure and the mean residual life from the reliability function. Finally, it lets the hazard depend on operating conditions as well as on age, and states what a hazard model can and cannot say about the physical cause of failure.

Jargon: lifetime, nonrepairable item, CDF, PDF, reliability function, hazard function, cumulative hazard, conditional reliability, memoryless property, MTTF, MRL, bathtub curve, proportional hazards

The *lifetime* or *time to failure* T of an item is the operating time from a defined starting condition to its first failure; it is a random variable. A *nonrepairable item* is one whose life ends at its first failure; for a repairable item the same description applies to the time to its first failure. The *cumulative distribution function* (CDF) $F(t) = \Pr(T \le t)$ is the probability that the failure has occurred by time t; in reliability it is called the cumulative failure probability. The *probability density function* (PDF) $f(t) = \mathrm{d}F/\mathrm{d}t$ is the failure density, in h^{-1}. The *reliability function*, also called the *survivor function*, is $R(t) = \Pr(T > t)$. The *hazard function* $\lambda(t)$, also called the *failure rate function* or simply the *failure rate*, is the instantaneous rate of failure at age t among the items that have survived to t, in h^{-1}. The *cumulative hazard* $\Lambda(t)$ is the integral of the hazard from 0 to t, a dimensionless number. *Conditional reliability* is the probability of surviving a further interval given survival to the present age. The *memoryless property* states that this probability does not depend on the age. *MTTF* (mean time to failure) is the expected lifetime $\mathbb{E}[T]$; *MRL* (mean residual life) is the expected remaining lifetime of an item that has survived to age t. The *bathtub curve* is a hazard that decreases in early life, stays nearly constant in a middle period and increases in wear-out. A *proportional hazards* model multiplies a baseline hazard by a factor that depends on measured operating conditions, the *covariates*.

Key idea

The four functions F, f, R and λ carry the same information about a lifetime, and each can be computed from any other. The hazard is a rate conditional on survival, in inverse hours, not a probability: the probability of failing within a finite interval is obtained by integrating the hazard over that interval, $1 - \exp[-\int_t^{t+\tau} \lambda(u)\,\mathrm{d}u]$.

This section states the questions a lifetime model answers, fixes what the random time to failure is, and names the four functions with which the rest of the chapter works.

5.1.1 Questions about when an item fails

Consider an industrial cooling-water pump of the kind installed in large numbers in a plant. Four questions about it recur in practice. What is the probability that the pump fails during its first 1000 h of operation? If the pump has already operated for 5000 h without failure, what is the probability that it survives another 1000 h? Is an old pump more likely to fail in the next week than a new one? And from the failure records of a fleet of such pumps, can random failures be told apart from failures caused by aging? The first question is about an unconditional probability; the second and third are about probabilities conditional on survival to a given age; the fourth is about the shape of a function of age. Answering all four requires a model of the time at which the pump fails, and four functions of that time.

5.1.2 The random time to failure

Let T denote the random time to failure, measured in operating hours from a defined starting condition, normally the moment the item is placed in service in working order. Throughout this chapter T is a continuous, nonnegative random variable that represents the first failure of a nonrepairable item, or the first failure of a repairable item before any repair. Failure is the termination of the item's ability to perform its required function, as defined in Chapter 2; a pump that runs but delivers less than the required flow has failed in this sense.

Three assumptions are made explicit because the results below depend on them. The distribution of T is continuous, so that the probability of failure at any single instant is zero and a density exists. No probability is concentrated at $t = 0$, so that the item is in working order when it starts; an item that may be dead on arrival needs a separate probability for that event. And the probability of surviving any finite time is positive, so that the conditional quantities below are defined. Repairable items, whose life is a sequence of failures and restorations, are treated in a later chapter on repairable systems; the functions of this chapter describe each of their times to failure.

5.1.3 The four lifetime functions

Four functions describe the distribution of T. Table 5.1 lists them: each row gives the symbol, the standard name, the definition, the units when time is in hours, and the question the function answers. The cumulative failure probability and the reliability function are probabilities and are dimensionless. The failure density and the hazard function are rates, in inverse hours, and are not probabilities. When the distribution is continuous and sufficiently regular, the four functions contain equivalent information: each one determines the other three. The next two sections derive the relations that make this so.

Table 5.1: The four functions of a lifetime T. Units assume time in hours.

Symbol	Name	Definition	Units	Question answered
$F(t)$	cumulative failure probability (CDF)	$\Pr(T \le t)$	–	What fraction of the population has failed by t?
$f(t)$	failure density (PDF)	$\mathrm{d}F(t)/\mathrm{d}t$	h^{-1}	How is the probability of failure distributed over time?
$R(t)$	reliability (survivor) function	$\Pr(T > t)$	–	What fraction survives beyond t?
$\lambda(t)$	hazard function (failure rate)	$f(t)/R(t)$	h^{-1}	How fast do the survivors at age t fail?

5.2 Cumulative failure probability, failure density and reliability

This section defines the first three functions, derives the relations among them, and uses them to compute the probability that an item which has already survived to some age survives a further interval, the step that leads to the hazard function.

5.2.1 Cumulative failure probability

The cumulative distribution function of the lifetime is the probability that failure has occurred by time t,

$$F(t) = \Pr(T \le t). \tag{5.1}$$

It has a direct empirical reading. Suppose that 1000 identical pumps start operating at $t = 0$ under comparable conditions, and that after 1000 h of operation 100 of them have failed. The empirical estimate of the cumulative failure probability is the fraction failed,

$$\hat{F}(1000\,\mathrm{h}) = \frac{100}{1000} = 0.10,$$

where the hat marks an estimate from data. If 250 pumps have failed cumulatively by 2000 h, then $\hat{F}(2000\,\mathrm{h}) = 250/1000 = 0.25$. These numbers are illustrative, and they are used throughout

this section. The simple fraction assumes that every pump is followed until it fails or until the time of interest: no pump is removed from observation while still working, which would be a censored observation, and none is lost to another cause. Real failure records rarely satisfy this, and the estimation of F from censored data is the subject of a later chapter on estimation and censoring.

For a nonnegative lifetime the function has three properties. It is a probability,

$$0 \le F(t) \le 1;$$

it is nondecreasing, because the event $\{T \le t_1\}$ is contained in $\{T \le t_2\}$ whenever $t_2 > t_1$,

$$F(t_2) \ge F(t_1), \qquad t_2 > t_1;$$

and, for a lifetime that is finite with probability one, every item eventually fails,

$$\lim_{t \to \infty} F(t) = 1.$$

The cumulative failure probability describes how much failure probability has accumulated by time t. It does not describe how fast items are failing at time t; that is the role of the next two functions.

5.2.2 Failure density

For a continuous lifetime, the probability density function is the derivative of the cumulative distribution,

$$f(t) = \frac{\mathrm{d}F(t)}{\mathrm{d}t}, \qquad F(t) = \int_0^t f(u)\,\mathrm{d}u, \tag{5.2}$$

where u is a variable of integration over ages from 0 to t. The integral form uses $F(0) = 0$, the assumption of no probability mass at $t = 0$. The density describes how the probability of failure is distributed over the possible failure times. Over a short interval of length Δt the probability of failure is approximately the density times the length,

$$\Pr(t < T \le t + \Delta t) = F(t + \Delta t) - F(t) \approx f(t)\,\Delta t,$$

an approximation that becomes exact in the limit $\Delta t \to 0$, since that limit is the definition of the derivative.

A probability density is not a probability. With time in hours, f has units of h^{-1}, and its numerical value can exceed one without contradiction: a population of items that all fail within the same half hour has a density near $2\,\mathrm{h}^{-1}$ over that half hour. What must lie between zero and one is the probability obtained by integrating the density over an interval, such as the probability of failure between 1000 and 1100 h,

$$\Pr(1000 < T \le 1100) = \int_{1000}^{1100} f(t)\,\mathrm{d}t.$$

This is the unconditional probability that failure occurs in those 100 h, measured as a fraction of the initial population.

Figure 5.1 shows these areas for the pump population. The horizontal axis is operating time in hours; the vertical axis is the failure density, in units of $10^{-4}\,\mathrm{h}^{-1}$. The curve is an illustrative smooth density chosen to pass through the two empirical values $F(1000) = 0.10$ and $F(2000) = 0.25$. The gray area under the curve up to 1000 h is $F(1000) = 0.10$; the blue area between 1000 and 2000 h is the probability of failure in that interval, $0.25 - 0.10 = 0.15$; and the green area beyond 2000 h is the probability of surviving beyond 2000 h, 0.75. The three areas add to one. Every probability about the lifetime is an area under the density, and the density itself is the height, a rate per hour.

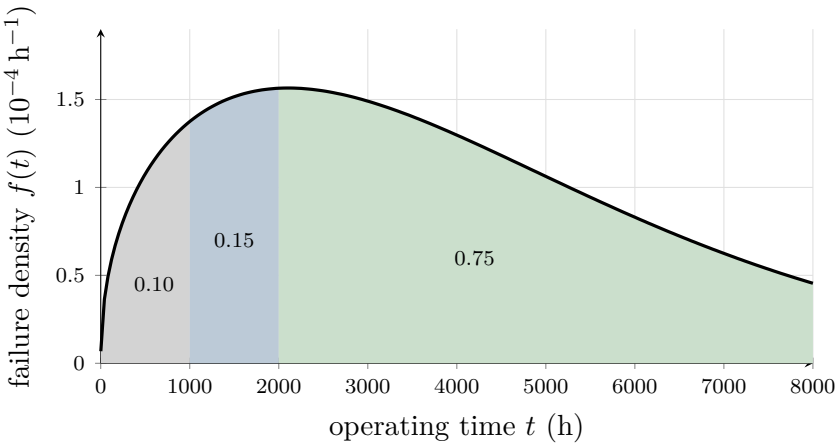


Figure 5.1: Probabilities as areas under the failure density, illustrative. The curve is a smooth density chosen to pass through $F(1000) = 0.10$ and $F(2000) = 0.25$, namely $f(t) = (\beta/\eta)(t/\eta)^{\beta-1}\exp[-(t/\eta)^{\beta}]$ with $\beta = 1.449$ and $\eta = 4725$ h, a member of the Weibull family treated in the next chapter. Gray: failure by 1000 h, 0.10. Blue: failure between 1000 and 2000 h, 0.15. Green: survival beyond 2000 h, $R(2000) = 0.75$.

5.2.3 Reliability function

The reliability function is the probability that the item survives beyond time t,

$$R(t) = \Pr(T > t). \tag{5.3}$$

Failure by time t and survival beyond t are complementary events: exactly one of them occurs. Their probabilities therefore add to one,

$$R(t) = 1 - F(t). \tag{5.4}$$

Differentiating equation (5.4) with respect to t and using equation (5.2) gives $\mathrm{d}R/\mathrm{d}t = -\mathrm{d}F/\mathrm{d}t = -f(t)$, so

$$f(t) = -\frac{\mathrm{d}R(t)}{\mathrm{d}t}. \tag{5.5}$$

The failure density is the negative slope of the survival curve: where the reliability function falls steeply, many items are failing per hour; where it is flat, few are. Because $f \geq 0$, the reliability function never increases, and it runs from $R(0) = 1$ toward zero.

For the pump population, $F(1000) = 0.10$ gives

$$R(1000) = 1 - 0.10 = 0.90,$$

and $F(2000) = 0.25$ gives $R(2000) = 0.75$. For a population of comparable pumps under the modeled conditions, 75 % are expected to survive beyond 2000 h. These are statements about new pumps. The probability that a pump which has already survived 1000 h survives another 1000 h is a different quantity, and it requires conditional probability.

5.2.4 Conditional survival probability

Suppose a pump has operated for 1000 h without failure. The probability that it survives to 2000 h is the conditional probability $\Pr(T > 2000 \mid T > 1000)$. For two events A and B with $\Pr(B) > 0$, conditional probability is defined by

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}.$$

Here $A = \{T > 2000\}$ and $B = \{T > 1000\}$. Survival beyond 2000 h implies survival beyond 1000 h, so $A \subset B$ and $A \cap B = A$. Hence

$$\Pr(T > 2000 \mid T > 1000) = \frac{\Pr(T > 2000)}{\Pr(T > 1000)} = \frac{R(2000)}{R(1000)}. \tag{5.6}$$

In general, for an age t and a further interval of length τ, the conditional reliability is

$$R(t + \tau \mid t) = \frac{R(t + \tau)}{R(t)}.$$

With the pump numbers,

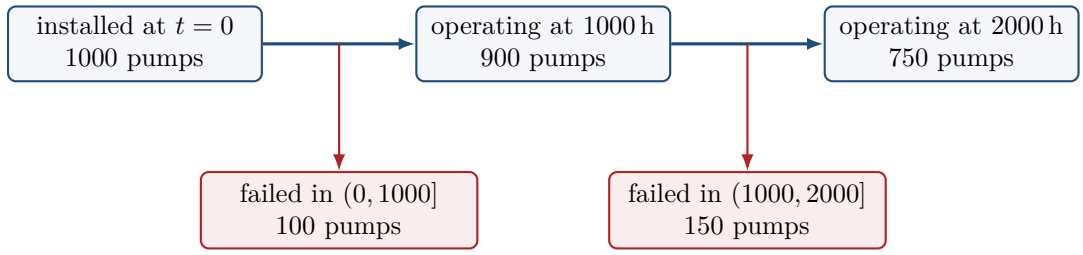
$$\Pr(T > 2000 \mid T > 1000) = \frac{0.75}{0.90} = 0.8333 :$$

a pump known to have survived its first 1000 h survives the next 1000 h with probability 0.8333. The conditional probability of failure in the second interval is the complement, $1-0.8333 = 0.1667$. It can also be computed directly, since the event of failure in $(t, t + \tau]$ is contained in survival to t:

$$\Pr(t < T \leq t + \tau \mid T > t) = \frac{F(t + \tau) - F(t)}{R(t)}, \tag{5.7}$$

which for the pumps is $(0.25 - 0.10)/0.90 = 0.15/0.90 = 0.1667$, the same value.

Figure 5.2 follows the 1000 pumps through the two intervals. The boxes on the top row are the pumps still operating at 0, 1000 and 2000 h; the boxes below are the pumps that failed in each interval. The two numbers at the bottom right are the same 150 failures divided by two different populations. Divided by the 1000 pumps installed, they give 0.15, the unconditional probability of failure between 1000 and 2000 h for a new pump. Divided by the 900 pumps that reached 1000 h, they give 0.1667, the probability of failure in that interval for a pump known to be working at 1000 h. The second probability is conditional on survival, and that distinction leads directly to the hazard function.



$R(1000) = 0.90$
$R(2000) = 0.75$
$R(2000)/R(1000) = 0.8333$

$150/1000 = 0.15$: unconditional, a fraction of all pumps
$150/900 = 0.1667$: conditional on survival to 1000 h

Figure 5.2: The illustrative population of 1000 pumps followed through two intervals of 1000 h. The same 150 failures in the second interval are 15 % of the pumps installed but 16.67 % of the pumps still operating at 1000 h (equation (5.7)).

5.3 The hazard function

The conditional probability of failure in a short interval, divided by the length of the interval, is the hazard function. This section defines it, interprets it, derives the relation between hazard and reliability, and collects the relations among all four functions.

5.3.1 Definition

The hazard function is the instantaneous rate of failure at age t among the items that have survived to t,

$$\lambda(t) = \lim_{\Delta t \to 0^+} \frac{\Pr(t < T \le t + \Delta t \mid T > t)}{\Delta t}. \tag{5.8}$$

The derivation of its working form takes three steps. First, the numerator is the conditional failure probability of equation (5.7) with $\tau = \Delta t$:

$$\lambda(t) = \lim_{\Delta t \to 0^+} \frac{F(t + \Delta t) - F(t)}{R(t)\,\Delta t}.$$

Second, $R(t)$ does not depend on Δt and can be taken out of the limit. Third, the remaining limit is the derivative of F, which is the density, wherever f is continuous:

$$\lim_{\Delta t \to 0^+} \frac{F(t + \Delta t) - F(t)}{\Delta t} = f(t).$$

Therefore

$$\lambda(t) = \frac{f(t)}{R(t)}. \tag{5.9}$$

This is one of the central equations of reliability theory. It connects the failure density, a rate per hour measured against the original population, with the probability that the item is still operating. The hazard is the density renormalized to the survivors.

5.3.2 Interpretation and units

Imagine 1000 pumps installed, of which 600 are still operating after several years. The hazard at that age describes the instantaneous propensity to fail among those 600 survivors, not among the original 1000. As an illustrative count, if 12 of the survivors fail in the next 100 h, the hazard at that age is estimated as $12/(600 \times 100\,\mathrm{h}) = 2.0 \times 10^{-4}\,\mathrm{h}^{-1}$, while the density is estimated as $12/(1000 \times 100\,\mathrm{h}) = 1.2 \times 10^{-4}\,\mathrm{h}^{-1}$; the ratio of the two is $R = 0.6$, as equation (5.9) requires. This normalization by the surviving population is what makes the hazard different from the density. Late in life, when few items survive, the density is small because few items are left to fail, while the hazard of each survivor can be large.

The hazard has units of inverse time, and it is not a probability. A hazard of $\lambda(t) = 0.002\,\mathrm{h}^{-1}$ does not mean that the probability of failure at time t is 0.002; the probability of failure at any single instant is zero. What the hazard gives is the conditional probability of failure over a short interval, approximately

$$\Pr(t < T \le t + \Delta t \mid T > t) \approx \lambda(t)\,\Delta t. \tag{5.10}$$

With $\Delta t = 0.1\,\mathrm{h}$, the probability that a pump which has survived to t fails in the next six minutes is about $0.002\,\mathrm{h}^{-1} \times 0.1\,\mathrm{h} = 0.0002$. The approximation assumes that the hazard changes negligibly over the interval and that the interval is short enough for the product to be small. For a longer interval the hazard must be integrated, which the next subsection prepares.

Common mistake

Reading a hazard as a probability. A hazard of $0.002\,\mathrm{h}^{-1}$ is a rate: multiplied by an interval it gives an approximate probability only when the product is small, and over 1000 h the product, 2, is not a probability at all. The correct probability over a finite interval is equation (5.14) of Section 5.4.1.

The name "failure rate" for $\lambda(t)$ is standard but has two other uses that should not be confused with it: the constant λ of the exponential model, which is the special case of a hazard that does not depend on age, and the rate at which a repairable system accumulates failures over calendar time, which is a property of a sequence of failures and is treated in a later chapter on repairable systems.

5.3.3 Reliability from the hazard

The hazard determines the reliability function completely. The derivation starts from the two relations already obtained, equation (5.9) and equation (5.5). Substituting the second into the first,

$$\lambda(t) = -\frac{1}{R(t)}\,\frac{\mathrm{d}R(t)}{\mathrm{d}t}.$$

By the chain rule, the derivative of the natural logarithm of R is $\frac{\mathrm{d}}{\mathrm{d}t}\ln R(t) = \frac{1}{R(t)}\frac{\mathrm{d}R(t)}{\mathrm{d}t}$, so

$$\lambda(t) = -\frac{\mathrm{d}}{\mathrm{d}t}\ln R(t). \tag{5.11}$$

Integrating both sides over ages from 0 to t,

$$\int_0^t \lambda(u)\,\mathrm{d}u = -\int_0^t \frac{\mathrm{d}}{\mathrm{d}u}\ln R(u)\,\mathrm{d}u = -\ln R(t) + \ln R(0).$$

The item starts in working order and has zero probability of immediate failure, so $R(0) = 1$ and $\ln R(0) = 0$. Then $\ln R(t) = -\int_0^t \lambda(u)\,\mathrm{d}u$, and exponentiating,

$$R(t) = \exp\!\left[-\int_0^t \lambda(u)\,\mathrm{d}u\right]. \tag{5.12}$$

Equation (5.12) is fundamental: if the hazard function is known, the reliability for any mission duration follows by one integration. It also shows that the hazard cannot be chosen freely: for $R(t)$ to fall to zero as $t \to \infty$, as it must for a lifetime that is finite with probability one, the integral of the hazard must grow without bound.

Principle 5.1. *A hazard is a rate conditional on survival, in inverse time, not a probability. Reliability over a mission is the exponential of minus the integral of the hazard over the mission, equation* (5.12)*.*

5.3.4 The cumulative hazard and the complete relations

The integral in equation (5.12) recurs often enough to have a name. The cumulative hazard is

$$\Lambda(t) = \int_0^t \lambda(u)\,\mathrm{d}u, \qquad R(t) = \mathrm{e}^{-\Lambda(t)}, \qquad \Lambda(t) = -\ln R(t). \tag{5.13}$$

The cumulative hazard is dimensionless, since it is a rate in h^{-1} integrated over hours. Unlike a probability, it can exceed one: $\Lambda = 1$ corresponds to $R = \mathrm{e}^{-1} = 0.368$, and $\Lambda = 3$ to $R = 0.050$. It is never negative and never decreases, because the hazard is never negative.

With Λ and λ, the four functions take the compact form

$$\Lambda(t) = \int_0^t \lambda(u)\,\mathrm{d}u, \quad R(t) = \mathrm{e}^{-\Lambda(t)}, \quad F(t) = 1 - \mathrm{e}^{-\Lambda(t)}, \quad f(t) = \lambda(t)\,\mathrm{e}^{-\Lambda(t)},$$

where the last follows from $f = \lambda R$, equation (5.9). These relations assume a continuous lifetime distribution with the usual regularity conditions and a positive survival probability over the times considered. Table 5.2 gives the full set of conversions: each row is the function to be computed, and each column is the function it is computed from. The table is the precise sense in which the four functions carry the same information.

Table 5.2: Each lifetime function in terms of each other one, with $\Lambda(t) = \int_0^t \lambda(u)\,\mathrm{d}u$. Continuous lifetime, $R(0) = 1$, $R(t) > 0$.

To compute	from f	from F	from R	from λ
$F(t)$	$\int_0^t f(u)\,\mathrm{d}u$	–	$1 - R(t)$	$1 - \mathrm{e}^{-\Lambda(t)}$
$f(t)$	–	$\mathrm{d}F(t)/\mathrm{d}t$	$-\mathrm{d}R(t)/\mathrm{d}t$	$\lambda(t)\,\mathrm{e}^{-\Lambda(t)}$
$R(t)$	$\int_t^\infty f(u)\,\mathrm{d}u$	$1 - F(t)$	–	$\mathrm{e}^{-\Lambda(t)}$
$\lambda(t)$	$\dfrac{f(t)}{\int_t^\infty f(u)\,\mathrm{d}u}$	$\dfrac{\mathrm{d}F(t)/\mathrm{d}t}{1 - F(t)}$	$-\dfrac{\mathrm{d}}{\mathrm{d}t}\ln R(t)$	–

5.4 Conditional failure probability over a finite interval

The short-interval approximation of equation (5.10) is not enough for planning, where the intervals are days or months. This section derives the exact conditional probability of failure over a finite interval, specializes it to a constant hazard, and compares it with the linear approximation in a worked table.

5.4.1 The exact conditional failure probability

The question is the probability that an item which has survived to age t fails before age $t + \tau$, where τ is a finite interval. From equation (5.6), the conditional probability of survival over the interval is $R(t + \tau)/R(t)$, so the conditional failure probability is

$$\Pr(t < T \le t + \tau \mid T > t) = 1 - \frac{R(t + \tau)}{R(t)}.$$

Writing both reliabilities with the cumulative hazard, equation (5.13), the ratio becomes

$$\frac{R(t + \tau)}{R(t)} = \frac{\mathrm{e}^{-\Lambda(t+\tau)}}{\mathrm{e}^{-\Lambda(t)}} = \mathrm{e}^{-[\Lambda(t+\tau) - \Lambda(t)]} = \exp\left[-\int_t^{t+\tau} \lambda(u)\,\mathrm{d}u\right],$$

since the integral from 0 to $t + \tau$ minus the integral from 0 to t is the integral from t to $t + \tau$. Therefore

$$\Pr(t < T \le t + \tau \mid T > t) = 1 - \exp\left[-\int_t^{t+\tau} \lambda(u)\,\mathrm{d}u\right]. \tag{5.14}$$

This is the exact conditional failure probability. It depends only on the hazard over the interval itself: the history before t enters only through the fact of survival. It is always between zero and one, whatever the size of the integral.

5.4.2 Constant hazard and the linear approximation

For a constant hazard $\lambda(u) = \lambda$, the integral in equation (5.14) is $\lambda\tau$ and the conditional failure probability is

$$\Pr(\text{failure in the next } \tau \mid \text{survival to } t) = 1 - \mathrm{e}^{-\lambda\tau}, \tag{5.15}$$

independent of the age t. For small $\lambda\tau$ it is approximately $\lambda\tau$. The approximation follows from the Taylor series of the exponential, $\mathrm{e}^{-x} = 1 - x + x^2/2 - x^3/6 + \cdots$, which gives, with $x = \lambda\tau$,

$$1 - \mathrm{e}^{-x} = x - \frac{x^2}{2} + \frac{x^3}{6} - \cdots.$$

Two inequalities bracket the exact value for every $x \geq 0$. The exponential is convex, so it lies above its tangent at zero, $\mathrm{e}^{-x} \geq 1 - x$; and the Taylor series stopped after the quadratic term has the remainder $-\mathrm{e}^{-\xi}x^3/6 \leq 0$ for some ξ between 0 and x, so $\mathrm{e}^{-x} \leq 1 - x + x^2/2$. Together,

$$x - \frac{x^2}{2} \;\leq\; 1 - \mathrm{e}^{-x} \;\leq\; x. \tag{5.16}$$

The linear approximation therefore always overestimates the failure probability: it errs on the conservative side, by at most $x^2/2$. Its relative error, the overestimate as a fraction of the exact value, is

$$\varepsilon(x) = \frac{x - (1 - \mathrm{e}^{-x})}{1 - \mathrm{e}^{-x}} = \frac{x}{1 - \mathrm{e}^{-x}} - 1 \approx \frac{x}{2} + \frac{x^2}{12}, \tag{5.17}$$

where the last form is the start of the series of $x/(1 - \mathrm{e}^{-x})$. To first order, the relative error in percent is half of $\lambda\tau$ in percent: an interval with $\lambda\tau = 0.02$ gives an error of about 1 %.

For a hazard that is not constant, $\lambda(t)\,\tau$ carries a second error in addition to the linearization: it uses the hazard at the start of the interval in place of its integral over the interval, and so underestimates the probability when the hazard is increasing and overestimates it when the hazard is decreasing. The exact form, equation (5.14), has neither error.

5.4.3 Exact against approximate probability

Table 5.3 compares the exact conditional failure probability $1 - \mathrm{e}^{-\lambda\tau}$ of equation (5.15) with the linear approximation $\lambda\tau$, for values of the product $\lambda\tau$ from 0.002 to 2. Only the product matters, so each row holds for any combination of hazard and interval with that product; the second column gives the interval τ for a pump with $\lambda = 0.002\,\mathrm{h}^{-1}$. The last two columns are the absolute difference and the relative error of equation (5.17). One row of the arithmetic, for $\lambda = 0.002\,\mathrm{h}^{-1}$ and $\tau = 100\,\mathrm{h}$: $\lambda\tau = 0.2$, the exact probability is $1 - \mathrm{e}^{-0.2} = 1 - 0.818731 = 0.181269$, the approximation gives 0.2, the difference is 0.018731, and the relative error is $0.018731/0.181269 = 10.33\,\%$. The exact probability over these 100 h is 18.13 %, and the approximation gives 20 %.

Table 5.3: Exact conditional failure probability against the linear approximation, for a constant hazard. The interval column assumes $\lambda = 0.002\,\mathrm{h}^{-1}$. The rows $\lambda\tau = 0.002$ and 2.0 extend the range to show both ends of the trend.

$\lambda\tau$	τ at $\lambda = 0.002\,\mathrm{h}^{-1}$	exact $1 - \mathrm{e}^{-\lambda\tau}$	approximation $\lambda\tau$	difference	relative error
0.002	1 h	0.001998	0.002	0.000002	0.10 %
0.01	5 h	0.009950	0.01	0.000050	0.50 %
0.05	25 h	0.048771	0.05	0.001229	2.52 %
0.1	50 h	0.095163	0.1	0.004837	5.08 %
0.2	100 h	0.181269	0.2	0.018731	10.33 %
0.5	250 h	0.393469	0.5	0.106531	27.07 %
1.0	500 h	0.632121	1.0	0.367879	58.20 %
2.0	1000 h	0.864665	2.0	1.135335	131.30 %

The trend is the one equation (5.17) predicts. For a one-hour interval the exact probability is 0.001998 against 0.002, a relative error of 0.1 %, negligible for any purpose. The relative error is close to $\lambda\tau/2$ up to $\lambda\tau \approx 0.2$, reaches 27 % at 0.5 and 58 % at 1, where the approximation reports a certain failure and the exact probability is 63 %; beyond 1 the approximation exceeds one and is no longer a probability at all. Figure 5.3 draws the same comparison. The horizontal axis is $\lambda\tau$; the vertical axis is the conditional failure probability. The solid blue curve is the

exact value, which bends over and approaches one from below; the dashed red line is the approximation, which rises without limit; the dots are the rows of Table 5.3. The two coincide near the origin and separate visibly from $\lambda\tau \approx 0.2$. The size of the interval, measured by $\lambda\tau$, decides whether the approximation can be used.

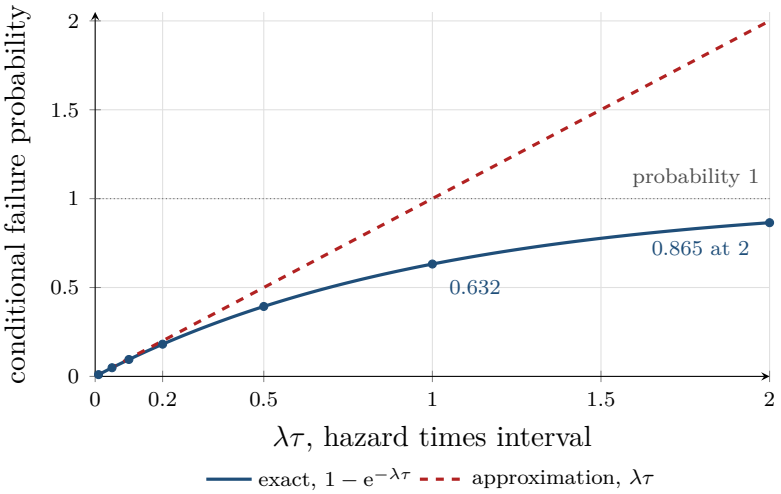


Figure 5.3: Exact conditional failure probability under a constant hazard, $1 - e^{-\lambda\tau}$ (equation (5.15)), against the linear approximation $\lambda\tau$. Dots are the rows of Table 5.3. The approximation always overestimates, equation (5.16), and exceeds one for $\lambda\tau > 1$.

Common mistake

Multiplying a failure rate by a long interval. For $\lambda = 0.002\,\mathrm{h}^{-1}$ over 500 h the product is 1, which would mean certain failure; the exact probability is 0.632. The product $\lambda\tau$ is the cumulative hazard over the interval, not the failure probability; it approximates the probability only when it is small, about 0.02 for a 1 % error.

5.5 Hazard shapes and two worked lifetime models

How the hazard changes with age is the most informative property of a lifetime model. This section classifies the three basic behaviors, discusses the bathtub curve that combines them, and then works two models through every function of the chapter: a pump with a constant hazard and an aging component with an increasing hazard.

5.5.1 Decreasing, constant and increasing hazard

Three broad behaviors are distinguished; specific distributions that produce each are the subject of the next chapter.

Decreasing hazard. The hazard of the survivors falls with age, a *decreasing failure rate* (DFR). This is the behavior of early-life failures, sometimes called infant mortality, caused by manufacturing defects, installation errors and weak units. In a population it often arises from heterogeneity: weak units fail early and leave a stronger surviving population. The practice of operating new units under controlled conditions before service, called *burn-in*, exploits this behavior to remove the weak units before they reach the field.

Constant hazard. The hazard does not depend on age: conditional on survival, a unit's propensity to fail is the same at every age. This is the exponential lifetime model of Section 5.5.3, appropriate for failures caused by external shocks that arrive at random rather than by accumulated damage.

Increasing hazard. The hazard of the survivors rises with age, an *increasing failure rate* (IFR). This is the behavior of wear-out and aging, produced by the accumulating damage of the mechanisms of Chapter 4: fatigue, corrosion, wear and insulation breakdown.

Figure 5.4 draws one example of each, using models that appear in this chapter. Each panel has operating age in hours on the horizontal axis and the hazard in units of $10^{-4}\,\mathrm{h}^{-1}$ on the vertical axis, on the same vertical scale. Panel (a) is a population of which 10 % are weak units with a constant hazard of $100 \times 10^{-4}\,\mathrm{h}^{-1}$ and 90 % strong units with $1 \times 10^{-4}\,\mathrm{h}^{-1}$. Panel (b) is the constant hazard of the pump of Section 5.5.3. Panel (c) is the linear hazard of the aging component of Section 5.5.4. In panel (a) the population hazard starts at 10.9, falls to 2.5 by 200 h and to 1.08 by 500 h, and settles at the strong units' value of 1: a tenfold decrease, although no individual unit's hazard changes at all. The decrease is produced entirely by the weak units leaving the surviving population, whose weak fraction falls from 10 % at the start to 1.5 % at 200 h and below 0.1 % at 500 h.

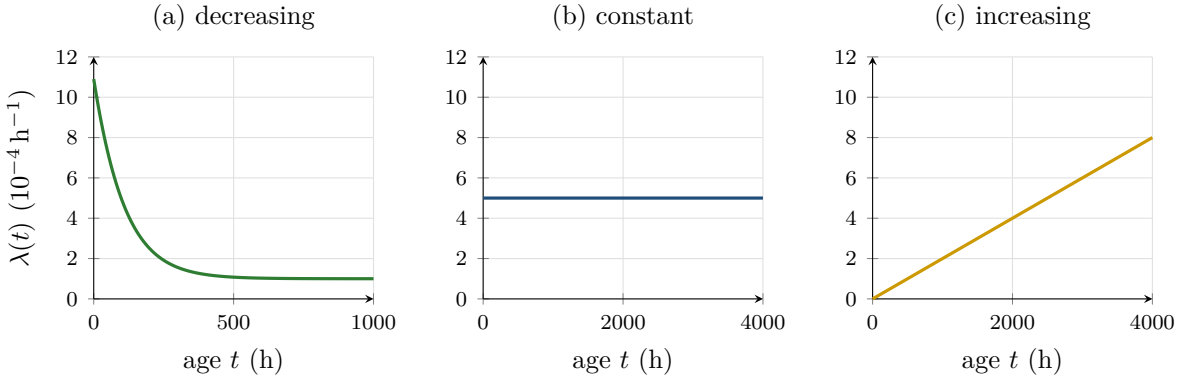


Figure 5.4: Three hazard behaviors, each from a model in this chapter. (a) Decreasing: a population of 10 % weak units with hazard $0.01\,\mathrm{h}^{-1}$ and 90 % strong units with $10^{-4}\,\mathrm{h}^{-1}$, $\lambda(t) = [0.1 \cdot 0.01\,\mathrm{e}^{-0.01t} + 0.9 \cdot 10^{-4}\mathrm{e}^{-10^{-4}t}]/[0.1\,\mathrm{e}^{-0.01t} + 0.9\,\mathrm{e}^{-10^{-4}t}]$, illustrative. (b) Constant: $\lambda = 5 \times 10^{-4}\,\mathrm{h}^{-1}$ (Section 5.5.3). (c) Increasing: $\lambda(t) = at$ with $a = 2 \times 10^{-7}\,\mathrm{h}^{-2}$ (Section 5.5.4).

5.5.2 The bathtub curve and its limits

The three behaviors are often combined into the *bathtub curve*, a hazard that decreases in early life, remains nearly constant during a period of useful life, and increases in wear-out. Figure 5.5 draws an illustrative one. The horizontal axis is age and the vertical axis is the hazard, both in arbitrary units, since the curve is a shape and not a model of any particular item. The solid curve is the total hazard; the three thin curves are the terms whose sum it is, one decreasing, one constant and one increasing; the dashed vertical lines separate the three periods named below the axis. Drawing the curve as a sum makes its usual reading explicit: the early-life, random and wear-out failure processes act at all ages, and each period is named after the process that dominates it.

The bathtub curve is a useful conceptual model, not a universal empirical law. Many components never exhibit all three periods: much electronic equipment shows no clear wear-out within its service life, and many mechanical components show wear-out from the start with no period of constant hazard. Failure mechanisms overlap in age, so the periods have no

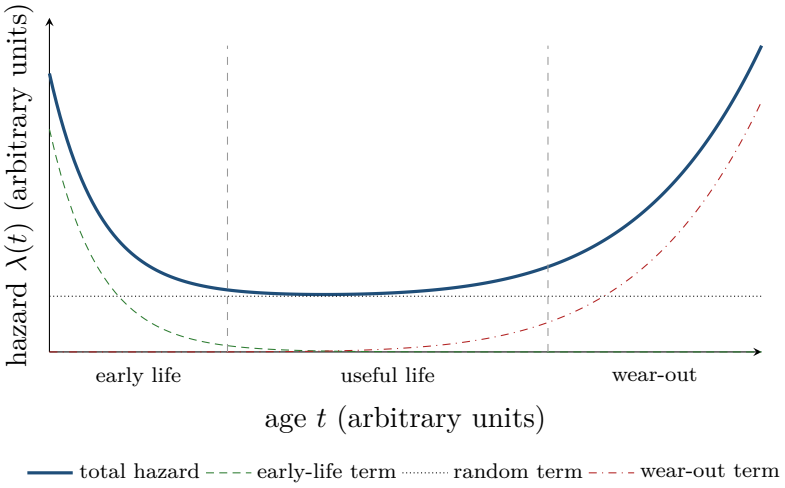


Figure 5.5: An illustrative bathtub-shaped hazard, $\lambda(t) = 0.8\,\mathrm{e}^{-t/0.07} + 0.2 + 0.9\,t^6$ in arbitrary units, drawn with its three terms. Each period is named after the term that dominates it; the boundaries at 0.25 and 0.7 are a matter of convention.

sharp boundaries. And a hazard estimated from a population describes the population, which can differ from the aging behavior of its individual members. Panel (a) of Figure 5.4 is the clearest case: the population hazard decreases tenfold while every unit keeps a constant hazard. A decreasing hazard in a population therefore does not mean that an individual component becomes physically stronger as it ages; it may mean only that the weaker components failed earlier and left a more robust surviving population.

Common mistake

Reading a population hazard as the aging of each unit. A decreasing population hazard can arise with no unit improving, from the removal of weak units; an increasing population hazard does not prove that every unit deteriorates at the same rate. The hazard is a property of the distribution of lifetimes in a population, not a measurement of any one unit's condition.

5.5.3 Worked example: the exponential lifetime model

Consider a pump with a constant hazard $\lambda = 0.0005\,\mathrm{h}^{-1}$. With $\lambda(u) = \lambda$ the cumulative hazard is $\Lambda(t) = \int_0^t \lambda\,\mathrm{d}u = \lambda t$, and the relations of Section 5.3.4 give the reliability function, the cumulative failure probability and the density:

$$R(t) = \mathrm{e}^{-0.0005t}, \qquad F(t) = 1 - \mathrm{e}^{-0.0005t}, \qquad f(t) = 0.0005\,\mathrm{e}^{-0.0005t}\,\mathrm{h}^{-1},$$

with t in hours. Its mean lifetime, derived in Section 5.6.1, is $\mathbb{E}[T] = 1/\lambda = 1/0.0005\,\mathrm{h}^{-1} = 2000\,\mathrm{h}$. Three questions are answered from these functions.

Survival of 1000 h. The probability that a new pump survives 1000 h is

$$R(1000) = \mathrm{e}^{-0.0005\times 1000} = \mathrm{e}^{-0.5} = 0.6065.$$

Failure within 1000 h. The probability that a new pump fails within 1000 h is the complement, $F(1000) = 1 - \mathrm{e}^{-0.5} = 0.3935$.

Survival of another 1000 h **after** 3000 h**.** For a pump that has already survived 3000 h, equation (5.6) gives

$$\Pr(T > 4000 \mid T > 3000) = \frac{R(4000)}{R(3000)} = \frac{\mathrm{e}^{-2}}{\mathrm{e}^{-1.5}} = \mathrm{e}^{-0.5} = 0.6065.$$

The probability of surviving the next 1000 h is the same for a new pump and for one that has already run 3000 h. The same cancellation happens for every age s and every interval t:

$$\Pr(T > s + t \mid T > s) = \frac{\mathrm{e}^{-\lambda(s+t)}}{\mathrm{e}^{-\lambda s}} = \mathrm{e}^{-\lambda t} = \Pr(T > t). \tag{5.18}$$

This is the *memoryless property* of the exponential distribution: a surviving unit is, probabilistically, as good as new. The exponential distribution is the only continuous lifetime distribution with this property. The argument is short: equation (5.18) requires $R(s+t) = R(s)\,R(t)$ for all $s, t \geq 0$, so $\ln R$ is additive, $\ln R(s+t) = \ln R(s) + \ln R(t)$; an additive function that is monotone is linear, $\ln R(t) = -\lambda t$ for some $\lambda \geq 0$, which is the exponential model. Equivalently, by equation (5.12), a constant hazard and an exponential reliability function are the same statement.

The model is mathematically convenient, and much of classical system reliability is built on it. For equipment undergoing physical aging or accumulating damage, however, memorylessness is unrealistic: a pump whose bearings have run 30 000 h is not as good as new, and an exponential model of it will understate the risk of its next month. Note also that the mean lifetime is not a time at which pumps fail: the fraction that survives beyond the mean is $R(2000) = \mathrm{e}^{-1} = 0.368$.

5.5.4 Worked example: an aging component

Now consider a component whose hazard grows linearly with operating age,

$$\lambda(t) = at, \qquad a = 2 \times 10^{-7}\,\mathrm{h}^{-2},$$

where the units of a make λ a rate in h^{-1} when t is in hours. At 1000 h the hazard is $\lambda(1000) = 2 \times 10^{-7} \times 1000 = 2 \times 10^{-4}\,\mathrm{h}^{-1}$; at 3000 h it is $6 \times 10^{-4}\,\mathrm{h}^{-1}$. The older component has three times the instantaneous propensity to fail. The model is illustrative; it is the Rayleigh distribution, the Weibull distribution of the next chapter with shape $\beta = 2$ and scale $\eta = \sqrt{2/a} = 3162\,\mathrm{h}$.

Step 1: cumulative hazard. Integrating the hazard,

$$\Lambda(t) = \int_0^t au\,\mathrm{d}u = \frac{at^2}{2}.$$

Step 2: reliability. From equation (5.13),

$$R(t) = \mathrm{e}^{-\Lambda(t)} = \mathrm{e}^{-at^2/2}.$$

Step 3: reliability at several ages. With $a/2 = 10^{-7}\,\mathrm{h}^{-2}$, the cumulative hazards at 1000, 2000 and 3000 h are 0.1, 0.4 and 0.9, and

$$R(1000) = \mathrm{e}^{-0.1} = 0.9048, \quad R(2000) = \mathrm{e}^{-0.4} = 0.6703, \quad R(3000) = \mathrm{e}^{-0.9} = 0.4066.$$

Step 4: future failure probability at two ages. For a component that has survived to 1000 h, the probability of failure in the next 500 h is, by equation (5.14),

$$1 - \frac{R(1500)}{R(1000)} = 1 - \mathrm{e}^{-[\Lambda(1500)-\Lambda(1000)]} = 1 - \mathrm{e}^{-(0.225-0.1)} = 1 - \mathrm{e}^{-0.125} = 0.1175,$$

or 11.75 %. For a component that has survived to 3000 h, the same interval gives

$$1 - \mathrm{e}^{-[\Lambda(3500)-\Lambda(3000)]} = 1 - \mathrm{e}^{-(1.225-0.9)} = 1 - \mathrm{e}^{-0.325} = 0.2775,$$

or 27.75 %. In general the cumulative hazard over the next 500 h from age t is $(a/2)[(t+500)^2 - t^2] = 10^{-4}\,t + 0.025$ with t in hours, which grows linearly with the age at which the interval starts.

Table 5.4 sets the two models side by side at five ages. The columns give, for each model, the hazard, the reliability of a new unit to that age, and the conditional probability of failure in the next 500 h for a unit that has reached that age. For the pump with constant hazard the last quantity is $1 - \mathrm{e}^{-0.25} = 0.2212$ at every age; for the aging component it rises from 0.025 at installation to 0.346 at 4000 h. Figure 5.6 plots the same comparison. In panel (a) the horizontal axis is age and the vertical axis the hazard: the pump's is flat at $5 \times 10^{-4}\,\mathrm{h}^{-1}$ and the aging component's rises through it at 2500 h. In panel (b) the vertical axis is the conditional failure probability over the next 500 h: flat for the pump, rising for the aging component, which crosses the pump's value at 2250 h, where its hazard averaged over the coming interval, $a(t + 250)$, equals $5 \times 10^{-4}\,\mathrm{h}^{-1}$. The two dots are the values of Step 4.

Table 5.4: The exponential pump ($\lambda = 5 \times 10^{-4}\,\mathrm{h}^{-1}$) and the aging component ($\lambda(t) = at$, $a = 2 \times 10^{-7}\,\mathrm{h}^{-2}$) at five ages. Hazards in $10^{-4}\,\mathrm{h}^{-1}$. The last two columns are the probabilities of failure in the next 500 h for a unit that has survived to the age in the first column.

Age t (h)	λ, pump	λ, aging	$R(t)$, pump	$R(t)$, aging	next 500 h, pump	next 500 h, aging
0	5	0	1.0000	1.0000	0.2212	0.0247
1000	5	2	0.6065	0.9048	0.2212	0.1175
2000	5	4	0.3679	0.6703	0.2212	0.2015
3000	5	6	0.2231	0.4066	0.2212	0.2775
4000	5	8	0.1353	0.2019	0.2212	0.3462

This is the fundamental difference between the exponential model and an aging model. Under a constant hazard, survival to an older age does not change the probability of failure over the next fixed interval. Under an increasing hazard, survival to an older age comes with a greater conditional failure probability over the next interval, because the integral of an increasing function over an interval of fixed length grows as the interval moves to later ages. The comparison also shows that the two models can rank the units in either order: the aging component is safer than the pump when young and riskier when old.

Principle 5.2. *Under a constant hazard the probability of failing in the next interval does not depend on age, which is the memoryless property. Under an increasing hazard it grows with age; under a decreasing hazard it falls. Which of the three holds is the first question to ask of any lifetime model.*

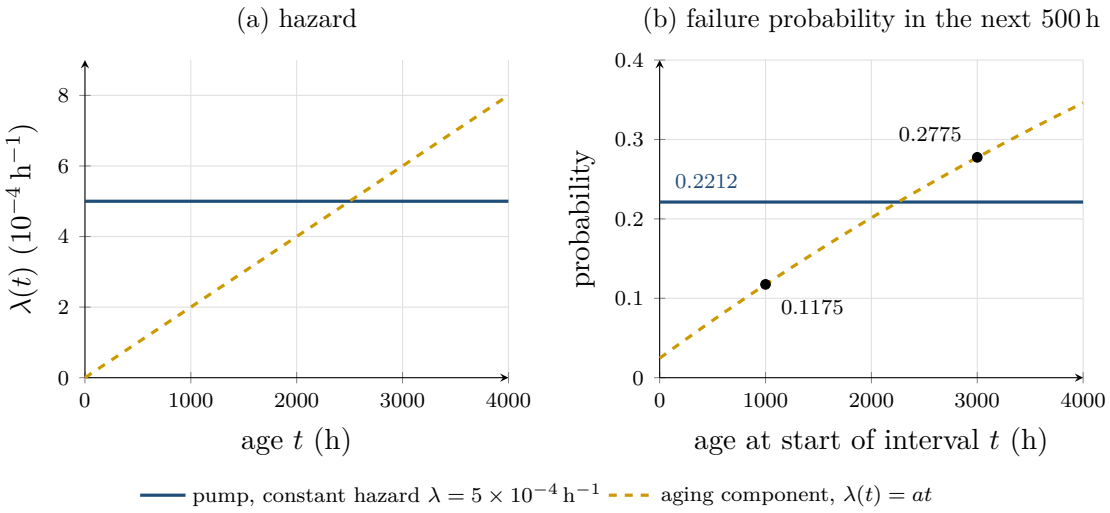


Figure 5.6: Constant against increasing hazard. (a) Hazard of the exponential pump, $\lambda = 5 \times 10^{-4}\,\mathrm{h}^{-1}$, and of the aging component, $\lambda(t) = at$ with $a = 2 \times 10^{-7}\,\mathrm{h}^{-2}$. (b) Conditional probability of failure in the next 500 h, from equation (5.14): $1 - \mathrm{e}^{-0.25} = 0.2212$ at every age for the pump, $1 - \exp[-(10^{-4}t + 0.025)]$ for the aging component. Dots: the ages 1000 and 3000 h of the worked example.

5.6 Mean life, condition-dependent hazard and the scope of hazard models

The chapter's functions so far describe probabilities at given ages. This section derives the expected lifetime and the expected remaining lifetime from the reliability function, lets the hazard depend on the operating conditions as well as on age, and states what a hazard model can and cannot reveal about the physics of failure.

5.6.1 Mean time to failure from the reliability function

The mean time to failure is the expected lifetime, which for a continuous lifetime is by definition

$$\mathrm{MTTF} = \mathbb{E}[T] = \int_0^\infty t\, f(t)\,\mathrm{d}t.$$

It can be computed directly from the reliability function,

$$\mathrm{MTTF} = \int_0^\infty R(t)\,\mathrm{d}t, \tag{5.19}$$

a general result that does not require any particular distribution. Two derivations are standard.

By integration by parts. Write $f = -\mathrm{d}R/\mathrm{d}t$ (equation (5.5)) and integrate by parts on $[0, b]$:

$$\int_0^b t\, f(t)\,\mathrm{d}t = -\int_0^b t\,\frac{\mathrm{d}R}{\mathrm{d}t}\,\mathrm{d}t = \big[-t\,R(t)\big]_0^b + \int_0^b R(t)\,\mathrm{d}t = -b\,R(b) + \int_0^b R(t)\,\mathrm{d}t.$$

When the mean is finite, the boundary term vanishes as $b \to \infty$, because $b\,R(b) = b\int_b^\infty f(t)\,\mathrm{d}t \le \int_b^\infty t\, f(t)\,\mathrm{d}t$, the tail of a convergent integral, which tends to zero. Letting $b \to \infty$ gives equation (5.19).

By the tail-integral identity. For a nonnegative T, write $T = \int_0^\infty \mathbf{1}\{T > t\}\,\mathrm{d}t$, where the indicator $\mathbf{1}\{\cdot\}$ is one when its condition holds and zero otherwise; the integrand is one for $t < T$ and zero after, so the integral is the length T. Taking expectations and exchanging expectation and integral, which is allowed because the integrand is nonnegative,

$$\mathbb{E}[T] = \int_0^\infty \mathbb{E}\big[\mathbf{1}\{T > t\}\big]\,\mathrm{d}t = \int_0^\infty \Pr(T > t)\,\mathrm{d}t = \int_0^\infty R(t)\,\mathrm{d}t.$$

The result says that the mean time to failure is the area under the reliability curve.

For the exponential model, $R(t) = \mathrm{e}^{-\lambda t}$ and

$$\mathrm{MTTF} = \int_0^\infty \mathrm{e}^{-\lambda t}\,\mathrm{d}t = \left[-\frac{\mathrm{e}^{-\lambda t}}{\lambda}\right]_0^\infty = \frac{1}{\lambda},$$

which gives the 2000 h of the pump with $\lambda = 0.0005\,\mathrm{h}^{-1}$. For the aging component, $R(t) = \mathrm{e}^{-at^2/2}$ and the Gaussian integral $\int_0^\infty \mathrm{e}^{-ct^2}\,\mathrm{d}t = \frac{1}{2}\sqrt{\pi/c}$ with $c = a/2$ gives

$$\mathrm{MTTF} = \sqrt{\frac{\pi}{2a}} = \sqrt{\frac{3.1416}{4 \times 10^{-7}\,\mathrm{h}^{-2}}} = 2802\,\mathrm{h}.$$

The aging component has the longer mean life although its hazard exceeds the pump's beyond 2500 h: its low hazard when young keeps most units alive through the first 2000 h. The same identity serves for the Weibull and lognormal distributions of the next chapter.

5.6.2 Mean residual life

For a component that has already survived to age t, the relevant expectation is its remaining life, the mean residual life,

$$\mathrm{MRL}(t) = \mathbb{E}[T - t \mid T > t].$$

Given survival to t, the remaining life $T - t$ is nonnegative, and the tail-integral identity applies to it under the conditional distribution. Its survival probability beyond s is the conditional reliability $\Pr(T - t > s \mid T > t) = R(t + s)/R(t)$, so

$$\mathrm{MRL}(t) = \int_0^\infty \frac{R(t+s)}{R(t)}\,\mathrm{d}s = \frac{1}{R(t)}\int_t^\infty R(u)\,\mathrm{d}u, \tag{5.20}$$

with the substitution $u = t + s$. At $t = 0$ it is the MTTF. The distinction is between the mean lifetime, $\mathbb{E}[T]$, a property of a new unit, and the mean remaining lifetime, $\mathbb{E}[T - t \mid T > t]$, which conditions on having survived to the present age. The second is the quantity that matters for a unit in service, and it is central to prognostics, treated in a later chapter.

For the exponential model the ratio in equation (5.20) is $\mathrm{e}^{-\lambda s}$ whatever t is, so $\mathrm{MRL}(t) = 1/\lambda$ at every age, the memoryless property once more. For an aging distribution the mean residual life generally depends on age. For the aging component, $\mathrm{MRL}(t) = \sqrt{\pi/(2a)}\,\mathrm{e}^{at^2/2}\,\mathrm{erfc}(t\sqrt{a/2})$, where $\mathrm{erfc}(x) = (2/\sqrt{\pi})\int_x^\infty \mathrm{e}^{-v^2}\,\mathrm{d}v$ is the complementary error function; evaluated, it gives 2802 h at installation, 2028 h at 1000 h, 1552 h at 2000 h, 1239 h at 3000 h and 1022 h at 4000 h.

Figure 5.7 shows both results. In panel (a) the horizontal axis is time and the vertical axis reliability; the shaded area under the pump's exponential reliability curve is its MTTF, and the dashed rectangle of height one and width 2000 h has the same area; the gold curve is the aging component's reliability, whose area is 2802 h. In panel (b) the horizontal axis is the age already survived and the vertical axis the mean residual life: flat at 2000 h for the pump, falling steadily for the aging component, which crosses below the pump's value just beyond an age of 1000 h.

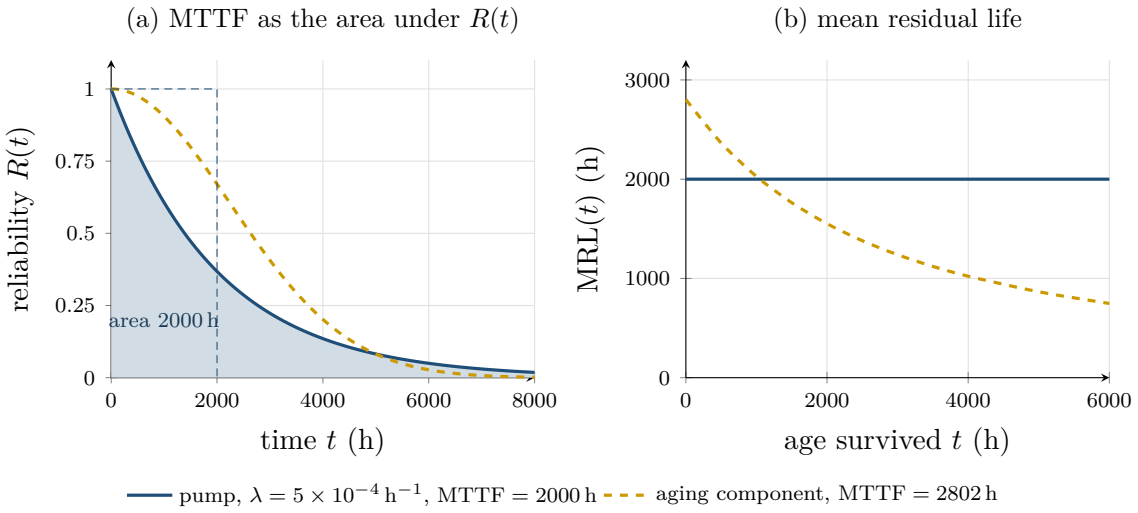


Figure 5.7: (a) Reliability of the exponential pump, $R(t) = \mathrm{e}^{-5\times10^{-4}t}$, and of the aging component, $R(t) = \mathrm{e}^{-10^{-7}t^2}$; the shaded area under the pump's curve equals the dashed rectangle, MTTF = 2000 h (equation (5.19)). (b) Mean residual life, equation (5.20): constant at $1/\lambda = 2000$ h for the pump, decreasing from 2802 h for the aging component, evaluated numerically at the plotted points.

5.6.3 From age-dependent to condition-dependent hazard

So far the hazard has been a function of operating age alone. Many engineering systems operate under changing environmental and loading conditions: a transmission line whose conductor temperature follows the current and the weather, a transformer whose insulation ages with its hot-spot temperature, a cooling pump whose bearings see a varying load, a battery whose degradation depends on its temperature and its cycling. The propensity to fail then depends on the conditions as well as on age,

$$\lambda(t, \theta, I, L, \ldots),$$

where θ is the operating temperature, I the electrical current, L the applied mechanical or thermal load, and t the operating age or accumulated exposure. The temperature is written θ here because T is reserved for the time to failure.

A standard way to write such a dependence collects the conditions in a vector of explanatory variables, the covariates $\mathbf{z}(t)$, and multiplies a baseline hazard by an exponential factor:

$$\lambda(t \mid \mathbf{z}(t)) = \lambda_0(t)\, \exp[\boldsymbol{\gamma}^{\mathsf{T}} \mathbf{z}(t)], \tag{5.21}$$

where $\lambda_0(t)$ is the baseline hazard, the hazard when every covariate is zero, and $\boldsymbol{\gamma}$ is the vector of coefficients that describe the effect of each covariate. With constant covariates this is the proportional hazards model: the covariates scale the whole hazard function by a constant factor, so that the hazards of two units under different constant conditions keep a fixed ratio. For a specified covariate history $\mathbf{z}(\cdot)$, the conditional reliability follows from equation (5.12) with the condition-dependent hazard inside the integral:

$$R(t \mid \mathbf{z}(\cdot)) = \exp\left[-\int_0^t \lambda(u \mid \mathbf{z}(u))\, \mathrm{d}u\right]. \tag{5.22}$$

The formulation assumes that the covariate history is known or specified for the whole interval, that it is not itself affected by the impending failure, and that the form of equation (5.21) holds for the component and the failure mechanism in question.

5.6.4 Example: temperature dependence

Consider a component whose hazard at a reference temperature $\theta_0 = 40°\mathrm{C}$ is constant, $\lambda_0 = 10^{-4}\,\mathrm{h}^{-1}$, and assume, for illustration, that every increase of 10°C doubles its hazard:

$$\lambda(\theta) = \lambda_0\, 2^{(\theta-\theta_0)/10} = \lambda_0 \exp\bigl[\gamma\,(\theta-\theta_0)\bigr], \qquad \gamma = \frac{\ln 2}{10°\mathrm{C}} = 0.0693\,(°\mathrm{C})^{-1}, \tag{5.23}$$

which is equation (5.21) with one covariate, $z = \theta - \theta_0$, and a constant baseline. For a 1000 h mission at constant temperature, the reliability is $R(1000 \mid \theta) = \exp[-\lambda(\theta) \times 1000\,\mathrm{h}]$. Table 5.5 evaluates the hazard, the 1000 h reliability and the mean time to failure $1/\lambda(\theta)$ at five temperatures; the rows at 40, 50 and 60°C are the core example, and the rows at 30 and 70°C are illustrative additions that show the trend. At 50°C, for instance, $\lambda = 10^{-4} \times 2 = 2 \times 10^{-4}\,\mathrm{h}^{-1}$ and $R(1000) = \mathrm{e}^{-0.2} = 0.8187$. Each 10°C halves the mean time to failure and squares the 1000 h reliability, since $\mathrm{e}^{-2x} = (\mathrm{e}^{-x})^2$: from 90.48 % at 40°C to 81.87 % at 50°C and 67.03 % at 60°C.

Table 5.5: Hazard, 1000 h reliability and mean time to failure of a component under the illustrative temperature model of equation (5.23), $\lambda_0 = 10^{-4}\,\mathrm{h}^{-1}$ at $\theta_0 = 40°\mathrm{C}$, hazard doubling every 10°C.

Temperature θ	Hazard $\lambda(\theta)$	$R(1000\,\mathrm{h} \mid \theta)$	MTTF $= 1/\lambda(\theta)$
30°C (added)	$0.5 \times 10^{-4}\,\mathrm{h}^{-1}$	0.9512	20 000 h
40°C	$1 \times 10^{-4}\,\mathrm{h}^{-1}$	0.9048	10 000 h
50°C	$2 \times 10^{-4}\,\mathrm{h}^{-1}$	0.8187	5000 h
60°C	$4 \times 10^{-4}\,\mathrm{h}^{-1}$	0.6703	2500 h
70°C (added)	$8 \times 10^{-4}\,\mathrm{h}^{-1}$	0.4493	1250 h

When the temperature varies during the mission, equation (5.22) integrates the hazard along the temperature history. As an illustrative case, let the same component spend 600 h at 40°C and 400 h at 60°C. The cumulative hazard over the mission is the sum over the two periods,

$$\Lambda(1000) = 10^{-4} \times 600 + 4 \times 10^{-4} \times 400 = 0.06 + 0.16 = 0.22, \qquad R(1000) = \mathrm{e}^{-0.22} = 0.8025.$$

Evaluating the hazard at the time-averaged temperature, $0.6 \times 40 + 0.4 \times 60 = 48°\mathrm{C}$, would give $\lambda = 10^{-4} \times 2^{0.8} = 1.741 \times 10^{-4}\,\mathrm{h}^{-1}$ and $R(1000) = \mathrm{e}^{-0.1741} = 0.8402$, an overestimate of the reliability. Because the hazard is a convex function of temperature, the hot periods contribute more than the average temperature suggests; the hazard must be integrated along the history, not evaluated at its mean.

The temperature-doubling rule is an assumed model, not a universal physical law; its validity must be established for the component and for the failure mechanism. The example shows how environmental conditions enter a probabilistic reliability model, and why a failure rate estimated under one set of conditions should not be applied under substantially different ones: a hazard measured at 40°C understates the hazard at 60°C by a factor of four under this model.

Principle 5.3. *A hazard estimated under one set of operating conditions describes those conditions. When the conditions vary, the reliability over a mission is obtained by integrating the condition-dependent hazard along the actual history of the conditions, equation* (5.22).

5.6.5 What a hazard model can and cannot tell

A hazard model describes the probability distribution of failure times. It does not, by itself, identify the physical cause of failure. Suppose an empirical model fitted to failure records gives

$$\lambda(t) = 0.0002\, t^{0.5},$$

with t in hours and the coefficient in $\mathrm{h}^{-1.5}$, so that λ is in h^{-1}; time units and coefficient dimensions must be kept consistent for such a formula to mean anything. The hazard increases with age, and it is reasonable to read that as evidence of an aging-related pattern of failure in the population. The model does not say whether the mechanism is fatigue, corrosion, insulation breakdown or a combination of processes; all of them can produce the same increasing hazard. Likewise, an increasing frequency of failure with age does not prove that every individual component deteriorates at the same rate, as the mixture of Section 5.5.2 showed for the decreasing case.

The distinction is between two kinds of model. *Statistical reliability modeling* describes the distribution of failure times, by the functions of this chapter fitted to failure data. *Physics-of-failure modeling* explains how stresses, material properties, damage and physical mechanisms produce failure, as in Chapter 4. The two can be combined. A fatigue crack-growth model, for example, predicts the time to failure as a function of the initial crack size, the material properties and the future load cycles. When those inputs are uncertain, the uncertainty induces a distribution of the failure time, from which $F(t)$, $R(t)$, $f(t)$ and $\lambda(t)$ follow by the relations of Table 5.2. In that combination the hazard is derived from the physics rather than assumed, and its shape inherits the meaning of the mechanism. This connection becomes central in the later chapters on degradation and prognostics, where the damage state of a component is followed in time and the remaining life is conditioned on it. The next chapter studies the specific lifetime distributions, exponential, Weibull and lognormal, whose shape parameters produce the decreasing, constant and increasing hazards described here.

5.7 What to retain

Definition 5.1 (Hazard and interval failure probability)**.** The hazard function $\lambda(t) = f(t)/R(t)$ is a rate of failure conditional on survival to age t, in inverse time, and not a probability; the probability that an item surviving to t fails within a further interval τ is obtained by integrating it, $1 - \exp[-\int_t^{t+\tau} \lambda(u)\,\mathrm{d}u]$, which the product $\lambda\tau$ approximates only when it is small.

Table 5.6 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 5.6: The terms of lifetime and hazard analysis and where each is defined. Units assume time in hours.

Term	Expression	Units	Meaning	Defined in
Time to failure	$T \geq 0$	h	random lifetime from a defined start	5.1.2
Cumulative failure probability	$F(t) = \Pr(T \leq t)$	–	failed by t	5.2.1
Failure density	$f(t) = \mathrm{d}F/\mathrm{d}t$	h^{-1}	distribution of failure times	5.2.2
Reliability function	$R(t) = 1 - F(t)$	–	survives beyond t	5.2.3
Conditional reliability	$R(t+\tau)/R(t)$	–	survives τ more, given age t	5.2.4
Hazard function	$\lambda(t) = f(t)/R(t)$	h^{-1}	failure rate of the survivors	5.3.1
Cumulative hazard	$\Lambda(t) = \int_0^t \lambda(u)\,\mathrm{d}u$	–	$R(t) = \mathrm{e}^{-\Lambda(t)}$	5.3.4
Interval failure probability	$1 - \exp[-\int_t^{t+\tau} \lambda\,\mathrm{d}u]$	–	fails within τ, given age t	5.4.1
Linear approximation	$\lambda\tau$, error $\approx \lambda\tau/2$	–	small intervals only	5.4.2
DFR, constant, IFR	sign of $\mathrm{d}\lambda/\mathrm{d}t$	–	early life, random, wear-out	5.5.1
Bathtub curve	sum of the three terms	–	conceptual, not a law	5.5.2
Memoryless property	$\Pr(T > s+t \mid T > s) = R(t)$	–	exponential model only	5.5.3
Mean time to failure	$\mathrm{MTTF} = \int_0^\infty R(t)\,\mathrm{d}t$	h	expected lifetime	5.6.1
Mean residual life	$\mathrm{MRL}(t) = \int_t^\infty R\,\mathrm{d}u/R(t)$	h	expected remaining life at age t	5.6.2
Proportional hazards	$\lambda_0(t)\exp[\boldsymbol{\gamma}^{\mathsf{T}}\mathbf{z}(t)]$	h^{-1}	hazard under operating conditions	5.6.3

Further reading. Rausand, Barros and Høyland, *System Reliability Theory*, third edition: lifetime distributions, reliability functions and failure-rate models. Modarres, Kaminskiy and Krivtsov, *Reliability Engineering and Risk Analysis*, third edition: statistical reliability and lifetime distributions. O'Connor and Kleyner, *Practical Reliability Engineering*, fifth edition: failure data and the practical interpretation of failure rates. The *NIST/SEMATECH e-Handbook of Statistical Methods*: reliability and survival analysis reference material.

5.8 Exercises

Exercise 5.1. (a) State what the hazard function $\lambda(t)$ represents, and why it is not the probability of failure at time t. (It is the instantaneous failure rate conditional on survival to t, in inverse time; the probability of failure at a single instant is zero, and only $\lambda(t)\,\Delta t$ over a short interval approximates a probability.) (b) For a constant hazard λ, give the exact probability of failure over an interval Δt for a unit that is working at its start. ($1 - \mathrm{e}^{-\lambda\Delta t}$; $\lambda\Delta t$ is only a short-interval approximation.) (c) Which property does the exponential lifetime model have: a hazard that increases with age, a survival probability over a fixed future duration that does not depend on the age already survived, or failure of every unit exactly at its MTTF? (The second, the memoryless property; only 36.8 % of units survive beyond the MTTF.)

Exercise 5.2. A component has $R(1000\,\mathrm{h}) = 0.90$ and $R(2000\,\mathrm{h}) = 0.72$. (a) Compute the probability that it survives the second 1000 h, given that it survived the first. ($0.72/0.90 = 0.80$.)

(b) Compute the conditional probability of failure in the second interval and the unconditional probability, for a new unit, of failure in the same interval, and explain why they differ. (0.20 and 0.18; the first is a fraction of the 90 % of units that survived to 1000 h, the second of all units.) (c) Compute $\Lambda(1000)$, $\Lambda(2000)$ and the average hazard over the second interval. (0.1054, 0.3285, and $0.2231/1000\,\mathrm{h} = 2.23 \times 10^{-4}\,\mathrm{h}^{-1}$.) (d) Over the first interval the survival probability was 0.90. Is the hazard constant, increasing or decreasing on average? (Increasing: the average hazard rose from 1.05×10^{-4} to $2.23 \times 10^{-4}\,\mathrm{h}^{-1}$.)

Exercise 5.3. A bearing has the hazard $\lambda(t) = at$ with $a = 5 \times 10^{-8}\,\mathrm{h}^{-2}$. (a) Derive $\Lambda(t)$ and $R(t)$, and compute $R(2000\,\mathrm{h})$. ($at^2/2$, $\mathrm{e}^{-at^2/2}$, $\mathrm{e}^{-0.1} = 0.9048$.) (b) Compute the probability of failure in the next 500 h for a bearing that has survived to 0, 2000 and 4000 h. (0.0062, 0.0547, 0.1008.) (c) Compute its MTTF with equation (5.19). ($\sqrt{\pi/(2a)} = 5605\,\mathrm{h}$.) (d) In general, if $\lambda(t)$ increases with age, what happens to the failure probability over the next fixed interval as the age at which the interval starts increases? (It increases with age, because the integral of the hazard over a fixed future interval increases when the hazard function itself is increasing.)

Exercise 5.4. A pump has a constant hazard $\lambda = 0.002\,\mathrm{h}^{-1}$ at its reference temperature of 40°C. (a) Compute the exact conditional failure probability and the linear approximation for intervals of 1, 100 and 500 h. (0.001998 against 0.002; 0.1813 against 0.2; 0.6321 against 1.0.) (b) Find the longest interval for which the relative error of the approximation, equation (5.17), stays below 1 %. ($\lambda\tau \approx 0.020$, so $\tau \approx 10\,\mathrm{h}$.) (c) Compute the MTTF, the probability of surviving beyond it, and the mean residual life of a pump that has run 300 h. (500 h, 0.368, 500 h.) (d) Under the doubling rule of equation (5.23), compute the hazard at 55°C and the reliability over a 100 h mission at that temperature. ($0.002 \times 2^{1.5} = 5.66 \times 10^{-3}\,\mathrm{h}^{-1}$; $\mathrm{e}^{-0.566} = 0.568$.)

Chapter 6

Exponential, Weibull, and Lognormal Lifetime Models

6.1 The choice of a lifetime distribution

Chapter 5 defined the four functions that describe a lifetime, the cumulative failure probability $F(t)$, the failure density $f(t)$, the reliability function $R(t)$ and the hazard function $\lambda(t)$, and showed that any one of them determines the others; in particular

$$R(t) = \exp\Big[-\int_0^t \lambda(u)\,\mathrm{d}u\Big].$$

These relations hold for a broad class of continuous lifetime distributions, but they do not say which mathematical form the hazard should take for a given component. That choice is the subject of this chapter. Its question is

how a probability distribution is selected to represent the lifetime of equipment, and what the selected distribution implies about how the equipment fails as it ages.

The chapter answers it in five steps. It explains why a distribution is needed at all and introduces three standard families. It derives the exponential distribution from a constant hazard and states the assumption that model makes. It derives the Weibull distribution from a power-law cumulative hazard and interprets its shape and scale parameters. It defines the lognormal distribution from a normally distributed logarithm of the lifetime and shows that its hazard rises and then falls. Finally, it compares the three models, shows how their parameters are estimated from failure times, and sets out a procedure for selecting among them that weighs statistical fit against the credibility of the prediction the model is asked to make.

Jargon: lifetime distribution, parameter, exponential, Weibull, lognormal, characteristic life, gamma function, standard normal distribution, likelihood, maximum likelihood estimation, probability plot, AIC, right-censoring

A *lifetime distribution* is a probability distribution of the time to failure T, written as a formula with a small number of *parameters*, constants that are fixed for a population and estimated from data. The *exponential distribution* has one parameter, the constant failure rate λ. The *Weibull distribution* has a *shape parameter* β, which sets how the hazard changes with age, and a *scale parameter* η, in hours, also called the *characteristic life*, the age by which 63.2 % of the population has failed. The *lognormal distribution* is the distribution of a lifetime whose natural logarithm is normally distributed; its parameters are the mean μ and the standard deviation σ of $\ln T$. The *gamma function* $\Gamma(z) = \int_0^\infty u^{z-1}\mathrm{e}^{-u}\,\mathrm{d}u$ appears in the Weibull mean. The *standard normal distribution* has mean zero and variance one; its density is written ϕ and its cumulative distribution function Φ. The *likelihood* $L(\boldsymbol{\theta})$ is the probability density of the observed data regarded as a function of the parameter vector $\boldsymbol{\theta}$, and *maximum likelihood estimation* (MLE) takes as the estimate the parameter value that maximizes it. A *probability plot* is a graph on which data from a given distribution fall near a straight line. The *Akaike information criterion* (AIC) is a score for comparing fitted models that penalizes the number of parameters. A *right-censored* observation is the lifetime of a unit that had not failed when observation stopped, so that only a lower bound of its T is known.

Key idea

The hazard function is the most informative description of a lifetime model: the exponential distribution fixes it constant, the Weibull lets it decrease, stay constant or increase according to its shape parameter, and the lognormal makes it rise and then fall. Two distributions with the same mean lifetime can give very different reliabilities over a mission, so a model is selected by its hazard behavior, the physics of failure, the data and the prediction it must support, never by its mean alone.

This section states why a lifetime needs a probability distribution rather than a single number, and introduces the three families that the rest of the chapter derives and compares.

6.1.1 Variation of lifetimes in a population

Consider again 1000 industrial cooling-water pumps placed into service under comparable operating conditions, the population of Chapter 5. They do not all fail at the same age. Some may fail after 200 h, others after 2000 h, and others after 20 000 h. Pumps of one nominal design still differ in manufacturing imperfections, material properties, operating stress, initial condition and installation quality, and in many other factors that no record captures. Even components that share a design can therefore have lifetimes that vary by two orders of magnitude.

This variation is represented by treating the time to failure T as a random variable, and a lifetime distribution assigns probabilities to its possible values. Once a distribution is chosen and its parameters known, every question of Chapter 5 has a numerical answer: the probability of failure within the first 1000 h, $\Pr(T < 1000)$; the probability that a pump which has already run 3000 h survives to 5000 h, $\Pr(T > 5000 \mid T > 3000)$; and the expected lifetime $\mathbb{E}[T]$. Two distributions can have the same mean lifetime and still assign very different probabilities to early and to late failure. That difference matters whenever equipment is selected or maintenance planned, and Section 6.5 quantifies it.

6.1.2 Three families of lifetime distributions

Three families cover most of classical reliability practice. The *exponential* distribution has a constant hazard. The *Weibull* distribution has a hazard that decreases, stays constant or increases with age, depending on one parameter. The *lognormal* distribution arises when lifetime results from many factors that combine multiplicatively, and its hazard rises to a maximum and then declines. Table 6.1 compares them property by property: each row is a property and each column a family. The rows give the parameters, the reliability function and the hazard, the qualitative behavior of the hazard, whether the model has the memoryless property of Chapter 5, the mean lifetime, and the usual motivation for using the model. Every entry is derived in the sections that follow.

Table 6.1: The three lifetime distributions of this chapter. Φ and ϕ are the standard normal distribution function and density, and Γ is the gamma function.

Property	Exponential	Weibull	Lognormal
Parameters	rate λ	shape β, scale η	log-location μ, log-scale σ of $\ln T$
Reliability $R(t)$	$\mathrm{e}^{-\lambda t}$	$\mathrm{e}^{-(t/\eta)^{\beta}}$	$1-\Phi\big((\ln t-\mu)/\sigma\big)$
Hazard $\lambda(t)$	λ	$(\beta/\eta)(t/\eta)^{\beta-1}$	$f(t)/R(t)$, no simpler form
Hazard behavior	constant	decreasing ($\beta<1$), constant ($\beta=1$), increasing ($\beta>1$)	zero at $t=0$, rises to a maximum, then decreases toward zero
Represents aging	no age dependence	yes, through the shape	a peaked failure propensity
Memoryless	yes	only when $\beta=1$	no
Mean lifetime	$1/\lambda$	$\eta\,\Gamma(1+1/\beta)$	$\mathrm{e}^{\mu+\sigma^2/2}$
Typical motivation	random, age-independent failures	flexible empirical lifetime modeling	multiplicative lifetime variability

One entry deserves emphasis at the outset. The lognormal hazard starts at zero, rises to a maximum and eventually decreases toward zero; it is not, in general, a model of monotone aging. A density curve of a lognormal lifetime looks like a skewed hump and does not reveal this, which is one reason the hazard function is often more informative than the density when models are compared. All three are statistical models, not physical laws. Their usefulness depends on whether they represent adequately the failure processes that have been observed and those that are expected.

6.2 The exponential distribution

The exponential distribution is the simplest lifetime model and the one on which much of classical system reliability is built. This section derives it from a constant hazard, interprets its rate parameter, works the pump example and states the assumption the model makes.

6.2.1 Derivation from a constant hazard

Suppose that a component has a hazard that does not depend on its age,

$$\lambda(t)=\lambda, \qquad \lambda>0. \tag{6.1}$$

The reliability function follows from the hazard by the relation of Chapter 5, which gives $R(t)$ as the exponential of minus the integral of the hazard, $\exp[-\int_0^t \lambda(u)\,\mathrm{d}u]$. With a constant hazard

the integral is $\int_0^t \lambda \, \mathrm{d}u = \lambda t$, which is also the cumulative hazard $\Lambda(t)$, so

$$R(t) = \mathrm{e}^{-\lambda t}. \tag{6.2}$$

The cumulative failure probability is the complement,

$$F(t) = 1 - \mathrm{e}^{-\lambda t}, \tag{6.3}$$

and differentiating it with respect to t gives the failure density,

$$f(t) = \frac{\mathrm{d}F(t)}{\mathrm{d}t} = \lambda \, \mathrm{e}^{-\lambda t}. \tag{6.4}$$

As a check, the ratio of density to reliability returns the hazard that was assumed,

$$\lambda(t) = \frac{f(t)}{R(t)} = \frac{\lambda \, \mathrm{e}^{-\lambda t}}{\mathrm{e}^{-\lambda t}} = \lambda.$$

Equations (6.1) to (6.4) are four statements of one model: a constant hazard and an exponential reliability function imply each other. The exponential distribution is a standard result of reliability theory; the derivation shows that it is nothing more than the assumption of an age-independent hazard.

6.2.2 Meaning of the rate parameter

The parameter λ has units of inverse time. Its reciprocal is the mean time to failure: by the identity $\mathrm{MTTF} = \int_0^\infty R(t) \, \mathrm{d}t$ of Chapter 5,

$$\mathrm{MTTF} = \int_0^\infty \mathrm{e}^{-\lambda t} \, \mathrm{d}t = \frac{1}{\lambda}. \tag{6.5}$$

For example, $\lambda = 0.0002\,\mathrm{h}^{-1}$ corresponds to $\mathrm{MTTF} = 1/0.0002\,\mathrm{h}^{-1} = 5000\,\mathrm{h}$.

This does not mean that a component normally fails at about 5000 h. The exponential distribution is skewed to the right: its density is largest at $t = 0$ and decreases steadily, so short lifetimes are the most probable and a long tail of survivors lasts far beyond the mean. Two numbers make this concrete. The fraction that fails before the mean is $F(1/\lambda) = 1 - \mathrm{e}^{-1} = 0.632$, so 63.2 % of the population fails before reaching the MTTF. The *median life* t_{50}, the age by which half the population has failed, solves $\mathrm{e}^{-\lambda t_{50}} = 0.5$, so

$$t_{50} = \frac{\ln 2}{\lambda} = 0.693\,\mathrm{MTTF},$$

which is 3466 h for $\lambda = 0.0002\,\mathrm{h}^{-1}$. The standard deviation of the exponential lifetime equals its mean, $1/\lambda$, so a spread as large as the mean itself is built into the model.

Figure 6.1 shows these properties for $\lambda = 0.0002\,\mathrm{h}^{-1}$. In both panels the horizontal axis is operating time in hours. In panel (a) the vertical axis is the failure density, in units of $10^{-4}\,\mathrm{h}^{-1}$; the shaded area to the left of the mean is 0.632, and the dashed lines mark the median at 3466 h and the mean at 5000 h. In panel (b) the vertical axis is the reliability; the dots mark the values used in the next subsection, $R(2000) = 0.670$, and the survival probability at the mean, $R(5000) = 0.368$. The density falls from its largest value at $t = 0$, and the median lies well to the left of the mean, which is the signature of a right-skewed distribution.

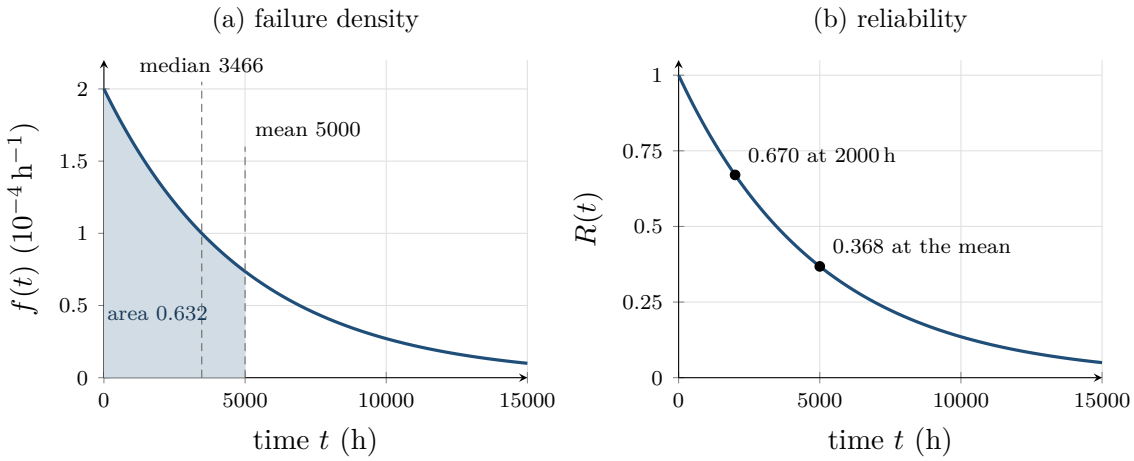


Figure 6.1: The exponential lifetime with $\lambda = 2\times10^{-4}\,\mathrm{h}^{-1}$, illustrative. (a) Density $f(t) = \lambda\,\mathrm{e}^{-\lambda t}$; the shaded area, $F(5000) = 1-\mathrm{e}^{-1} = 0.632$, is the fraction failing before the mean $1/\lambda = 5000\,\mathrm{h}$; the median is $\ln 2/\lambda = 3466\,\mathrm{h}$. (b) Reliability $R(t) = \mathrm{e}^{-\lambda t}$.

6.2.3 Worked example: the cooling-water pump

Consider a cooling-water pump with constant hazard $\lambda = 0.0002\,\mathrm{h}^{-1}$, an illustrative value that gives an MTTF of 5000 h. The probability that the pump survives a 2000 h mission is

$$R(2000) = \mathrm{e}^{-0.0002\times2000} = \mathrm{e}^{-0.4} = 0.6703,$$

a 67.03 % probability of completing the mission without failure. The probability of failure during the mission is the complement,

$$F(2000) = 1 - 0.6703 = 0.3297,$$

or 32.97 %. For comparison, the pump of Chapter 5, with $\lambda = 0.0005\,\mathrm{h}^{-1}$ and MTTF = 2000 h, has $R(2000) = \mathrm{e}^{-1} = 0.3679$ over the same mission: a rate two and a half times larger changes the exponent by the same factor, and the mission reliability from 0.67 to 0.37.

6.2.4 The memoryless assumption

Suppose now that the pump has already operated for 8000 h without failure. The probability that it survives a further 2000 h is the conditional reliability

$$\Pr(T > 10\,000 \mid T > 8000) = \frac{R(10\,000)}{R(8000)} = \frac{\mathrm{e}^{-\lambda\times10\,000}}{\mathrm{e}^{-\lambda\times8000}} = \mathrm{e}^{-\lambda\times2000} = 0.6703,$$

the same value as for a new pump. The additional survival probability does not depend on the age already reached. This is the memoryless property of Chapter 5,

$$\Pr(T > s+t \mid T > s) = \Pr(T > t) \qquad \text{for all } s,t \geq 0, \tag{6.6}$$

and the exponential is the only continuous lifetime distribution that has it.

Memorylessness is the model's most important assumption, and it is an assumption about the physics. It is reasonable when failures are caused by events that arrive at random and do not depend on the condition of the item: an external shock, a voltage transient, a foreign object in the flow, a human error during an intervention. It is questionable whenever the mechanism involves accumulated wear or damage, as fatigue, corrosion and insulation aging do, because then a pump that has run 8000 h is not as good as new, and a model that treats it as new understates the risk of its next interval.

Principle 6.1. *The exponential distribution is the constant-hazard model. It is convenient and often adequate for randomly arriving failures, but it is memoryless: it cannot describe physical aging through its hazard, and applied to a wearing component it understates the risk of an old unit.*

6.3 The Weibull distribution

The Weibull distribution adds one parameter to the exponential and with it the ability to represent a hazard that changes with age. This section defines the distribution, derives its functions from a power-law cumulative hazard, interprets its two parameters, and works an example of aging bearings and a table of conditional failure probabilities for several shapes.

6.3.1 Motivation and definition

Many engineering components fail in ways that depend on age. Three mechanisms illustrate the range. A manufacturing defect causes early failure and preferentially removes the weaker units from the population. Random external events cause failures without systematic dependence on age. Mechanical wear and fatigue make the propensity to fail grow with accumulated use. These are the decreasing, constant and increasing hazards of Chapter 5. The exponential distribution represents the second pattern only; it cannot represent the first or the third, because its hazard does not depend on age.

The two-parameter Weibull distribution introduces a shape parameter that allows all three. Its parameters are the *shape* β, dimensionless, which controls how the hazard changes with age, and the *scale* η, in hours, which sets the characteristic lifetime; both are positive,

$$\beta > 0, \qquad \eta > 0.$$

The Weibull cumulative distribution function is

$$F(t) = 1 - \exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right], \qquad t \geq 0. \tag{6.7}$$

The Weibull model is a standard result of reliability theory; the next subsection shows where its form comes from.

6.3.2 Reliability, density and hazard from a power-law cumulative hazard

The structure of equation (6.7) is clearest through the cumulative hazard. Since $F(t) = 1 - \mathrm{e}^{-\Lambda(t)}$ for any lifetime (Chapter 5), the Weibull distribution is exactly the lifetime whose cumulative hazard is a power of age,

$$\Lambda(t) = \left(\frac{t}{\eta}\right)^{\beta}. \tag{6.8}$$

Every other function follows from this one assumption in three steps.

Step 1: reliability. From $R(t) = 1 - F(t)$, or equally from $R(t) = \mathrm{e}^{-\Lambda(t)}$,

$$R(t) = \exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right]. \tag{6.9}$$

Step 2: density. The density is $f(t) = -\mathrm{d}R/\mathrm{d}t$. By the chain rule, the derivative of $\exp[-g(t)]$ is $-g'(t)\exp[-g(t)]$, with $g(t) = (t/\eta)^{\beta}$ and $g'(t) = \beta\,(t/\eta)^{\beta-1} \cdot (1/\eta)$. Hence

$$f(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right]. \tag{6.10}$$

Step 3: hazard. The hazard is the ratio of density to reliability, and the exponential factors cancel:

$$\lambda(t) = \frac{f(t)}{R(t)} = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}. \tag{6.11}$$

The same result follows directly as the derivative of the cumulative hazard, $\lambda(t) = \mathrm{d}\Lambda/\mathrm{d}t$, which is $g'(t)$ above. The units are consistent: t/η is dimensionless, so λ has the units of $1/\eta$, inverse hours.

Equation (6.11) is the most revealing equation of the model. The hazard is a power of age, and the exponent $\beta - 1$ determines how it changes with time: it falls when $\beta < 1$, is constant when $\beta = 1$ and rises when $\beta > 1$. The aging component of Chapter 5, with the linear hazard $\lambda(t) = at$ and $a = 2 \times 10^{-7}\,\mathrm{h}^{-2}$, is the Weibull case $\beta = 2$, for which $\lambda(t) = (2/\eta^2)\,t$; matching $2/\eta^2 = a$ gives $\eta = \sqrt{2/a} = 3162\,\mathrm{h}$.

6.3.3 The shape parameter

The shape parameter divides the Weibull family into three regimes.

$\beta < 1$: decreasing hazard. The propensity to fail of the survivors declines with age. This is the regime often associated with early-life failure populations, in which defective units fail first and leave a stronger population behind.

$\beta = 1$: constant hazard. Equation (6.11) reduces to $\lambda(t) = 1/\eta$, and equation (6.9) to $\mathrm{e}^{-t/\eta}$: the Weibull distribution becomes the exponential distribution with $\lambda = 1/\eta$. The exponential model is the special case of the Weibull model with unit shape, and only that case is memoryless.

$\beta > 1$: increasing hazard. The propensity to fail grows with age, which can represent wear-out or accumulated damage. For $\beta = 2$ the hazard grows linearly with age; for $\beta = 3$ it grows with the square of age; for $1 < \beta < 2$ it grows, but ever more slowly.

Figure 6.2 compares the hazards of four Weibull distributions with the same scale, $\eta = 5000\,\mathrm{h}$, and shapes 0.5, 1, 2 and 3. The horizontal axis is operating time in hours; the vertical axis is the hazard in units of $10^{-3}\,\mathrm{h}^{-1}$, that is, failures per 1000 operating hours among the survivors. The shape alone changes the pattern completely. At $t = 500\,\mathrm{h}$ the hazard is 0.316, 0.200, 0.040 and 0.006 per 1000 h for $\beta = 0.5$, 1, 2 and 3; at $t = 10\,000\,\mathrm{h}$ it is 0.071, 0.200, 0.800 and 2.400. The order of the four curves is reversed between early and late life. At the common scale $t = \eta$ every hazard equals β/η, which is 0.1, 0.2, 0.4 and 0.6 per 1000 h. The curves cross at ages fixed by the ratio t/η: the $\beta = 2$ curve crosses the constant one at $t = \eta/2 = 2500\,\mathrm{h}$, and the $\beta = 0.5$ curve falls below it at $t = \eta/4 = 1250\,\mathrm{h}$.

For $\beta < 1$ the hazard diverges as t approaches zero: at $t = 1\,\mathrm{h}$ the $\beta = 0.5$ curve of Figure 6.2 is already $7.1 \times 10^{-3}\,\mathrm{h}^{-1}$, and it grows without limit below that. The distribution remains well defined, because the cumulative hazard $(t/\eta)^\beta$ tends to zero as $t \to 0$ for any positive β, so that $F(0) = 0$ and every probability is finite. In practice, a fitted model with a very high early hazard requires scrutiny: it may describe a genuine population of weak units, or it may be an artifact of how the start of life, the installation records or the early failures were defined.

Principle 6.2. *The Weibull shape parameter states how the propensity to fail changes with age: decreasing for $\beta < 1$, constant for $\beta = 1$, where the model is exponential, and increasing for $\beta > 1$. Two Weibull models with the same scale and different shapes describe different failure processes.*

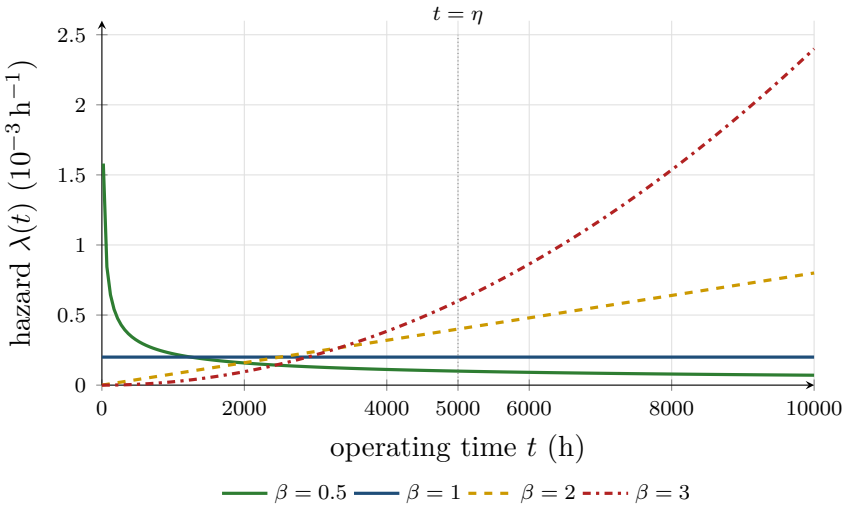


Figure 6.2: Weibull hazard functions with common scale $\eta = 5000\,\mathrm{h}$, $\lambda(t) = (\beta/\eta)(t/\eta)^{\beta-1}$ (equation (6.11)), for $\beta = 0.5$, 1, 2 and 3, in failures per 1000 h among survivors. The $\beta = 0.5$ curve is drawn from $t = 20\,\mathrm{h}$; it grows without bound as $t \to 0$. Illustrative parameters.

6.3.4 The scale parameter and the mean

The scale parameter η is called the *characteristic life*. Its meaning follows by setting $t = \eta$ in equation (6.9):

$$R(\eta) = \exp\left[-\left(\frac{\eta}{\eta}\right)^{\beta}\right] = \mathrm{e}^{-1} = 0.3679, \qquad F(\eta) = 1 - \mathrm{e}^{-1} = 0.6321.$$

About 63.2 % of the population is expected to fail by age η, whatever the shape: η is the 63.2 % life, the age at which the cumulative hazard reaches one. Figure 6.3 shows this. The horizontal axis is operating time in hours and the vertical axis is the reliability, for $\eta = 5000\,\mathrm{h}$ and the four shapes of Figure 6.2. All four curves pass through the same point, $(5000\,\mathrm{h}, 0.368)$, marked by the dotted lines. Before η the larger shapes keep the reliability higher, because their hazard is low when the item is young; after η they drop it faster, because their hazard is then high. The $\beta = 3$ curve is almost a step that falls sharply around η, while the $\beta = 0.5$ curve loses a third of the population in the first 1000 h and keeps a long tail of survivors.

The scale parameter is in general not the mean lifetime. The mean is derived from $\mathrm{MTTF} = \int_0^\infty R(t)\,\mathrm{d}t$:

$$\mathrm{MTTF} = \int_0^\infty \exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right] \mathrm{d}t.$$

Substitute $u = (t/\eta)^{\beta}$, so that $t = \eta\, u^{1/\beta}$ and $\mathrm{d}t = (\eta/\beta)\, u^{1/\beta-1}\,\mathrm{d}u$; the limits $t = 0$ and $t \to \infty$ become $u = 0$ and $u \to \infty$. Then

$$\mathrm{MTTF} = \frac{\eta}{\beta}\int_0^\infty u^{1/\beta-1}\,\mathrm{e}^{-u}\,\mathrm{d}u = \frac{\eta}{\beta}\,\Gamma\left(\frac{1}{\beta}\right),$$

by the definition of the gamma function, $\Gamma(z) = \int_0^\infty u^{z-1}\mathrm{e}^{-u}\,\mathrm{d}u$. The gamma function satisfies $\Gamma(z+1) = z\,\Gamma(z)$, which follows from an integration by parts; with $z = 1/\beta$ it gives $(1/\beta)\,\Gamma(1/\beta) = \Gamma(1 + 1/\beta)$, and therefore

$$\mathrm{MTTF} = \mathbb{E}[T] = \eta\,\Gamma\left(1 + \frac{1}{\beta}\right). \tag{6.12}$$

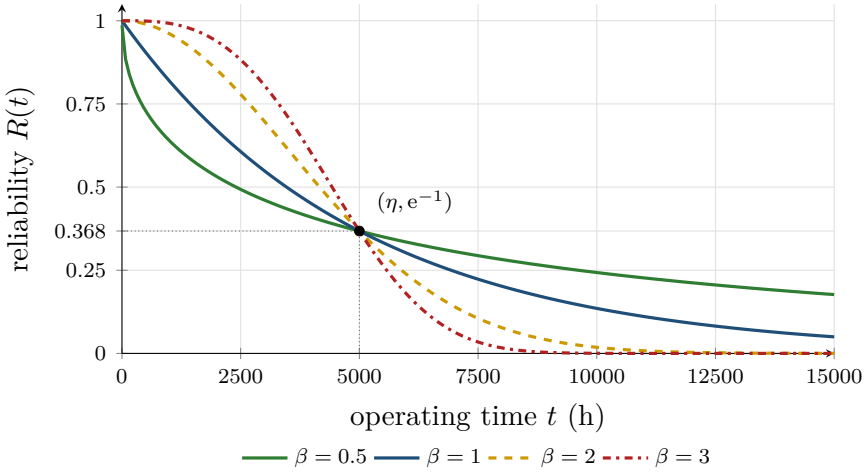


Figure 6.3: Weibull reliability functions $R(t) = \exp[-(t/\eta)^{\beta}]$ (equation (6.9)) with $\eta = 5000\,\mathrm{h}$ and $\beta = 0.5$, 1, 2, 3, illustrative. Every curve passes through $R(\eta) = \mathrm{e}^{-1} = 0.368$: the scale is the 63.2 % life for any shape.

For $\beta = 1$, $\Gamma(2) = 1$ and MTTF $= \eta$, the exponential result. For $\beta = 2$, $\Gamma(1.5) = \sqrt{\pi}/2 = 0.8862$ and MTTF $= 0.8862\,\eta$. For $\beta = 3$, $\Gamma(4/3) = 0.8930$ and MTTF $= 0.8930\,\eta$. For $\beta = 0.5$, $\Gamma(3) = 2$ and the mean is twice the characteristic life, carried by the long tail of survivors. The median follows from $R(t_{50}) = 0.5$ as $t_{50} = \eta\,(\ln 2)^{1/\beta}$.

Table 6.2 evaluates these quantities for $\eta = 5000\,\mathrm{h}$ and seven shapes. The columns are the gamma factor, the mean, the median, and the probability of surviving beyond the mean, $R(\mathrm{MTTF}) = \exp[-\Gamma(1 + 1/\beta)^{\beta}]$. For shapes of one and above, the mean lies between $0.886\,\eta$ and η, the factor having its minimum near $\beta = 2.2$; for shapes below one it exceeds η and grows quickly. The probability of surviving beyond the mean rises with the shape, from 0.24 at $\beta = 0.5$ to about one half at $\beta = 4$, where the distribution is nearly symmetric and the median overtakes the mean. The distinction between η, the mean and the median matters whenever a model is fitted to failure data, where the three are easily confused.

Table 6.2: Mean, median and survival beyond the mean of Weibull distributions with $\eta = 5000\,\mathrm{h}$, from equation (6.12) and $t_{50} = \eta(\ln 2)^{1/\beta}$.

Shape β	$\Gamma(1 + 1/\beta)$	MTTF (h)	median t_{50} (h)	R(MTTF)
0.5	2.0000	10 000	2402	0.2431
0.7	1.2658	6329	2962	0.3075
1	1.0000	5000	3466	0.3679
1.5	0.9027	4514	3916	0.4241
2	0.8862	4431	4163	0.4559
3	0.8930	4465	4425	0.4906
4	0.9064	4532	4562	0.5092

6.3.5 Worked example: aging pump bearings

Suppose that the lifetime of the bearings of a pump is modeled as

$$T \sim \mathrm{Weibull}(\beta = 2,\ \eta = 5000\,\mathrm{h}),$$

an illustrative model of wear-out. The reliability function is

$$R(t) = \exp\left[-\left(\frac{t}{5000}\right)^2\right],$$

with t in hours, and the hazard, $\lambda(t) = (2/5000)(t/5000) = 8 \times 10^{-8}\, t\,\mathrm{h}^{-1}$, grows linearly: $2.4 \times 10^{-4}\,\mathrm{h}^{-1}$ at 3000 h and $4.8 \times 10^{-4}\,\mathrm{h}^{-1}$ at 6000 h. The mean life is $5000\,\Gamma(1.5) = 4431$ h. Four questions are answered.

A. Survival of 3000 h**.**

$$R(3000) = \exp\left[-(3000/5000)^2\right] = \mathrm{e}^{-0.36} = 0.6977.$$

B. Survival of 6000 h**.**

$$R(6000) = \exp\left[-(6000/5000)^2\right] = \mathrm{e}^{-1.44} = 0.2369.$$

C. Survival of another 1000 h **after** 3000 h**.** For a bearing that has already survived 3000 h,

$$\begin{aligned}\Pr(T > 4000 \mid T > 3000) = \frac{R(4000)}{R(3000)} &= \exp\left[-\left(\frac{4000}{5000}\right)^2 + \left(\frac{3000}{5000}\right)^2\right] \\ &= \mathrm{e}^{-(0.64-0.36)} = \mathrm{e}^{-0.28} = 0.7558.\end{aligned}$$

The conditional probability of failure over the next 1000 h is $1 - 0.7558 = 0.2442$, or 24.42 %.

D. Survival of another 1000 h **after** 6000 h**.** For a bearing that has already survived 6000 h,

$$\begin{aligned}\Pr(T > 7000 \mid T > 6000) &= \exp\left[-\left(\frac{7000}{5000}\right)^2 + \left(\frac{6000}{5000}\right)^2\right] \\ &= \mathrm{e}^{-(1.96-1.44)} = \mathrm{e}^{-0.52} = 0.5945,\end{aligned}$$

and the conditional probability of failure over the next 1000 h is $1 - 0.5945 = 0.4055$, or 40.55 %.

The older bearing has a substantially greater probability of failing over the same future interval, 40.55 % against 24.42 %. The reason is the increasing hazard of $\beta = 2$: the exponent in each conditional reliability is the increase of the cumulative hazard over the interval, $(t + 1000)^2/\eta^2 - t^2/\eta^2 = (2000\,t + 10^6)/\eta^2$, which grows linearly with the age t at which the interval starts. Under the exponential model of Section 6.2.4 the same calculation would give one value at every age.

6.3.6 Conditional failure probability for several shapes

The bearing example fixes $\beta = 2$. How the answers change with the shape is seen by evaluating the same model over a grid. For a Weibull lifetime with scale η, the probability of surviving to age t is equation (6.9), and the probability of failing within a further mission of length τ, given survival to age t and no other information about the item's condition, is

$$\Pr(t < T \le t + \tau \mid T > t) = 1 - \exp\left\{-\left[\left(\frac{t+\tau}{\eta}\right)^\beta - \left(\frac{t}{\eta}\right)^\beta\right]\right\}. \tag{6.13}$$

Table 6.3 evaluates both for $\eta = 5000$ h and $\tau = 1000$ h, with the shape varying from 0.5 to 4: $\beta = 0.7$ represents early failures, $\beta = 1$ a constant hazard, $\beta = 2$ the bearing and $\beta = 3$ wear-out; 0.5, 1.5 and 4 are added to show the trend. The first four numerical columns are the

Table 6.3: Weibull reliability and conditional failure probability over the next 1000 h, equations (6.9) and (6.13), for $\eta = 5000$ h. The row $\beta = 2$ is the bearing of Section 6.3.5.

	reliability $R(t)$				failure in next 1000 h, given age		
β	1000 h	3000 h	5000 h	10 000 h	age 0	age 3000 h	age 6000 h
0.5	0.6394	0.4609	0.3679	0.2431	0.3606	0.1129	0.0840
0.7	0.7232	0.4969	0.3679	0.1970	0.2768	0.1445	0.1214
1	0.8187	0.5488	0.3679	0.1353	0.1813	0.1813	0.1813
1.5	0.9144	0.6283	0.3679	0.0591	0.0856	0.2218	0.2896
2	0.9608	0.6977	0.3679	0.0183	0.0392	0.2442	0.4055
3	0.9920	0.8057	0.3679	0.0003	0.0080	0.2562	0.6380
4	0.9984	0.8784	0.3679	$< 10^{-4}$	0.0016	0.2442	0.8293

reliability at four ages; the last three are the conditional failure probability over the next 1000 h for a unit of age 0, 3000 and 6000 h. One row of the arithmetic, for $\beta = 3$ at age 3000 h: the increase of the cumulative hazard is $(4000/5000)^3 - (3000/5000)^3 = 0.512 - 0.216 = 0.296$, and the conditional failure probability is $1 - \mathrm{e}^{-0.296} = 0.2562$.

Three trends appear in the table. First, the column at 5000 h is constant, the characteristic-life property of Section 6.3.4. Second, the shape changes the qualitative behavior of the conditional failure probability: for $\beta < 1$ it falls with age, from 0.2768 to 0.1214 for $\beta = 0.7$; for $\beta = 1$ it is $1 - \mathrm{e}^{-0.2} = 0.1813$ at every age; for $\beta > 1$ it rises, from 0.0080 to 0.6380 for $\beta = 3$. Third, the effect of a larger shape depends on where the interval lies relative to η. For a new unit, a larger shape lowers the probability of failure in the first 1000 h; for a unit aged 6000 h, beyond η, it raises it; and at 3000 h, before η, the probability first rises with the shape and then falls again. The equal values at 3000 h for $\beta = 2$ and $\beta = 4$ are a coincidence of these numbers: $0.8^4 - 0.6^4 = (0.8^2 - 0.6^2)(0.8^2 + 0.6^2)$, and $0.8^2 + 0.6^2 = 1$.

Figure 6.4 draws the last three columns as continuous functions of age for four of the shapes. The horizontal axis is the age at the start of the 1000 h interval and the vertical axis is the conditional failure probability. The constant-hazard line is flat at 0.1813; the $\beta = 0.7$ curve falls; the $\beta = 2$ and $\beta = 3$ curves rise, the second more steeply. The two dots are questions C and D of the bearing example. The figure is the graphical form of the question of Chapter 5, whether an old unit is more likely to fail in the next interval than a new one: the answer is fixed by the shape parameter.

6.4 The lognormal distribution

The lognormal distribution describes a lifetime produced by many factors that multiply. This section motivates it, defines its density, distribution, reliability and hazard, shows that the hazard rises and then falls, derives the median, mean and variance, and works an example of electronic equipment.

6.4.1 Multiplicative variability

Suppose that a component's lifetime is influenced by several factors that combine multiplicatively,

$$T = T_0 \, X_1 \, X_2 \, X_3,$$

where T_0 is a nominal lifetime and X_1, X_2 and X_3 are positive random factors associated with material quality, operating environment and manufacturing variability; a factor of 0.8 shortens the life by 20 % and a factor of 1.3 lengthens it by 30 %. Taking logarithms turns the product into a sum,

$$\ln T = \ln T_0 + \ln X_1 + \ln X_2 + \ln X_3.$$

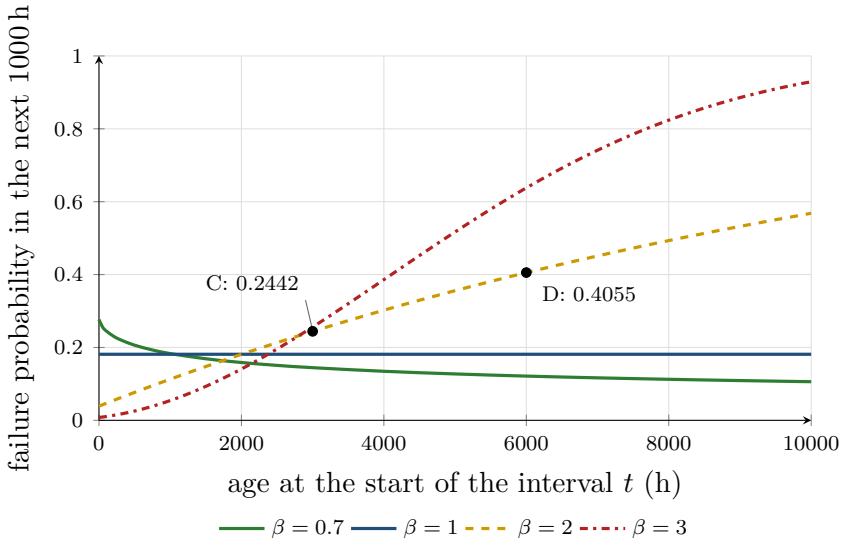


Figure 6.4: Conditional probability of failure in the next 1000 h as a function of the age already survived, equation (6.13) with $\eta = 5000$ h, $\tau = 1000$ h and $\beta = 0.7$, 1, 2, 3, illustrative. Dots: questions C and D of the bearing example ($\beta = 2$).

A sum of several independent effects, none of which dominates, tends to be approximately normally distributed, by the central limit theorem. If the sum of the log-effects is approximately normal, then $\ln T$ is approximately normal, and the lifetime T itself is approximately lognormal. This is one motivation for the model, not a proof that it holds. The lognormal has been found useful for some fatigue and degradation processes, for repair durations and in accelerated life testing. A right-skewed set of lifetimes does not by itself establish that it is appropriate, since the exponential and most Weibull distributions are right-skewed as well.

6.4.2 Density, distribution and reliability

A lifetime T is lognormally distributed when its natural logarithm is normally distributed,

$$\ln T \sim \mathcal{N}(\mu, \sigma^2), \tag{6.14}$$

where μ is the mean of the logarithm of the lifetime and $\sigma > 0$ its standard deviation. Strictly, a logarithm is taken of a dimensionless number, so with hours as the reference unit $\ln T$ means $\ln[T/(1\,\text{h})]$, and μ is in natural-log hours; changing the time unit shifts μ and leaves σ unchanged. The parameter μ is not the arithmetic mean lifetime, and e^{μ} is not the mean either, as Section 6.4.4 shows.

Remark 6.1. In this chapter μ and σ are the mean and the standard deviation of $\ln T$. The symbol μ here is not the repair rate μ of the notation table, which appears in the chapters on repairable systems, and σ is not a stress amplitude.

Cumulative distribution. Let $Z = (\ln T - \mu)/\sigma$, the standardized logarithm, which has the standard normal distribution. Because the logarithm is increasing, $T \le t$ exactly when $\ln T \le \ln t$, and therefore

$$F(t) = \Pr(\ln T \le \ln t) = \Pr\left(Z \le \frac{\ln t - \mu}{\sigma}\right) = \Phi\left(\frac{\ln t - \mu}{\sigma}\right), \qquad t > 0, \tag{6.15}$$

where $\Phi(z) = \int_{-\infty}^{z} \phi(v)\,\mathrm{d}v$ is the standard normal cumulative distribution function and $\phi(v) = \mathrm{e}^{-v^2/2}/\sqrt{2\pi}$ the standard normal density. Φ has no closed form and is evaluated from tables or software.

Density. Differentiating equation (6.15) by the chain rule, with $\mathrm{d}\Phi(z)/\mathrm{d}z = \phi(z)$ and $\mathrm{d}z/\mathrm{d}t = 1/(\sigma t)$ for $z = (\ln t - \mu)/\sigma$,

$$f(t) = \frac{1}{t\sigma}\,\phi\!\left(\frac{\ln t - \mu}{\sigma}\right) = \frac{1}{t\sigma\sqrt{2\pi}}\exp\left[-\frac{(\ln t - \mu)^2}{2\sigma^2}\right], \qquad t > 0. \tag{6.16}$$

The factor $1/t$ is the Jacobian of the change of variable from $\ln t$ to t; it is what makes the density skewed to the right.

Reliability. From $R(t) = 1 - F(t)$,

$$R(t) = 1 - \Phi\!\left(\frac{\ln t - \mu}{\sigma}\right). \tag{6.17}$$

6.4.3 The lognormal hazard function

The hazard is the ratio of equations (6.16) and (6.17),

$$\lambda(t) = \frac{f(t)}{R(t)} = \frac{\dfrac{1}{t\sigma}\,\phi\!\left(\dfrac{\ln t - \mu}{\sigma}\right)}{1 - \Phi\!\left(\dfrac{\ln t - \mu}{\sigma}\right)}. \tag{6.18}$$

Unlike the exponential and Weibull hazards, it does not simplify to a power of age, and its shape is found by examining the formula at its two ends. As $t \to 0$, $z = (\ln t - \mu)/\sigma \to -\infty$, the denominator tends to one and the numerator to zero, because $\phi(z)$ decays like $\mathrm{e}^{-z^2/2}$, faster than $1/t = \mathrm{e}^{-\mu-\sigma z}$ grows; the hazard starts at zero. As $t \to \infty$, $z \to +\infty$ and the standard approximation of the normal tail, $1 - \Phi(z) \approx \phi(z)/z$ for large z, gives

$$\lambda(t) \approx \frac{z}{\sigma t} = \frac{\ln t - \mu}{\sigma^2\, t} \;\longrightarrow\; 0,$$

because $\ln t$ grows more slowly than t. A hazard that is zero at both ends and positive in between must rise to a maximum and then decline. It is a standard result that the lognormal hazard has exactly one maximum.

Figure 6.5 draws the density and the hazard from equations (6.16) and (6.18) for two lognormal distributions with the same median, $\mathrm{e}^{\mu} = 5000\,\mathrm{h}$, and $\sigma = 0.5$ and 1.0. In both panels the horizontal axis is operating time in hours. In panel (a) the vertical axis is the density in $10^{-4}\,\mathrm{h}^{-1}$; in panel (b) it is the hazard in $10^{-3}\,\mathrm{h}^{-1}$, and the dots mark the maxima. With $\sigma = 0.5$ the density is a concentrated hump with its mode at $\mathrm{e}^{\mu-\sigma^2} = 3894\,\mathrm{h}$, and the hazard rises from essentially zero below 1000 h to a maximum of $0.371 \times 10^{-3}\,\mathrm{h}^{-1}$ at about 8800 h, then declines slowly, to $0.255 \times 10^{-3}\,\mathrm{h}^{-1}$ at 30 000 h. With $\sigma = 1.0$ the density is broader and its hazard peaks earlier and lower, $0.168 \times 10^{-3}\,\mathrm{h}^{-1}$ at about 3100 h. The decline after the peak is not an improvement of the units: it reflects that, in a population with a wide spread of multiplicative factors, the units that survive far beyond the median are those with favorable factors.

> **Principle 6.3.** *The lognormal distribution represents multiplicative variability of the lifetime. Its hazard is nonmonotone: zero at the start, rising to a single maximum and then declining, so it is not a model of steadily increasing wear-out.*

6.4.4 Median, mean and variance

The three summary quantities of a lognormal lifetime are derived from the normal distribution of $X = \ln T$.

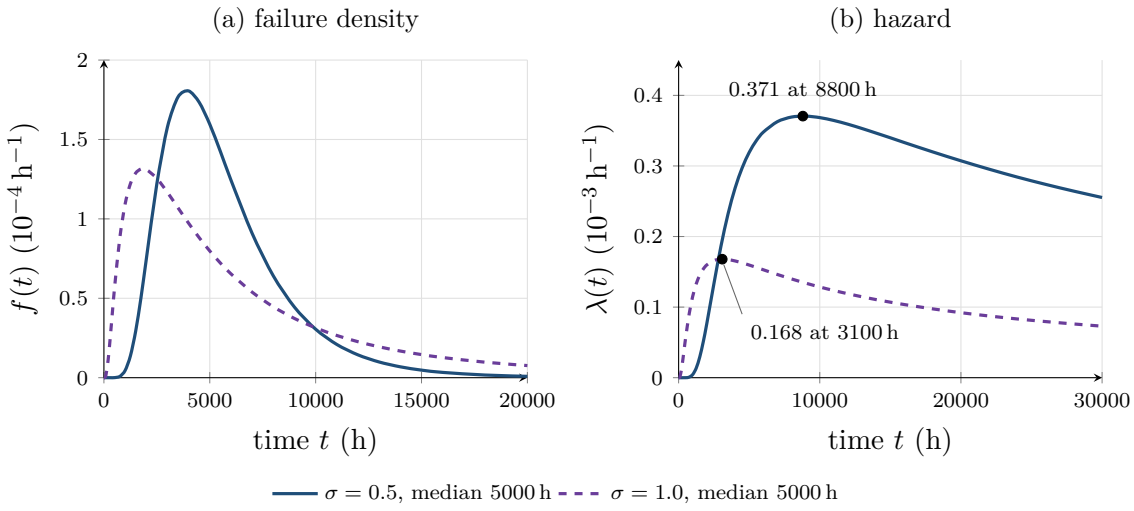


Figure 6.5: Lognormal lifetimes with median $\mathrm{e}^{\mu} = 5000\,\mathrm{h}$ ($\mu = \ln 5000 = 8.517$) and $\sigma = 0.5$ (solid) or 1.0 (dashed), illustrative. (a) Density, equation (6.16). (b) Hazard, equation (6.18), evaluated numerically: it is zero at $t = 0$, rises to a single maximum (dots) and then declines toward zero.

Median. The median of a normal variable is its mean, so $\Pr(\ln T \le \mu) = \Phi(0) = 0.5$ and, since the logarithm is increasing, the median life is

$$t_{50} = \mathrm{e}^{\mu}. \tag{6.19}$$

This is the most direct reading of μ: it is the logarithm of the median life.

Mean. The mean lifetime is $\mathbb{E}[T] = \mathbb{E}[\mathrm{e}^{X}]$ with $X \sim \mathcal{N}(\mu, \sigma^2)$:

$$\mathbb{E}[\mathrm{e}^{X}] = \int_{-\infty}^{\infty} \mathrm{e}^{x}\, \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \mathrm{d}x.$$

Collect the exponents and complete the square:

$$x - \frac{(x-\mu)^2}{2\sigma^2} = -\frac{(x-\mu)^2 - 2\sigma^2 x}{2\sigma^2} = -\frac{\big(x - (\mu+\sigma^2)\big)^2}{2\sigma^2} + \mu + \frac{\sigma^2}{2},$$

which is checked by expanding $(x-\mu-\sigma^2)^2 = (x-\mu)^2 - 2\sigma^2(x-\mu) + \sigma^4$. The factor $\mathrm{e}^{\mu+\sigma^2/2}$ comes out of the integral, and what remains is the integral of a normal density with mean $\mu+\sigma^2$ and variance σ^2, which is one. Therefore

$$\mathrm{MTTF} = \mathbb{E}[T] = \mathrm{e}^{\mu+\sigma^2/2}. \tag{6.20}$$

Variance. The same computation with $2X \sim \mathcal{N}(2\mu, 4\sigma^2)$ gives $\mathbb{E}[T^2] = \mathbb{E}[\mathrm{e}^{2X}] = \mathrm{e}^{2\mu+2\sigma^2}$, and

$$\mathrm{Var}(T) = \mathbb{E}[T^2] - \mathbb{E}[T]^2 = \mathrm{e}^{2\mu+2\sigma^2} - \mathrm{e}^{2\mu+\sigma^2} = \left(\mathrm{e}^{\sigma^2} - 1\right)\mathrm{e}^{2\mu+\sigma^2}. \tag{6.21}$$

The parameter σ controls the spread and the skewness of the lifetime. Holding μ fixed, a larger σ produces a broader distribution with a longer right tail, and it moves the mean away from the median by the factor $\mathrm{e}^{\sigma^2/2}$: the mean always exceeds the median, and the mode, $\mathrm{e}^{\mu-\sigma^2}$, always lies below it.

6.4.5 Worked example: electronic equipment

Suppose that the median lifetime of an electronic component is 5000 h and that $\sigma = 0.5$, illustrative values. From equation (6.19),

$$\mu = \ln 5000 = 8.517.$$

The mean lifetime, from equation (6.20), is

$$\mathbb{E}[T] = \mathrm{e}^{\mu}\,\mathrm{e}^{\sigma^2/2} = 5000\,\mathrm{e}^{0.5^2/2} = 5000\,\mathrm{e}^{0.125} = 5666\,\mathrm{h}.$$

The mean exceeds the median because the distribution is skewed to the right. Equation (6.21) gives a standard deviation of $\sqrt{(\mathrm{e}^{0.25}-1)\,\mathrm{e}^{2(8.517)+0.25}} = 3020\,\mathrm{h}$.

The probability that the component fails by 3000 h follows from equation (6.15):

$$F(3000) = \Phi\left(\frac{\ln 3000 - \ln 5000}{0.5}\right) = \Phi\left(\frac{-0.5108}{0.5}\right) = \Phi(-1.0217) = 0.153,$$

about 15.3 %, and therefore $R(3000) = 0.847$, or 84.7 %. These are the solid curves of Figure 6.5: the hazard at 3000 h is $0.186 \times 10^{-3}\,\mathrm{h}^{-1}$, still rising, and it will reach its maximum near 8800 h. The lognormal model predicts a different lifetime pattern from an exponential or a Weibull model even when their means or medians are made comparable, which the next section shows directly.

6.5 Comparing the three models

The three distributions can be made to agree on one summary number and still disagree on the probabilities that decisions depend on. This section compares them at equal mean lifetime and states the limits of the most widely used of them, the Weibull.

6.5.1 Equal mean lifetime, different reliability

Choose the parameters of the three models so that each has the mean lifetime $\mathbb{E}[T] = 5000\,\mathrm{h}$:

1. exponential with $\lambda = 1/5000\,\mathrm{h}^{-1}$;
2. Weibull with $\beta = 2$ and $\eta = 5000/\Gamma(1.5) = 5642\,\mathrm{h}$, from equation (6.12);
3. lognormal with $\sigma = 0.5$ and $\mu = \ln 5000 - 0.125 = 8.392$, from equation (6.20), so that the median is $\mathrm{e}^{\mu} = 4412\,\mathrm{h}$ and the mean is $\mathrm{e}^{8.392+0.125} = 5000\,\mathrm{h}$.

Figure 6.6 draws their reliability functions. The horizontal axis is operating time in hours; the vertical axis is the reliability. The areas under the three curves are equal, since each area is the MTTF, but the curves are not. Table 6.4 lists their values at nine ages. One entry of the arithmetic, for the lognormal at 1000 h: $z = (\ln 1000 - 8.392)/0.5 = (6.908 - 8.392)/0.5 = -2.969$ and $R(1000) = 1 - \Phi(-2.969) = 0.9985$.

The exponential model has by far the most early failures, 18.1 % by 1000 h against 3.1 % for the Weibull and 0.15 % for the lognormal, and by far the most long-lived survivors, 13.5 % at 10 000 h against 4.3 % and 5.1 %. The exponential and Weibull curves cross at 6366 h, the exponential and lognormal curves at about 5450 h, and the Weibull and lognormal curves twice, near 3460 and 9130 h. The practical implication follows directly: a component with a higher probability of surviving the first 1000 h may have a lower probability of surviving 10 000 h, although both have the same expected lifetime. The lognormal component survives 1000 h with probability 0.9985 against 0.8187 for the exponential one, and 10 000 h with probability 0.0509 against 0.1353. A procurement or maintenance decision based only on the MTTF can therefore be misleading; what matters is the reliability over the mission or the interval that the decision concerns.

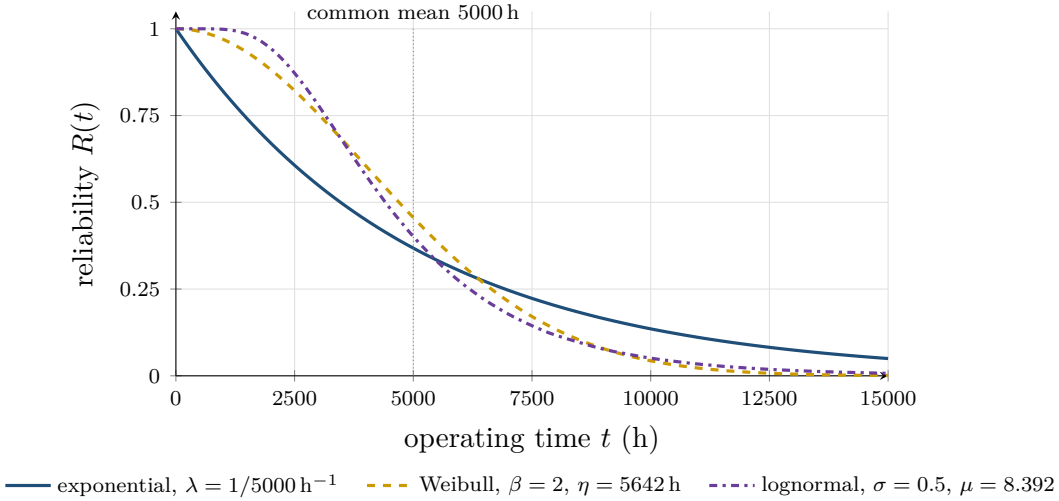


Figure 6.6: Reliability functions of three lifetime models with the same mean, $\mathbb{E}[T] = 5000\,\mathrm{h}$: exponential, $\mathrm{e}^{-t/5000}$; Weibull, $\exp[-(t/5642)^2]$; lognormal, $1 - \Phi[(\ln t - 8.392)/0.5]$, evaluated numerically. Illustrative parameters. The areas under the curves are equal; the survival probabilities are not.

Table 6.4: Reliability, in percent, of the three models of Figure 6.6, all with mean lifetime 5000 h.

Time t (h)	exponential	Weibull $\beta = 2$	lognormal $\sigma = 0.5$
1000	81.87	96.91	99.85
2000	67.03	88.19	94.32
3000	54.88	75.37	77.98
4000	44.93	60.49	57.78
5000	36.79	45.59	40.13
6000	30.12	32.27	26.94
8000	20.19	13.39	11.70
10 000	13.53	4.32	5.09
12 000	9.07	1.08	2.27

Common mistake

Comparing components by their MTTF alone. Equal means do not imply equal mission reliabilities: in Table 6.4 three models with MTTF = 5000 h give 1000 h reliabilities from 0.819 to 0.999. The mission reliability, $R(H)$ for the horizon H of interest, is the quantity to compare.

6.5.2 Limits of the Weibull model

The Weibull model is flexible, mathematically convenient and widely used, and it is natural to ask why it is not always the model of choice. Flexibility is not the same as correctness. The two-parameter Weibull hazard, equation (6.11), is a power of age, and a power is monotone: it can decrease, stay constant or increase, but it cannot rise and then fall. Consider an electronic component whose empirical hazard rises early, reaches a maximum and then declines. A standard two-parameter Weibull distribution cannot represent that nonmonotone pattern, while a lognormal distribution can, as Figure 6.5 shows.

Likewise, a population of components may experience several distinct failure mechanisms at once: early manufacturing defects, random external disturbances and late-life mechanical wear, the three terms of the bathtub curve of Chapter 5. A single Weibull distribution has one shape and so one monotone trend, and it may not represent the combined behavior adequately. The standard alternatives are *mixture models*, in which the population is a weighted combination of subpopulations each with its own distribution, as in the weak and strong units of Chapter 5; *competing-risk models*, in which each unit fails at the earliest of several mechanism-specific failure times, so that the hazards of the mechanisms add; *piecewise hazard models*, which use different hazard forms over different age ranges; and *physics-based degradation models*, which derive the lifetime from a damage process. The model should be selected on the evidence and the engineering question, not on familiarity.

6.6 Estimating and selecting a lifetime model

So far the parameters of each model were assumed known. In practice they are estimated from failure data, and the model itself must be chosen. This section derives the maximum likelihood estimate of the exponential rate, describes Weibull and lognormal estimation, and sets out a procedure for selecting a model that distinguishes statistical fit from the credibility of a prediction.

6.6.1 Maximum likelihood estimation

Suppose that five components are run to failure and fail at the operating ages of Table 6.5, illustrative values. The table gives, for each component, its failure time t_i, its logarithm, and a plotting position used in Section 6.6.4.

A common method of estimation is maximum likelihood. Let $t_1, t_2, \ldots, t_n$ be independent failure times, all observed completely, from a distribution with density $f(t \mid \boldsymbol{\theta})$ and parameter vector $\boldsymbol{\theta}$. The likelihood is the joint density of the observations, regarded as a function of the parameters,

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} f(t_i \mid \boldsymbol{\theta}), \tag{6.22}$$

where the product expresses the independence of the observations. The maximum likelihood estimate $\hat{\boldsymbol{\theta}}$ is the parameter value that maximizes L, the value under which the observed data are most probable. Because the logarithm is increasing, it is equivalent and simpler to maximize the log-likelihood $\ell(\boldsymbol{\theta}) = \ln L(\boldsymbol{\theta}) = \sum_i \ln f(t_i \mid \boldsymbol{\theta})$.

Table 6.5: Five complete failure times, illustrative. $\hat{F}_i = (i - 0.3)/(n + 0.4)$ is the median-rank plotting position of the i-th smallest time, and $\ln \hat{\Lambda}_i = \ln[-\ln(1 - \hat{F}_i)]$.

Component	failure time t_i (h)	$\ln t_i$	$\hat{F}_i$	$\ln \hat{\Lambda}_i$
A	1200	7.0901	0.1296	−1.9745
B	1800	7.4955	0.3148	−0.9727
C	2400	7.7832	0.5000	−0.3665
D	4000	8.2940	0.6852	0.1448
E	5600	8.6305	0.8704	0.7145

The exponential estimate. For the exponential model, $f(t_i) = \lambda \, \mathrm{e}^{-\lambda t_i}$, so

$$L(\lambda) = \prod_{i=1}^{n} \lambda \, \mathrm{e}^{-\lambda t_i} = \lambda^n \, \mathrm{e}^{-\lambda \sum_i t_i}, \qquad \ell(\lambda) = n \ln \lambda - \lambda \sum_i t_i.$$

Differentiating with respect to λ,

$$\frac{\mathrm{d}\ell}{\mathrm{d}\lambda} = \frac{n}{\lambda} - \sum_i t_i,$$

and setting the derivative to zero gives $n/\lambda - \sum_i t_i = 0$, so

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} t_i}. \tag{6.23}$$

The second derivative, $-n/\lambda^2$, is negative, so the stationary point is a maximum. The estimate is the number of failures divided by the total operating time accumulated by all units. In the notation of Chapter 7, with r failures and a total time on test T_{tot}, it reads $\hat{\lambda} = r/T_{\mathrm{tot}}$; here every unit failed, so $r = n$ and $T_{\mathrm{tot}} = \sum_i t_i$. For the data of Table 6.5,

$$\sum_i t_i = 1200 + 1800 + 2400 + 4000 + 5600 = 15\,000 \,\mathrm{h}, \qquad \hat{\lambda} = \frac{5}{15\,000 \,\mathrm{h}} = 3.333 \times 10^{-4} \,\mathrm{h}^{-1},$$

and the estimated mean time to failure is $\widehat{\mathrm{MTTF}} = 1/\hat{\lambda} = 3000 \,\mathrm{h}$, the sample mean.

This estimate rests on five failures and is not a physical constant. The size of its sampling uncertainty can be judged from the curvature of the log-likelihood: the Fisher information is $\mathcal{I}(\lambda) = -\mathrm{d}^2\ell/\mathrm{d}\lambda^2 = n/\lambda^2$, and the standard large-sample approximation to the standard error is $1/\sqrt{\mathcal{I}(\hat{\lambda})} = \hat{\lambda}/\sqrt{n}$, which is 45 % of the estimate for $n = 5$. Interval estimates are treated in Chapter 7.

6.6.2 Weibull and lognormal estimation

For the Weibull model the likelihood is the product of densities (6.10),

$$L(\beta, \eta) = \prod_{i=1}^{n} \frac{\beta}{\eta} \left(\frac{t_i}{\eta} \right)^{\beta - 1} \exp\left[- \left(\frac{t_i}{\eta} \right)^{\beta} \right],$$

and taking logarithms term by term, $\ln[(\beta/\eta)(t_i/\eta)^{\beta-1}] = \ln \beta - \beta \ln \eta + (\beta - 1) \ln t_i$, gives the log-likelihood

$$\ell(\beta, \eta) = n \ln \beta - n\beta \ln \eta + (\beta - 1) \sum_i \ln t_i - \sum_i \left(\frac{t_i}{\eta} \right)^{\beta}. \tag{6.24}$$

Setting the derivative with respect to η to zero, $-n\beta/\eta + (\beta/\eta)\sum_i (t_i/\eta)^\beta = 0$, gives the scale in terms of the shape,

$$\hat{\eta}^\beta = \frac{1}{n}\sum_i t_i^\beta .$$

Setting the derivative with respect to β to zero and substituting this relation leaves one equation for the shape,

$$\frac{\sum_i t_i^\beta \ln t_i}{\sum_i t_i^\beta} - \frac{1}{\beta} - \frac{1}{n}\sum_i \ln t_i = 0,$$

which has no closed-form solution. The parameters are therefore found by numerical optimization of equation (6.24) or by solving the likelihood equations numerically. For the five failure times the solution is $\hat{\beta} = 2.019$ and $\hat{\eta} = 3407\,\mathrm{h}$, which gives a mean of $3407\,\Gamma(1 + 1/2.019) = 3019\,\mathrm{h}$.

A similar approach applies to the lognormal model, and there it has a closed form: since $\ln t_i$ are normal observations, the estimates are the sample mean and the standard deviation of the logarithms, with divisor n,

$$\hat{\mu} = \frac{1}{n}\sum_i \ln t_i = 7.859, \qquad \hat{\sigma} = \sqrt{\frac{1}{n}\sum_i (\ln t_i - \hat{\mu})^2} = 0.550,$$

a median of $\mathrm{e}^{7.859} = 2588\,\mathrm{h}$ and a mean of $\mathrm{e}^{7.859 + 0.550^2/2} = 3011\,\mathrm{h}$.

Table 6.6 collects the three fits. Its columns are the estimated parameters, the fitted mean, the maximized log-likelihood $\hat{\ell}$, the Akaike information criterion $\mathrm{AIC} = 2k - 2\hat{\ell}$, where k is the number of estimated parameters, and the fitted reliability at three ages. All three models give a mean of about 3000 h. The log-likelihoods differ by less than 1.7, and the AIC values, 92.06, 91.11 and 90.80, by less than 1.3, a difference too small to prefer any model. The reliabilities at 1000 h, inside the range of the data, differ by 0.24. Those at 10 000 and 20 000 h, far beyond the largest observed failure, differ by orders of magnitude, a point taken up in Section 6.6.6.

Table 6.6: Maximum likelihood fits of the three models to the five failure times of Table 6.5. $\mathrm{AIC} = 2k - 2\hat{\ell}$; lower is better.

Model	estimates	mean (h)	$\hat{\ell}$	AIC	$R(1000)$	$R(10\,000)$	$R(20\,000)$
exponential	$\hat{\lambda} = 3.333 \times 10^{-4}\,\mathrm{h}^{-1}$	3000	-45.03	92.06	0.7165	3.6×10^{-2}	1.3×10^{-3}
Weibull	$\hat{\beta} = 2.019$, $\hat{\eta} = 3407\,\mathrm{h}$	3019	-43.56	91.11	0.9193	1.5×10^{-4}	3×10^{-16}
lognormal	$\hat{\mu} = 7.859$, $\hat{\sigma} = 0.550$	3011	-43.40	90.80	0.9580	7.0×10^{-3}	1.0×10^{-4}

All of this assumes that every unit was observed until it failed. A complication arises when some components have not failed when the study ends. Such observations are *right-censored*: a unit still running at 5000 h shows that its lifetime exceeds 5000 h, without showing by how much. Censored observations still contain information about reliability, and ignoring them, or recording them as failures at the end of observation, can bias the lifetime estimates severely, usually toward a shorter life. The likelihood of censored data, the Kaplan–Meier estimator and interval estimates are the subject of Chapter 7.

6.6.3 Four perspectives on model selection

The choice among exponential, Weibull and lognormal distributions is best made by examining the problem from four perspectives. Table 6.7 lists them: each row gives the criterion, the

question it asks, and why the answer matters. The physical mechanism proposes plausible hazard behavior; the data constrain the parameters; the statistical fit detects systematic departures; and the engineering objective decides which part of the distribution must be right.

Table 6.7: Four perspectives from which a lifetime model is selected.

Criterion	Question	Why it matters
Physical mechanism	Is failure associated with aging, random events or accumulated damage?	Establishes plausible hazard behavior
Observed data	What do the failure times and the censored observations indicate?	Constrains the parameter estimates
Statistical fit	Which model represents the observations adequately?	Detects systematic departures from the model
Engineering objective	Is the prediction about early failures, long-term survival or remaining life?	Determines which part of the distribution matters most

The selection then proceeds in four steps, of which the first is taken here and the other three in the next two subsections.

Step 1: establish plausible failure behavior. The mechanism and any prior experience suggest a hazard shape, and the shape suggests a candidate model. Table 6.8 gives a preliminary guide: each row is an observed or expected behavior and the model that represents it. The rows are starting points, not automatic selection rules; the remaining steps test them.

Table 6.8: Preliminary guide from hazard behavior to candidate model.

Observed or expected behavior	Candidate model
Approximately constant hazard	exponential
Decreasing hazard	Weibull with $\beta < 1$
Increasing hazard	Weibull with $\beta > 1$
Hazard that rises and eventually declines	lognormal
Several distinct failure mechanisms	mixture or competing-risk model
Strong dependence on operating conditions	covariate-dependent or physics-informed model

6.6.4 Probability plotting

Step 2: examine the data graphically. A probability plot shows whether a proposed distribution is consistent with the data by transforming the axes so that the distribution becomes a straight line. For the Weibull model the transformation comes from the cumulative hazard, equation (6.8). Taking logarithms of $\Lambda(t) = (t/\eta)^\beta$,

$$\ln \Lambda(t) = \beta \ln t - \beta \ln \eta. \tag{6.25}$$

A plot of $\ln \Lambda(t)$ against $\ln t$ therefore follows a straight line if the Weibull model is appropriate, and its slope is the shape parameter β; the line reaches $\ln \Lambda = 0$ at $t = \eta$. This is the basis of Weibull probability plotting. The cumulative hazard at each failure time is estimated from the data through $\Lambda = -\ln(1 - F)$, with F estimated at the i-th smallest of n failure times by a plotting position. The median rank, approximately $\hat{F}_i = (i - 0.3)/(n + 0.4)$, is the usual choice; Table 6.5 lists the resulting coordinates. The exponential model is the special case of slope one.

Figure 6.7 is the Weibull plot of the five failure times. The horizontal axis is the failure time on a logarithmic scale; the vertical axis is $\ln \hat{\Lambda}$. The dots are the data. The solid line is equation (6.25) with the maximum likelihood estimates $\hat{\beta} = 2.019$ and $\hat{\eta} = 3407\,\text{h}$; the dashed line is the exponential fit, $\ln(\hat{\lambda}t)$, whose slope is one. The points lie close to the Weibull line and visibly steeper than the exponential one, which suggests an increasing hazard. With five points, however, the slope is poorly determined: a least-squares line through the same five points has slope 1.66, not 2.02.

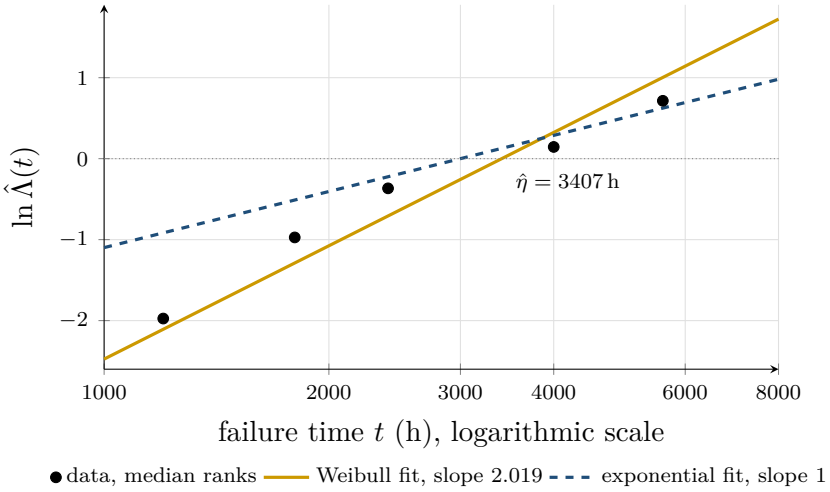


Figure 6.7: Weibull probability plot of the five failure times of Table 6.5, illustrative. On these axes a Weibull distribution is the straight line $\ln \Lambda = \beta \ln t - \beta \ln \eta$ (equation (6.25)); solid: maximum likelihood fit, $\hat{\beta} = 2.019$, $\hat{\eta} = 3407\,\text{h}$, which crosses zero at $t = \hat{\eta}$; dashed: exponential fit, $\hat{\lambda} = 1/3000\,\text{h}^{-1}$.

The same transformation gives a quick two-point fit. The illustrative pump population of Chapter 5 had $F(1000) = 0.10$ and $F(2000) = 0.25$, so $\Lambda(1000) = -\ln 0.90 = 0.1054$ and $\Lambda(2000) = -\ln 0.75 = 0.2877$. The slope of the line through the two points is $\beta = \ln(0.2877/0.1054)/\ln(2000/1000) = 1.0045/0.6931 = 1.449$, and $\eta = 1000 \times 0.1054^{-1/1.449} = 4725\,\text{h}$: the smooth density used for that population was a Weibull distribution with these two parameters, an increasing hazard. Probability plots and cumulative hazard plots are practical diagnostic tools, but small data sets often cannot distinguish clearly between competing models; the lognormal and Weibull fits of Table 6.6 would both pass through the five points of Figure 6.7 acceptably.

6.6.5 Comparing fitted models and their uncertainty

Step 3: compare the fitted models. After the candidate distributions are fitted, they are examined together, on:

- the maximized log-likelihood and information criteria such as the AIC, which rewards fit and penalizes parameters;
- goodness-of-fit diagnostics;
- residual and probability plots;
- the stability of the fitted parameters when data are added or removed;
- the sensitivity of the predicted reliability to the choice of distribution;
- the credibility of any extrapolation beyond the observed failure ages.

A model with a slightly better AIC is not necessarily the better engineering model if it conflicts with a well-established physical mechanism or produces implausible extrapolations. In Table 6.6 the lognormal has the lowest AIC, but by 0.31 over the Weibull; if the mechanism is known to

be wear-out, the Weibull, whose hazard keeps increasing, is the more defensible choice, and the data give no reason to reject it.

Step 4: evaluate the uncertainty of the prediction. Suppose that a Weibull model is fitted and the shape is estimated as $\hat{\beta} = 2.5$. That point estimate does not say how precisely β is known. With limited failure data several values of β may be statistically plausible, and the predicted reliability may then be uncertain even though the fitted curve looks smooth. As an illustration, hold $\eta = 5000\,\text{h}$ and let the shape range over 1.5, 2.0, 2.5, 3.0 and 3.5: the probability of failure within the first 1000 h, $1 - \exp[-(1000/5000)^{\beta}]$, is then 0.0856, 0.0392, 0.0177, 0.0080 and 0.0036, a factor of 24 between the ends of the range, around the value 0.0177 at the point estimate. In a real fit β and η are estimated together and their errors are correlated, but the conclusion holds: the farther into the tail the prediction, here early failure under an increasing hazard, the more it depends on the shape. This matters most when the quantity of interest is a rare failure probability. A sound reliability study reports both the predicted probability and an assessment of its uncertainty.

6.6.6 Statistical fit and predictive credibility

Consider a data set of 20 observed failures, all between 1000 and 5000 h, from which the reliability at 20 000 h is to be predicted. An exponential, a Weibull and a lognormal model might all fit the observed data reasonably well, and their extrapolated survival probabilities could still differ greatly. The five failure times of Table 6.6 show the effect in miniature: three fits that the AIC cannot separate predict $R(20\,000)$ as 1.3×10^{-3}, 1.0×10^{-4} and 3×10^{-16}, a spread of thirteen orders of magnitude. Each number follows correctly from its model; they differ because each model continues its own hazard into an age range where no failure was observed: constant for the exponential, rising for the Weibull, falling for the lognormal.

The difficulty is therefore not computational. The model is being asked to predict behavior in a range of lifetimes where there are no observations, and agreement within the observed range does not guarantee a credible long-term extrapolation. For high-consequence systems an engineer should examine the physical failure mechanisms, the alternative models, the uncertainty of the parameters and the intended operating environment before relying on such a prediction. This is where statistical reliability modeling and physics-of-failure modeling become complementary: the mechanism decides how the hazard should continue beyond the data, and the data calibrate the mechanism within it.

> **Principle 6.4.** *Model selection must consider the physics, the data and the intended prediction. A good statistical fit within the observed range of lifetimes does not by itself make a reliability prediction credible, least of all an extrapolation beyond the observed failure ages.*

The estimates in this chapter assumed complete data, every unit observed to failure. The next chapter turns to the realistic case, in which many units have not failed when the data are collected, and to the estimation methods that handle it.

6.7 What to retain

Definition 6.1 (Lifetime models)**.** The exponential, Weibull and lognormal distributions differ in the shape of their hazard: constant; decreasing, constant or increasing according to the shape β; and rising to a single maximum, then declining. A lifetime model is chosen from the failure mechanism, the data and the prediction it must support, since models that agree on the mean, or fit the same data equally well, can give very different mission reliabilities and extrapolations.

Table 6.9 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 6.9: The terms of the three lifetime models and of their estimation, and where each is defined. Units assume time in hours.

Term	Expression	Units	Meaning	Defined in
Exponential model	$R(t) = \mathrm{e}^{-\lambda t}$, $\lambda = \lambda$	–	constant hazard	6.2.1
Failure rate	$\lambda = 1/\mathrm{MTTF}$	h^{-1}	exponential parameter	6.2.2
Median life	t_{50}, $R(t_{50}) = 0.5$	h	half the population failed	6.2.2
Memoryless property	$\Pr(T > s + t \mid T > s) = R(t)$	–	exponential only	6.2.4
Weibull model	$R(t) = \exp[-(t/\eta)^{\beta}]$	–	power-law $\Lambda(t)$	6.3.2
Weibull hazard	$(\beta/\eta)(t/\eta)^{\beta-1}$	h^{-1}	monotone in age	6.3.2
Shape parameter	β	–	< 1 DFR, $= 1$ constant, > 1 IFR	6.3.3
Characteristic life	η, $F(\eta) = 0.632$	h	63.2 % life	6.3.4
Weibull mean	$\eta\,\Gamma(1 + 1/\beta)$	h	MTTF	6.3.4
Lognormal model	$\ln T \sim \mathcal{N}(\mu, \sigma^2)$	–	multiplicative variability	6.4.2
Lognormal reliability	$1 - \Phi[(\ln t - \mu)/\sigma]$	–	survival beyond t	6.4.2
Lognormal hazard	$f(t)/R(t)$	h^{-1}	rises, peaks, declines	6.4.3
Lognormal median, mean	e^{μ}, $\mathrm{e}^{\mu+\sigma^2/2}$	h	mean exceeds median	6.4.4
Likelihood	$L(\boldsymbol{\theta}) = \prod_i f(t_i \mid \boldsymbol{\theta})$	–	data as a function of parameters	6.6.1
Exponential MLE	$\hat{\lambda} = n/\sum_i t_i$	h^{-1}	failures per unit time	6.6.1
AIC	$2k - 2\hat{\ell}$	–	fit penalized by parameters	6.6.2
Right-censoring	$T > t_{\mathrm{end}}$	–	lower bound only	6.6.2
Weibull plot	$\ln \Lambda = \beta \ln t - \beta \ln \eta$	–	slope β	6.6.4
Predictive credibility	–	–	extrapolation beyond data	6.6.6

Further reading. Rausand, Barros and Høyland, *System Reliability Theory*, third edition: lifetime distributions and component reliability. Modarres, Kaminskiy and Krivtsov, *Reliability Engineering and Risk Analysis*, third edition: statistical lifetime models. O'Connor and Kleyner, *Practical Reliability Engineering*, fifth edition: Weibull analysis and practical reliability data. The *NIST/SEMATECH e-Handbook of Statistical Methods*: reliability distributions and models, and reliability data analysis, covering model estimation, censoring and diagnostics.

6.8 Exercises

Exercise 6.1. (a) A Weibull distribution has shape $\beta = 3$. State what this implies about its hazard. (The hazard increases with age; for $\beta > 1$ the Weibull hazard is increasing, here with the square of age.) (b) State what percentage of a Weibull population has failed by its characteristic life η. (63.2 %, since $F(\eta) = 1 - \mathrm{e}^{-1} = 0.6321$ for every β.) (c) Two components have identical mean lifetimes. State whether their mission reliabilities must be identical. (No; distributions with the same mean can have different survival and hazard functions, as Table 6.4 shows.) (d) For a lognormal lifetime, state which quantity is normally distributed. (The logarithm of the lifetime, $\ln T$, not T and not the hazard.)

Exercise 6.2. A pump bearing has a Weibull lifetime with $\beta = 1.5$ and $\eta = 8000\,\text{h}$. (a) Compute $R(2000\,\text{h})$. ($\exp[-(0.25)^{1.5}] = \text{e}^{-0.125} = 0.8825$.) (b) Compute the MTTF from equation (6.12). ($8000\,\Gamma(1.667) = 8000 \times 0.9027 = 7222\,\text{h}$.) (c) Compute the hazard at 1000 and 4000 h and their ratio. (6.63×10^{-5} and $1.33 \times 10^{-4}\,\text{h}^{-1}$; the ratio is $4^{0.5} = 2$.) (d) Compute the probability of failure in the next 1000 h for a new bearing and for one that has survived 4000 h. (0.0432; $1 - \text{e}^{-(0.4941-0.3536)} = 0.1311$.) (e) For an exponential model with the same MTTF, compute $R(2000\,\text{h})$ and the probability of failure in any 1000 h interval, and comment. ($\text{e}^{-2000/7222} = 0.7581$ and $1 - \text{e}^{-1000/7222} = 0.1293$ at every age; the exponential model overstates the early risk and understates the risk of the aged bearing.)

Exercise 6.3. An electronic module has a lognormal lifetime with a median of 20 000 h and $\sigma = 0.8$. (a) Compute μ, and state why it is not a repair rate. ($\mu = \ln 20\,000 = 9.904$; it is the mean of $\ln T$, in log-hours, unrelated to the repair rate μ.) (b) Compute the mean lifetime. ($20\,000\,\text{e}^{0.32} = 27\,540\,\text{h}$.) (c) Compute $R(10\,000\,\text{h})$ and $F(5000\,\text{h})$. ($z = \ln 0.5/0.8 = -0.866$, $R = 1 - \Phi(-0.866) = 0.807$; $z = \ln 0.25/0.8 = -1.733$, $F = \Phi(-1.733) = 0.042$.) (d) State whether the hazard of this module keeps increasing at very long ages, and what this implies for a replacement policy based on the model alone. (No; the lognormal hazard rises to a single maximum, here near 20 000 h at about $5.0 \times 10^{-5}\,\text{h}^{-1}$, and then declines, to $3.0 \times 10^{-5}\,\text{h}^{-1}$ at 100 000 h, so the model alone gives no age beyond which an older module is worse than a younger one; an age-based replacement must rest on the mechanism.)

Exercise 6.4. Six units are run to failure at 800, 1500, 2100, 2900, 3600 and 5100 h. (a) Compute the exponential maximum likelihood estimate of λ, the estimated MTTF, and the approximate relative standard error $1/\sqrt{n}$. ($6/16\,000 = 3.75 \times 10^{-4}\,\text{h}^{-1}$; 2667 h; 41 %.) (b) Holding the Weibull shape at $\beta = 2$, derive the estimate of η from the likelihood equation of Section 6.6.2 and compute it, with the corresponding mean. ($\hat{\eta} = (\sum_i t_i^2/n)^{1/2} = (54\,680\,000/6)^{1/2} = 3019\,\text{h}$; mean $0.8862 \times 3019 = 2675\,\text{h}$. The full two-parameter fit gives $\hat{\beta} = 2.001$, so the fixed shape is consistent with the data.) (c) Compute $R(1000\,\text{h})$ under both fits. (0.687 and 0.896.) (d) A seventh unit is still running at 6000 h when the test ends. State what this observation contributes and how it must be treated. (It shows that its lifetime exceeds 6000 h, a lower bound; it must be handled as a right-censored observation, not ignored and not recorded as a failure at 6000 h.)

Chapter 7

Reliability Data and Statistical Estimation

7.1 Reliability from incomplete operating records

Chapter 5 defined the functions that describe a random lifetime, and Chapter 6 studied the exponential, Weibull and lognormal models of it. In both chapters the parameters of the model were taken as known: the constant failure rate λ, the Weibull shape β and scale η, and the lognormal parameters μ and σ of $\ln T$ (here μ is a lognormal parameter, not the repair rate μ). In engineering practice these quantities are rarely known in advance. What is available is a set of observations: failure records, operating-hour counters, inspection reports, and equipment that is still in service. The parameters must be inferred from that evidence, and the evidence is almost always incomplete, because many of the units observed have not failed yet. The question of this chapter is

how the reliability of a component is estimated from operating records in which some units have failed, others are still operating, and the lifetime of most units has been observed only in part.

The chapter answers it in five steps. It describes reliability data and the right-censored observation, the record of a unit that was still operating when its observation ended, and shows why such records cannot be discarded. It derives the Kaplan–Meier estimator, which estimates the reliability function without assuming a lifetime distribution. It constructs the likelihood of a set of failure and censored records, and from it derives the maximum likelihood estimate of the exponential failure rate. It quantifies the statistical uncertainty of these estimates by standard errors and confidence limits, for the exponential model, for the Kaplan–Meier estimate and for the two-parameter Weibull model. Finally, it states the diagnostic checks that decide whether a fitted model deserves engineering credence.

Jargon: reliability data, censoring, risk set, Kaplan–Meier estimator, likelihood, MLE, total time on test, standard error, confidence interval, Fisher information

Reliability data (life data) are records of the operating times of units of a population and of whether each record ended in a failure. A *right-censored* observation is a record that ends before failure, so that only a lower bound of the lifetime is known; the end of its observation is the *censoring time*. In *interval censoring* the failure is known to lie between two inspections, and in *left censoring* it had already occurred at the first inspection. *Left truncation* (delayed entry) means that a unit enters the sample only after surviving to some age. The *risk set* at a time t is the set of units still under observation and not yet failed just before t. The *Kaplan–Meier estimator*, also called the *product-limit estimator*, estimates the reliability function as a product of conditional survival fractions over the risk sets; it is *nonparametric*, meaning that it assumes no family of lifetime distributions. The *likelihood* $L(\theta)$ is the probability or density of the observed data, viewed as a function of the model parameters θ; the *maximum likelihood estimate* (MLE) is the parameter value that maximizes it. The *total time on test*, or total exposure, T_{tot}, is the sum of the operating times of all units observed, failed or not. The *standard error* is the standard deviation of an estimator, estimated from the data. A $100(1-\alpha)\,\%$ *confidence interval* is an interval computed from the data by a procedure that contains the true parameter in a fraction $1-\alpha$ of repeated samples. The *Fisher information* $\mathcal{I}(\theta)$ is the expected curvature of the log-likelihood, whose inverse is the large-sample variance of the MLE.

Key idea

A unit still operating when its observation ends is evidence: it shows that its lifetime exceeds the observed time. A correct estimator uses that evidence, through the risk sets of the Kaplan–Meier estimator and through the survival factor $R(c)$ in the likelihood, and neither discards the record nor counts it as a failure.

This section states why lifetime parameters must be estimated, presents the dataset used throughout the chapter, and lists the questions and assumptions of the analysis.

7.1.1 Parameters estimated from evidence

Three objects must be kept apart in what follows. The physical system is the population of components and the conditions under which they operate; its lifetimes have some distribution, which nobody observes directly. The model is a mathematical description of that distribution: a parametric family such as the exponential or the Weibull, or, without a family, a reliability function estimated point by point. The evidence is the set of records actually collected, with whatever gaps the observation process left in them. Estimation is the step from evidence to model: it chooses the model's parameters, or its estimated reliability function, so that the model is consistent with the records, and it states how far the records constrain that choice.

The records constrain the choice only through the way they were collected. A unit removed from service while still working, a counter read at an inspection, a test stopped at a planned time: each of these observation mechanisms determines what a record says about the lifetime, and an estimator that ignores the mechanism draws the wrong conclusion. This is why the chapter begins with the data before any formula.

7.1.2 The eight-pump dataset

As a running example, an industrial facility installs eight nominally identical cooling-water pumps. Each begins service in a known operating condition, and its cumulative operating hours are recorded from installation until its first failure or until the observation of that pump ends.

Four pumps fail; four are still operating when their observation ends. Table 7.1 lists the records: for each pump, the observed operating hours and whether the record ended in a failure or with the pump still operating. The data are synthetic and illustrative; they are used for every calculation in the chapter.

Table 7.1: The eight-pump dataset, illustrative. Observed operating hours from installation to the first failure or to the end of observation.

Pump	Observed operating hours	Status	Failure indicator δ_i
P1	400	failed	1
P2	700	failed	1
P3	900	still operating	0
P4	1200	failed	1
P5	1500	still operating	0
P6	1600	failed	1
P7	1800	still operating	0
P8	2000	still operating	0

Figure 7.1 draws the same records as observed histories. The horizontal axis is operating hours since installation; each row is one pump, and its line runs from installation to the end of its observed history. A red cross at the end of a line marks an observed failure; a green dot marks the end of observation with the pump still operating. The dots are not failures. Each says only that the pump survived at least as long as its line, and the four longest histories, which carry most of the operating time, all end in dots.

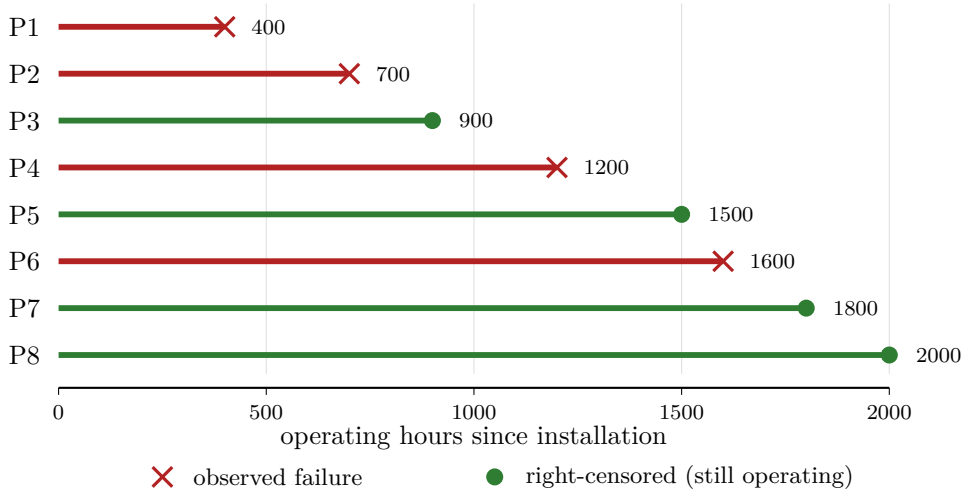


Figure 7.1: Observed histories of the eight pumps of Table 7.1, illustrative. Each line runs from installation to the end of observation; a cross is a failure, a dot is the end of observation with the pump still operating, which establishes only that its lifetime exceeds the observed time.

7.1.3 Engineering questions and assumptions

The engineering questions asked of the dataset are five. What is the probability that a new pump survives 1000 h? What is the probability that it survives 2000 h? What failure rate is consistent with the observations? How uncertain are these estimates? And do the observations support a constant-hazard model? All five are answered in this chapter, and the last two receive

the most attention, because a number without its uncertainty and without a check of its model is not yet an engineering result.

Two assumptions underlie every estimator below. First, the first-failure lifetimes $T_1, \ldots, T_8$ of the pumps are independent and drawn from the same population distribution: the pumps are exchangeable members of one population under comparable conditions. Second, the process that ends each observation is independent of the pump's failure time: a pump's record does not end because the pump was about to fail. The second assumption is examined in Section 7.3.5 and the first in Section 7.6.3. Under them, the central difficulty is that the four surviving pumps still carry reliability information, and together they account for 6200 of the 10 100 pump-hours observed.

7.2 Censored lifetime data

A record that ends before failure is the normal case in reliability data, not an exception. This section defines the right-censored observation, shows what goes wrong when such records are dropped or miscounted, and classifies the forms of incomplete lifetime data.

7.2.1 Right-censoring

Pump P8 has operated for 2000 h without failure when its observation ends. What is known about its lifetime T_8 is the inequality

$$T_8 > 2000\,\text{h},$$

and nothing more: it may fail at 2100 h or survive beyond 20 000 h. Such a record is a *right-censored observation*: the lifetime is known to exceed a particular value, the censoring time, but its exact value is unknown. Right-censoring is a standard feature of life testing and of field data, and the methods of this chapter incorporate such records without treating them as failures.

The record of unit i is written as a pair. Let T_i be its lifetime and C_i its censoring time, the time at which its observation would end if it had not failed (in this chapter C_i is a censoring time, not a capacity). What is observed is the shorter of the two and an indicator of which one it was,

$$y_i = \min(T_i, C_i), \qquad \delta_i = \begin{cases} 1, & \text{failure observed, } y_i = T_i, \\ 0, & \text{right-censored, } T_i > y_i = C_i. \end{cases} \tag{7.1}$$

The last column of Table 7.1 is δ_i, and the operating hours are y_i.

Three censoring schemes are distinguished, because the confidence limits of Section 7.5.3 depend on them. In a *time-terminated* (Type I) test all units are observed to a common time fixed in advance, and the number of failures is random. In a *failure-terminated* (Type II) test the observation stops at the r-th failure, and the total time on test is random. In *random* or *progressive* censoring each unit has its own censoring time, set by when it entered service, when it was withdrawn or when the data were extracted. Field data, and the eight pumps, are of the third kind.

7.2.2 Consequences of ignoring or miscounting censored records

The simplest summary of the dataset averages the four observed failure times,

$$\frac{400 + 700 + 1200 + 1600}{4} = \frac{3900}{4} = 975\,\text{h}.$$

This is not a valid estimate of the mean lifetime of the population. It ignores four pumps that survived substantial periods without failure, and it is biased toward short lives by construction:

a lifetime appears among the observed failures only if it was shorter than its censoring time, so the averaged values are a selection of the short lifetimes.

The opposite treatment, counting the censoring times as failures, is also wrong. It replaces each inequality $T_i > y_i$ by the equality $T_i = y_i$, which shortens every censored lifetime to its lower bound. The average of all eight observed times, $10\,100/8 = 1262.5$ h, is again too short. The two naive treatments differ in how they handle the censored records, but both bias the estimate toward short lives; the maximum likelihood estimate of Section 7.4.4, which uses the censored records as lower bounds, gives 2525 h under the exponential model. What is needed is an estimator that uses both kinds of record for what each says.

Common mistake

Averaging only the observed failure times, or treating censoring times as failure times. The first discards the survival evidence and averages a selection of short lifetimes (975 h for the pumps); the second replaces lower bounds by exact values (1262.5 h). Both understate the lifetime; a censored record must enter as the inequality $T_i > y_i$.

Principle 7.1. *Surviving components are data. A unit still operating at the end of its observation contributes evidence about the lifetime distribution even though its failure time is unknown, and censoring must be handled explicitly: ignoring censored records, or treating them as failures, leads to invalid inference.*

7.2.3 Forms of incomplete lifetime data

Right-censoring is one of several observation mechanisms that leave a lifetime partly unknown. Table 7.2 lists them: each row gives the type of record, what it establishes about the lifetime T, an example for a pump, and the factor the record contributes to the likelihood, derived in Section 7.4.2. An exact failure fixes T. A right-censored record bounds it from below. An interval-censored record, typical of equipment checked only at inspections, places the failure between two inspection times a and b. A left-censored record establishes only that the failure had already occurred at the first inspection, at time c.

Table 7.2: Forms of lifetime records, with what each establishes and its factor in the likelihood (Section 7.4.2). Pump examples illustrative.

Record	What is known	Example	Likelihood factor
Exact failure	$T = t$	pump failed at 700 h	$f(t)$
Right-censored	$T > c$	pump survived at least 1800 h	$R(c)$
Interval-censored	$a < T \le b$	inspection found the failure between 1000 and 1200 h	$F(b) - F(a)$
Left-censored	$T \le c$	failure had already occurred at the first inspection	$F(c)$

A different issue is *left truncation*, or delayed entry: a unit enters the observed sample only after it has survived to some age a, for example a pump already in service, with a hours on its counter, when record keeping began. A left-truncated unit is not a censored record. Units that failed before age a never entered the sample at all, so the sample over-represents long lives unless each such record is conditioned on survival to its entry age, which divides its likelihood factor by $R(a)$. Left-censoring, by contrast, concerns a unit that is in the sample and whose failure time is bounded from above. For the rest of the chapter the records are exact failures and right-censored observations, as in the eight-pump dataset.

7.3 The Kaplan–Meier estimator

The first estimator of the chapter assumes no lifetime distribution at all. This section states its purpose, derives it from a product of conditional survival probabilities, applies it to the eight pumps, draws the resulting step function, and states the assumption on which it rests and the limit it places on estimating the mean lifetime.

7.3.1 Nonparametric estimation of the reliability function

The reliability function can be estimated without first assuming that lifetimes are exponential, Weibull or lognormal. The Kaplan–Meier estimator does this: it is a nonparametric estimator, which answers the question of what the estimated survival probability is, given the observed failures and the number of units still at risk at each failure time. It is the standard first step of a reliability data analysis, because it provides a reference survival curve, determined by the data alone, against which parametric models are later checked (Section 7.6.1).

Without censoring, the reliability function would be estimated by the empirical fraction of survivors, $\hat{R}(t) = (\text{number with } T > t)/n$, as in the population examples of Chapter 5. With censoring the fraction is undefined: whether pump P3, last seen working at 900 h, is a survivor at 1000 h is unknown. The Kaplan–Meier estimator resolves this by using each unit only while it is observed.

7.3.2 Derivation from conditional survival probabilities

Let $t_1 < t_2 < \cdots < t_k$ be the distinct times at which failures are observed, and set $t_0 = 0$. Two counts are attached to each failure time t_j:

- n_j, the number of units at risk immediately before t_j: units neither failed nor censored before t_j;
- d_j, the number of failures observed at t_j.

The derivation takes three steps.

Step 1: survival as a product of conditional survivals. Survival beyond t_m requires survival through every earlier failure time. By the chain rule of conditional probability, the conditional reliability of Chapter 5 applied interval by interval,

$$R(t_m) = \Pr(T > t_m) = \prod_{j=1}^{m} \Pr(T > t_j \mid T > t_{j-1}),$$

since $\Pr(T > t_j) = \Pr(T > t_j \mid T > t_{j-1})\ \Pr(T > t_{j-1})$ and $R(t_0) = 1$.

Step 2: estimating each factor. Each factor is split at the failure time. Between t_{j-1} and t_j no failure was observed, so the data give no evidence of failure there, and the conditional survival over the open interval is estimated as one. At t_j itself, n_j units were at risk and d_j of them failed, so the conditional probability of surviving t_j, given survival to just before it, is estimated by the fraction of the risk set that survived,

$$\hat{p}_j = \frac{n_j - d_j}{n_j} = 1 - \frac{d_j}{n_j}.$$

A unit censored between t_{j-1} and t_j was at risk at t_{j-1} but is not counted in n_j: it contributes to every risk set while it is observed, and to none afterward. When a censoring time coincides with a failure time, the convention is that the censored unit was at risk at that failure.

Step 3: the product. Multiplying the estimated factors over all failure times up to t,

$$\hat{R}(t) = \prod_{j:\, t_j \le t} \left(1 - \frac{d_j}{n_j}\right), \tag{7.2}$$

with $\hat{R}(t) = 1$ before the first failure. This is the Kaplan–Meier product-limit estimator. It is a step function: it drops only at observed failure times and is constant between them. Censored observations do not lower it; they reduce the risk sets at later failure times, which makes each later failure count for more.

Without censoring the estimator reduces to the empirical fraction. Each risk set is then the previous one minus its failures, $n_{j+1} = n_j - d_j$, and the product telescopes: $\prod_{j \le m}(n_j - d_j)/n_j = (n_m - d_m)/n_1$, the number of survivors divided by the number of units. The estimator is therefore the natural generalization of the fraction surviving to data with censoring.

7.3.3 Calculation for the eight pumps

The estimator is applied to the pumps in time order. Initially eight pumps are at risk. At 400 h pump P1 fails, with $n_1 = 8$ and $d_1 = 1$; the conditional survival is $1 - 1/8 = 7/8$, so $\hat{R}(400) = 0.875$. At 700 h P2 fails; seven pumps were at risk just before, and

$$\hat{R}(700) = \frac{7}{8} \cdot \frac{6}{7} = 0.75.$$

At 900 h P3 is censored. The estimate does not decrease, because no failure occurred, but P3 is no longer counted among the pumps at risk at later failure times. At 1200 h P4 fails with five pumps at risk,

$$\hat{R}(1200) = 0.75 \left(1 - \frac{1}{5}\right) = 0.60.$$

At 1500 h P5 is censored. At 1600 h P6 fails with three pumps at risk,

$$\hat{R}(1600) = 0.60 \left(1 - \frac{1}{3}\right) = 0.40.$$

P7 and P8 are censored at 1800 and 2000 h and leave the estimate unchanged.

Table 7.3 collects the whole calculation. Each row is one record in time order; the columns give the time, the pump and the kind of event, the number at risk immediately before the event, the number of failures, the conditional survival factor $1 - d_j/n_j$ (one for a censoring), and the running product, which is the Kaplan–Meier estimate from that time until the next failure.

The first two engineering questions now have nonparametric answers. The estimated probability that a new pump survives 1000 h is $\hat{R}(1000) = 0.75$, the value of the step that runs from 700 to 1200 h. The estimated probability that it survives 2000 h is $\hat{R}(2000) = 0.40$. The last value does not establish that the survival probability stays at 0.40 beyond 2000 h; it means only that no pump was observed past that time.

7.3.4 The step-function survival curve

Figure 7.2 draws the estimate. The horizontal axis is operating hours and the vertical axis the estimated reliability. The curve is constant between failures and drops at each of the four failure times; the vertical tick marks on it are the four censoring times, where the curve does not drop. The drops grow larger with time: 1/8 of the risk set at 400 h, 1/7 at 700 h, 1/5 at 1200 h and 1/3 at 1600 h, because each failure is divided by a smaller risk set.

The final estimate is 0.40, although only four of the eight pumps failed. The fraction $4/8 = 0.50$ would count P3, P5 and P7 as survivors to 2000 h, but they were observed only to 900, 1500 and 1800 h, and nothing is known of them beyond. As pumps leave observation,

Table 7.3: Kaplan–Meier calculation for the eight pumps, equation (7.2). Censored rows have no failures; they leave the estimate unchanged and reduce the next risk set.

Time (h)	Pump	Event	At risk before event	Failures	Factor $1 - d_j/n_j$	$\hat{R}(t)$
0	–	start	8	0	–	1.000
400	P1	failure	8	1	7/8	0.875
700	P2	failure	7	1	6/7	0.750
900	P3	censored	6	0	1	0.750
1200	P4	failure	5	1	4/5	0.600
1500	P5	censored	4	0	1	0.600
1600	P6	failure	3	1	2/3	0.400
1800	P7	censored	2	0	1	0.400
2000	P8	censored	1	0	1	0.400

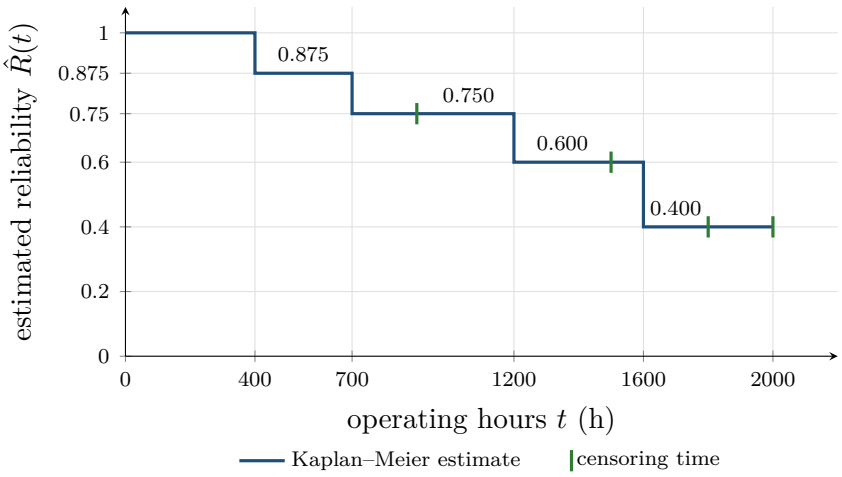


Figure 7.2: Kaplan–Meier estimate of the reliability function of the eight pumps (Table 7.3). The estimate is a right-continuous step function that drops only at failure times; censoring times, marked by ticks, reduce later risk sets but do not lower the curve.

fewer remain to inform later survival: at the failure at 1600 h only three pumps were at risk, so that failure removes a third of the remaining survival probability. Equivalently, the estimator starts with a probability of 1/8 on each pump and, at each censoring, shares the censored pump's probability equally among the pumps still under observation beyond it: after P3 is censored each of the five remaining pumps carries $1/8 + 1/40 = 0.15$, the drop at 1200 h; after P5 is censored each of the three remaining carries $0.15 + 0.05 = 0.20$, the drop at 1600 h. The Kaplan–Meier estimator accounts for these changing risk sets; the simple fraction does not.

7.3.5 The censoring assumption

The estimator requires *noninformative* (independent) censoring. Units removed from observation must not have systematically different failure prospects from comparable units that remain under observation, because the estimator lets the units still at risk stand for those that left. Step 2 of Section 7.3.2 estimated the conditional survival at t_j from the risk set alone; that estimate describes the whole population only if the censored units would have failed at the same rate as those that stayed.

Two examples mark the boundary. Ending a test at a predetermined time, or extracting field data on a fixed date, is generally compatible with independent censoring, since the end of observation has nothing to do with the condition of the unit. Removing a pump because its vibration has become dangerously high is informative censoring: the removed pump was closer to failure than the pumps left running. If deteriorating pumps are systematically removed before they fail, the failures that would have occurred never appear in the data, and a naive Kaplan–Meier analysis overestimates reliability. This limitation is important in real maintenance datasets, where preventive replacement and condition-based removal are common; such removals must be recorded as a separate outcome, and modeled as such, rather than as ordinary censoring.

7.3.6 Mean lifetime and restricted mean survival time

Chapter 5 showed that the mean time to failure is the area under the reliability function, $\mathrm{MTTF} = \mathbb{E}[T] = \int_0^\infty R(t)\,\mathrm{d}t$. The natural nonparametric estimate would replace R by $\hat{R}$, $\widehat{\mathrm{MTTF}} = \int_0^\infty \hat{R}(t)\,\mathrm{d}t$. For the pumps this integral does not exist as a number: the last records are censored, and the estimate remains at 0.40 beyond the last observed failure without any information about when the remaining pumps fail. The full mean lifetime cannot be determined nonparametrically from these records.

What can be estimated is the *restricted mean survival time* (RMST) over a specified horizon τ,

$$\mathrm{RMST}(\tau) = \int_0^\tau R(t)\,\mathrm{d}t = \mathbb{E}\big[\min(T, \tau)\big], \tag{7.3}$$

where the second form follows from the tail-integral identity of Chapter 5 applied to the lifetime capped at τ. For $\tau = 2000$ h, the integral of the step function is a sum of rectangles, one per step, of width the length of the step and height its value:

$$\begin{aligned}\widehat{\mathrm{RMST}}(2000) &= 400\,(1.000) + 300\,(0.875) + 500\,(0.750) + 400\,(0.600) + 400\,(0.400)\\ &= 400 + 262.5 + 375 + 240 + 160 = 1437.5\,\mathrm{h}.\end{aligned}$$

This is the estimated expected number of failure-free hours accumulated during the first 2000 h, with each unit credited with at most 2000 h. It is not the unrestricted MTTF. Since $\min(T, \tau) \le T$, it is a lower bound on the mean lifetime, and the gap is the expected life beyond the horizon, about which these data say nothing without a model. The distinction matters whenever the available data do not extend far enough to establish long-term reliability; a parametric model can extend the curve beyond the data, but only by assumption, a point taken up in Section 7.6.4.

7.4 Likelihood and maximum likelihood estimation with censoring

The Kaplan–Meier estimator gives a curve but no parameters. To estimate the exponential failure rate λ, or the Weibull parameters β and η, a method is needed that determines which parameter values are most consistent with the records. This section defines the likelihood, derives the contribution of each kind of record, specializes it to the exponential model, and derives the maximum likelihood estimate of the failure rate.

7.4.1 The likelihood function

Let θ denote the parameters of a candidate lifetime distribution (a vector, $\boldsymbol{\theta}$, when there are several), such as λ for the exponential model or (β, η) for the Weibull. The *likelihood* is the joint probability or density of the observed data under the model, considered as a function of θ with the data held fixed:

$$L(\theta) = L(\theta \mid \text{data}).$$

The *maximum likelihood estimate* is the parameter value that makes the observed failure and censoring records most probable under the chosen model,

$$\hat{\theta} = \arg\max_{\theta} L(\theta) = \arg\max_{\theta} \ell(\theta), \qquad \ell(\theta) = \ln L(\theta), \tag{7.4}$$

where the log-likelihood ℓ has the same maximizer, because the logarithm is increasing, and turns products into sums.

The likelihood is not the probability distribution of the unknown parameters. It is a function of the parameter values given the observed data; it does not integrate to one over θ, and the parameters are treated as fixed but unknown constants. Assigning probability distributions to the parameters is the approach of Chapter 8.

7.4.2 Likelihood contributions of failures and censored records

Each record contributes one factor to the likelihood, and the factor is the probability of what the record says, as Table 7.2 listed. For a failure observed at y_i the record says $T_i = y_i$; for a continuous lifetime the probability of that exact value is replaced by the density, and the contribution is $f(y_i \mid \theta)$. For a record censored at y_i the record says $T_i > y_i$, an event whose probability is the reliability function, and the contribution is $R(y_i \mid \theta)$.

The censoring times also have a distribution, and the derivation shows when it can be left out. Let the censoring time C_i have density f_C and survivor function R_C, independent of T_i. A failure observed at y requires $T_i = y$ and $C_i > y$, with probability density $f(y \mid \theta)\, R_C(y)$; a censoring at y requires $C_i = y$ and $T_i > y$, with density $f_C(y)\, R(y \mid \theta)$. Multiplying over independent units,

$$\prod_{i=1}^{n} \left[f(y_i \mid \theta)\, R_C(y_i) \right]^{\delta_i} \left[f_C(y_i)\, R(y_i \mid \theta) \right]^{1-\delta_i} = \underbrace{\prod_{i=1}^{n} R_C(y_i)^{\delta_i} f_C(y_i)^{1-\delta_i}}_{\text{free of } \theta} \prod_{i=1}^{n} f(y_i \mid \theta)^{\delta_i}\, R(y_i \mid \theta)^{1-\delta_i}.$$

When the censoring distribution does not involve θ, which is the noninformative-censoring assumption of Section 7.3.5, the first product is a constant that does not move the maximizer, and the likelihood for independent exact failures and right-censored observations is

$$L(\theta) = \prod_{i=1}^{n} f(y_i \mid \theta)^{\delta_i}\, R(y_i \mid \theta)^{1-\delta_i}. \tag{7.5}$$

The indicator does the bookkeeping: if $\delta_i = 1$ only the density term remains, and if $\delta_i = 0$ only the survival term remains. The censored records are not discarded, and no failure times are invented for them. The same reasoning gives the other factors of Table 7.2: an interval-censored record contributes $F(b) - F(a)$, a left-censored record $F(c)$, and a record left-truncated at age a has its factor divided by $R(a)$.

Figure 7.3 shows the two factors for two of the pumps, under the exponential model fitted below. The horizontal axis is operating time and the vertical axis the failure density. The failure of P2 at 700 h contributes the height of the density at 700 h; the censoring of P7 at 1800 h contributes the shaded area under the density beyond 1800 h, which is $R(1800)$, the probability of surviving past the censoring time. A failure record pins the lifetime to a point; a censored record spreads it over the whole tail.

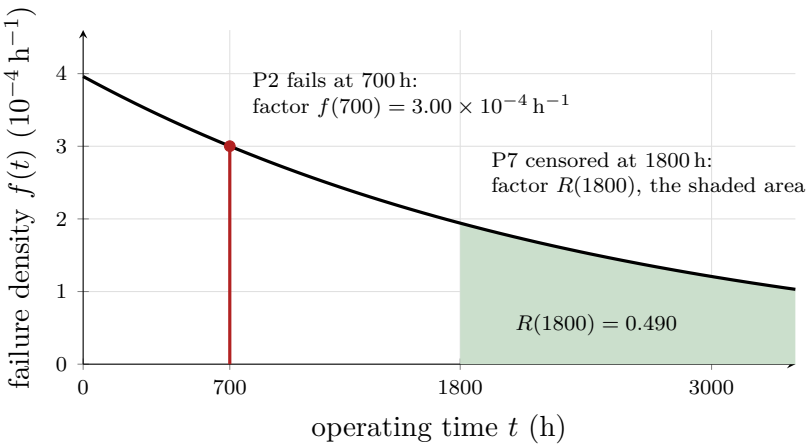


Figure 7.3: Likelihood contributions of a failure and of a censored record, under the exponential model fitted to the pumps, $f(t) = \hat{\lambda}\,\mathrm{e}^{-\hat{\lambda}t}$ with $\hat{\lambda} = 3.9604 \times 10^{-4}\,\mathrm{h}^{-1}$ (Section 7.4.4). A failure contributes the density at its time; a censored record contributes the probability of the tail beyond its censoring time.

Principle 7.2. *Maximum likelihood uses the probability model of the observation process. Failure records contribute densities, $f(y_i)$; right-censored records contribute survival probabilities, $R(y_i)$; and the censoring mechanism may be left out of the likelihood only when it carries no information about the lifetime parameters.*

7.4.3 Exponential likelihood and total exposure

Assume the exponential lifetime model, $f(t \mid \lambda) = \lambda\,\mathrm{e}^{-\lambda t}$ and $R(t \mid \lambda) = \mathrm{e}^{-\lambda t}$. Substituting into equation (7.5),

$$L(\lambda) = \prod_{i=1}^{n} \left(\lambda\,\mathrm{e}^{-\lambda y_i}\right)^{\delta_i} \left(\mathrm{e}^{-\lambda y_i}\right)^{1-\delta_i}.$$

Every record, failed or censored, carries the factor $\mathrm{e}^{-\lambda y_i}$, because $\delta_i + (1 - \delta_i) = 1$, and only the failed records carry the factor λ. Collecting the powers,

$$L(\lambda) = \lambda^{\sum_i \delta_i}\,\mathrm{e}^{-\lambda \sum_i y_i}.$$

Define the number of observed failures and the total exposure, or total time on test,

$$r = \sum_{i=1}^{n} \delta_i, \qquad T_{\mathrm{tot}} = \sum_{i=1}^{n} y_i. \tag{7.6}$$

Then the likelihood and the log-likelihood are

$$L(\lambda) = \lambda^{r}\,\mathrm{e}^{-\lambda T_{\mathrm{tot}}}, \qquad \ell(\lambda) = r\ln\lambda - \lambda\,T_{\mathrm{tot}}. \tag{7.7}$$

This is a useful reduction: for the exponential model the whole dataset enters the likelihood through two numbers, the count of observed failures and the total operating exposure of all units. They are the *sufficient statistics* of the exponential model; any two datasets with the same r and T_{tot} give the same inference about λ.

For the eight pumps the total exposure is

$$T_{\mathrm{tot}} = 400 + 700 + 900 + 1200 + 1500 + 1600 + 1800 + 2000 = 10\,100 \text{ pump-hours},$$

of which 3900 come from the failed pumps and 6200 from the censored ones, and the number of failures is $r = 4$. The exposure includes the operating time of the censored pumps; that is exactly how the surviving pumps enter the exponential estimate.

7.4.4 The maximum likelihood estimate

The maximum of equation (7.7) is found by setting the derivative of the log-likelihood to zero. The derivative, called the *score*, is

$$\frac{\mathrm{d}\ell}{\mathrm{d}\lambda} = \frac{r}{\lambda} - T_{\mathrm{tot}}.$$

Setting it to zero, $r/\lambda - T_{\mathrm{tot}} = 0$, gives

$$\hat{\lambda} = \frac{r}{T_{\mathrm{tot}}}\,: \tag{7.8}$$

the number of failures divided by the total time on test. The second derivative,

$$\frac{\mathrm{d}^2\ell}{\mathrm{d}\lambda^2} = -\frac{r}{\lambda^2},$$

is negative for $r > 0$, so the stationary point is the interior maximum of a concave function. For the pumps,

$$\hat{\lambda} = \frac{4}{10\,100\,\mathrm{h}} = 3.9604 \times 10^{-4}\,\mathrm{h}^{-1},$$

or about 0.396 failures per 1000 pump-hours.

The maximum likelihood estimate of any function of λ is that function of $\hat{\lambda}$, the *invariance property* of maximum likelihood. The estimated mean lifetime under the exponential model is therefore

$$\widehat{\mathrm{MTTF}} = \frac{1}{\hat{\lambda}} = \frac{T_{\mathrm{tot}}}{r} = \frac{10\,100}{4} = 2525\,\mathrm{h},$$

very different from the naive average of the failure times, 975 h. The difference arises because the likelihood includes the survival information of the censored pumps. The estimated reliability at 1000 h is

$$\hat{R}(1000) = \mathrm{e}^{-\hat{\lambda}\times 1000} = \mathrm{e}^{-0.39604} = 0.6730,$$

about 67.3 %, and at 2000 h it is $\mathrm{e}^{-0.79208} = 0.4529$.

7.4.5 Comparison with the Kaplan–Meier estimate

At 1000 h the two methods give different answers from the same records: the Kaplan–Meier estimate is $\hat{R}_{\mathrm{KM}}(1000) = 0.75$ and the exponential estimate is $\hat{R}_{\mathrm{exp}}(1000) = 0.6730$. At 2000 h they are 0.40 and 0.4529. The estimates differ because they rest on different assumptions. The Kaplan–Meier estimate is nonparametric and changes only at observed failure times; at 1000 h

it still reflects the fact that only two failures had occurred among the pumps at risk. The exponential model imposes a constant hazard over the whole lifetime, and so spreads the four failures evenly over the 10 100 pump-hours. The difference does not by itself establish which estimate is better. That requires the uncertainty of each (Section 7.5), a check of the fit of the model (Section 7.6.1) and a statement of the intended application.

7.4.6 Effect of censored exposure on the estimate

The role of the censored exposure is clearest when it is varied with the failures held fixed. Table 7.4 evaluates equation (7.8), the reliability $\mathrm{e}^{-\hat{\lambda}\times 1000}$ and the mean $1/\hat{\lambda}$ for several cases. The upper block keeps the four failures and their 3900 h of exposure and changes only the censored exposure: the actual dataset, a case in which the four survivors were withdrawn at 900 h each, cases in which the survivors were followed without failure to 3000 and to 5000 h each, and the case in which the censored exposure is ignored. The lower block keeps the exposure of the dataset, 10 100 h, and changes the number of failures. The cases other than the actual dataset are illustrative. One row of the arithmetic, for survivors followed to 3000 h: the censored exposure is $4\times 3000 = 12\,000$ h, $T_{\mathrm{tot}} = 3900 + 12\,000 = 15\,900$ h, $\hat{\lambda} = 4/15\,900 = 2.516\times 10^{-4}\ \mathrm{h}^{-1}$, and $\hat{R}(1000) = \mathrm{e}^{-0.2516} = 0.7776$.

Table 7.4: Exponential maximum likelihood estimate for varying censored exposure (upper block, four failures with 3900 h of failed-unit exposure) and for a varying number of failures (lower block, $T_{\mathrm{tot}} = 10\,100$ h). The eight-pump dataset is marked; the other cases are illustrative.

Case	r	Censored exposure (h)	T_{tot} (h)	$\hat{\lambda}$ ($10^{-4}\,\mathrm{h}^{-1}$)	$\hat{R}(1000)$	$1/\hat{\lambda}$ (h)
survivors withdrawn at 900 h	4	3600	7500	5.333	0.5866	1875
eight-pump dataset	4	6200	10 100	3.960	0.6730	2525
survivors followed to 3000 h	4	12 000	15 900	2.516	0.7776	3975
survivors followed to 5000 h	4	20 000	23 900	1.674	0.8459	5975
censored exposure ignored	4	0	3900	10.256	0.3586	975
no failure	0	–	10 100	0	1	–
one failure	1	–	10 100	0.990	0.9057	10 100
two failures	2	–	10 100	1.980	0.8203	5050
four failures (dataset)	4	–	10 100	3.960	0.6730	2525
eight failures	8	–	10 100	7.921	0.4529	1262.5

The upper block shows the trend: with the failure count fixed, each additional failure-free hour lowers the estimated failure rate, and ignoring the 6200 censored pump-hours multiplies it by $10\,100/3900 = 2.59$. The last row of the lower block, eight failures in the same exposure, is also what results from counting the four censoring times as failures; its mean, 1262.5 h, is the average of all eight times of Section 7.2.2. The table is a sensitivity calculation; an actual estimate must come from valid histories of the units, not from exposure assumed for them.

The row with no failures needs care. With $r = 0$ the log-likelihood $-\lambda T_{\mathrm{tot}}$ decreases for every $\lambda > 0$, and the maximum lies at the boundary, $\hat{\lambda} = 0$. This does not show that the failure rate is zero; it shows that the data cannot give a positive point estimate. The meaningful

statement is a one-sided upper confidence bound, derived in Section 7.5.3: with no failure in 10 100 h the 95 % upper bound on λ is $2.966 \times 10^{-4}\,\mathrm{h}^{-1}$.

7.5 Statistical uncertainty

A point estimate from four failures can look precise while resting on very little evidence. This section quantifies how much the data constrain the estimates: the standard error from the curvature of the log-likelihood, exact chi-square confidence limits for the exponential failure rate and the reliability limits they imply, Greenwood intervals for the Kaplan–Meier estimate, and the larger uncertainty of the two-parameter Weibull model.

7.5.1 Precision of a point estimate

The estimate $\hat{\lambda} = 3.9604 \times 10^{-4}\,\mathrm{h}^{-1}$ rests on four failures. Table 7.5 sets it beside two other studies with the same point estimate: study A is the eight-pump dataset, study B observes 400 failures in 1 010 000 h, a hundred times the failures and the exposure, and study C, an illustrative addition, lies in between with 40 failures in 101 000 h. The columns give the failures, the exposure, the estimate, its standard error and relative standard error (Section 7.5.2), and the 95 % chi-square confidence interval with the ratio of its limits (Section 7.5.3).

Table 7.5: Three studies with the same estimated failure rate. Rates in $10^{-4}\,\mathrm{h}^{-1}$. Study C is an illustrative addition; the intervals are the chi-square limits of equation (7.10) with $\alpha = 0.05$.

Study	r	T_{tot} (h)	$\hat{\lambda}$	SE	relative SE	95 % interval	λ_U/λ_L
A	4	10 100	3.960	1.980	50 %	1.079 to 10.140	9.40
C	40	101 000	3.960	0.626	15.8 %	2.829 to 5.393	1.91
B	400	1 010 000	3.960	0.198	5 %	3.582 to 4.368	1.22

The three studies produce the same point estimate, but study B contains a hundred times more failure information. If the model assumptions hold, its estimate is much more precise: the true rate lies, with 95 % confidence, within about ±10 % of the estimate, while for study A the interval spans almost a factor of ten. A point estimate reported without its uncertainty hides this difference completely. Confidence intervals quantify it.

Principle 7.3. *Point estimates are not enough. The precision of an estimated failure rate is set by the number of failures observed, and a small dataset can leave an uncertainty of the same order as the estimate itself.*

7.5.2 Fisher information and standard errors

The precision of a maximum likelihood estimate is measured by the curvature of the log-likelihood at its maximum: a sharply peaked log-likelihood singles out a narrow range of parameter values, a flat one does not. For the exponential log-likelihood of equation (7.7), the *observed information* is the negative second derivative,

$$J(\lambda) = -\frac{\mathrm{d}^2\ell}{\mathrm{d}\lambda^2} = \frac{r}{\lambda^2}.$$

The *Fisher information* is its expectation over repeated samples,

$$\mathcal{I}(\lambda) = \mathbb{E}\left[-\frac{\mathrm{d}^2\ell}{\mathrm{d}\lambda^2}\right] = \frac{\mathbb{E}[r]}{\lambda^2},$$

where the expected number of failures depends on the censoring design. A standard result of likelihood theory states that, for large samples, the MLE is approximately normally distributed about the true value with variance $1/\mathcal{I}(\lambda)$. Since $\mathcal{I}$ involves the unknown λ and the design, it is estimated by the observed information at the estimate,

$$J(\hat\lambda) = \frac{r}{\hat\lambda^2} = \frac{T_{\text{tot}}^2}{r},$$

and the approximate standard error (SE) of $\hat\lambda$ is

$$\operatorname{SE}(\hat\lambda) \approx \frac{1}{\sqrt{J(\hat\lambda)}} = \frac{\hat\lambda}{\sqrt{r}}. \tag{7.9}$$

For the four pump failures,

$$\operatorname{SE}(\hat\lambda) \approx \frac{3.9604 \times 10^{-4}}{\sqrt{4}} = 1.9802 \times 10^{-4}\,\text{h}^{-1},$$

half the estimate itself. The relative standard error is $1/\sqrt{r}$: it depends on the number of failures only, not on the exposure, so it is the failures that carry the information about the rate. Four failures give 50 %, 40 give 15.8 % and 400 give 5 %, the column of Table 7.5. The uncertainty propagates to any function of λ by the first-order (delta-method) rule $\operatorname{SE}[g(\hat\lambda)] \approx |g'(\hat\lambda)|\operatorname{SE}(\hat\lambda)$; for $\hat R(t) = \mathrm{e}^{-\hat\lambda t}$ it gives $\operatorname{SE}[\hat R(1000)] \approx 1000 \times 0.6730 \times 1.9802 \times 10^{-4} = 0.133$.

Figure 7.4 shows the curvature directly. The horizontal axis is the failure rate, in $10^{-4}\,\text{h}^{-1}$; the vertical axis is the log-likelihood relative to its maximum, $\ell(\lambda) - \ell(\hat\lambda)$, which for both studies is $r[\ln(\lambda/\hat\lambda) - \lambda/\hat\lambda + 1]$. Both curves peak at $\hat\lambda = 3.96$; the curve of study A, with four failures, is broad and asymmetric, and the curve of study B, with 400, is a hundred times more sharply curved. The dotted line at -1.92 marks the likelihood-ratio interval of Section 7.5.3: the parameter values whose log-likelihood lies within 1.92 of the maximum.

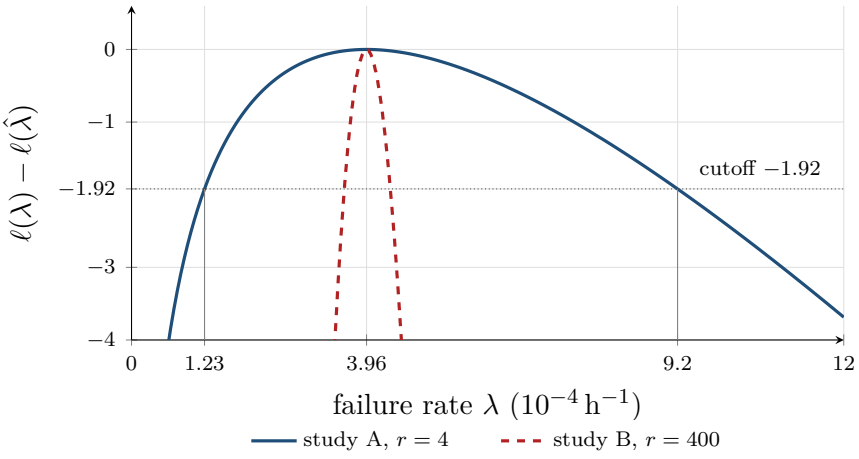


Figure 7.4: Relative log-likelihood of the exponential failure rate, $\ell(\lambda) - \ell(\hat\lambda) = r[\ln(\lambda/\hat\lambda) - \lambda/\hat\lambda + 1]$, for study A ($r = 4$, $T_{\text{tot}} = 10\,100\,\text{h}$) and study B ($r = 400$, $T_{\text{tot}} = 1\,010\,000\,\text{h}$), both with $\hat\lambda = 3.96 \times 10^{-4}\,\text{h}^{-1}$. The curvature at the peak is the observed information. The values within 1.92 of the maximum form the 95 % likelihood-ratio interval: 1.23 to 9.20 for A, 3.58 to 4.36 for B.

A conventional large-sample interval, $\hat\lambda \pm 1.96\operatorname{SE}(\hat\lambda)$, rests on the normal approximation, which is questionable with only four failures: the log-likelihood of study A in Figure 7.4 is visibly not a parabola, and the symmetric interval reaches almost to zero. A small-sample method is preferable.

7.5.3 Chi-square confidence limits for the exponential model

For the exponential model, confidence limits can be computed from the exact distribution of the data rather than from a large-sample approximation. Write $\chi^2_{\nu,p}$ for the p-quantile of the chi-square distribution with ν degrees of freedom, so that $\Pr(\chi^2_\nu \le \chi^2_{\nu,p}) = p$. Consider a time-terminated test with a total exposure T_{tot} fixed in advance, in which failures occur as a Poisson process in accumulated exposure with rate λ; this holds exactly for exponential lifetimes when each failed unit is replaced at once. The number of failures N then has a Poisson distribution with mean λT_{tot}. The limits follow in four steps.

Step 1: counts and waiting times. Let W_r be the exposure accumulated at the r-th failure. At least r failures occur within T_{tot} exactly when the r-th failure comes before T_{tot}: $\Pr(N \ge r) = \Pr(W_r \le T_{\text{tot}})$.

Step 2: the distribution of the waiting time. W_r is the sum of r independent exponential inter-failure exposures with rate λ. For one of them, X, the scaled variable $2\lambda X$ satisfies $\Pr(2\lambda X > x) = \mathrm{e}^{-x/2}$, which is the chi-square distribution with two degrees of freedom; a sum of r independent such variables is chi-square with $2r$ degrees of freedom. Hence $2\lambda W_r \sim \chi^2_{2r}$ and

$$\Pr(N \ge r) = \Pr\bigl(\chi^2_{2r} \le 2\lambda T_{\text{tot}}\bigr).$$

Step 3: the lower limit. The lower limit λ_L is the rate at which observing r or more failures is just improbable enough, $\Pr(N \ge r) = \alpha/2$, so that $2\lambda_L T_{\text{tot}} = \chi^2_{2r,\alpha/2}$.

Step 4: the upper limit. The upper limit λ_U is the rate at which observing r or fewer failures has probability $\alpha/2$, $\Pr(N \le r) = 1 - \Pr(N \ge r+1) = \alpha/2$, so that $\Pr(\chi^2_{2r+2} \le 2\lambda_U T_{\text{tot}}) = 1 - \alpha/2$ and $2\lambda_U T_{\text{tot}} = \chi^2_{2r+2,1-\alpha/2}$.

The central two-sided $100(1-\alpha)\,\%$ confidence interval for the constant failure rate is therefore

$$\lambda_L = \frac{\chi^2_{2r,\,\alpha/2}}{2T_{\text{tot}}}, \qquad \lambda_U = \frac{\chi^2_{2(r+1),\,1-\alpha/2}}{2T_{\text{tot}}}, \qquad r > 0. \tag{7.10}$$

Because the failure count is discrete, the coverage of this interval is at least $1-\alpha$; it is exact in the sense that it uses the exact distribution of the count, and conservative in the sense that it may cover more often than stated. For a failure-terminated (Type II) test the roles are reversed: T_{tot} is random, $2\lambda T_{\text{tot}} \sim \chi^2_{2r}$ exactly, and both limits use $2r$ degrees of freedom. For a time-terminated test without replacement, and for data with individual censoring times such as the pumps, the stopping and censoring design differs from both, equation (7.10) is an approximation, and the familiar interval should not automatically be called exact. Likelihood-ratio intervals, or other methods calibrated for the actual design, can be used instead.

For the pumps, with $r = 4$, $T_{\text{tot}} = 10\,100\,\text{h}$ and $\alpha = 0.05$, the quantiles are $\chi^2_{8,0.025} = 2.180$ and $\chi^2_{10,0.975} = 20.483$, and

$$\lambda_L = \frac{2.180}{20\,200\,\text{h}} = 1.079 \times 10^{-4}\,\text{h}^{-1}, \qquad \lambda_U = \frac{20.483}{20\,200\,\text{h}} = 1.0140 \times 10^{-3}\,\text{h}^{-1}.$$

Since the pumps have individual censoring times, these values are a reference calculation, not an exact 95 % interval justified by this censoring design. The interval is wide: its upper limit is 9.4 times its lower limit, which reflects how little information four failures provide. The failure-terminated form would give $\lambda_U = \chi^2_{8,0.975}/(2T_{\text{tot}}) = 17.535/20\,200 = 8.68 \times 10^{-4}\,\text{h}^{-1}$, with the same lower limit.

Table 7.6 compares four 95 % intervals for the pumps' failure rate. Each row is a method; the columns give its lower and upper limit on λ and the reliability limits at 1000 h that follow

from them (Section 7.5.4). The Wald interval is $\hat\lambda \pm 1.96\,\mathrm{SE}$. The log-transformed Wald interval applies the same rule to $\ln\hat\lambda$, whose standard error is $1/\sqrt{r} = 0.5$, giving $\hat\lambda\,\mathrm{e}^{\pm 1.96/\sqrt{r}}$. The likelihood-ratio interval is the set of λ with $2[\ell(\hat\lambda) - \ell(\lambda)] \le \chi^2_{1,0.95} = 3.841$, read off Figure 7.4; for a one-parameter model it is the profile-likelihood interval. The chi-square row is equation (7.10).

Table 7.6: Four 95 % confidence intervals for the failure rate of the eight pumps ($r = 4$, $T_{\mathrm{tot}} = 10\,100$ h), with the reliability limits at 1000 h obtained from them by equation (7.12).

Method	λ_L ($10^{-4}\,\mathrm{h}^{-1}$)	λ_U ($10^{-4}\,\mathrm{h}^{-1}$)	$R_L(1000)$	$R_U(1000)$
Wald, $\hat\lambda \pm 1.96\,\mathrm{SE}$	0.079	7.842	0.457	0.992
log-transformed Wald	1.486	10.552	0.348	0.862
likelihood ratio	1.230	9.200	0.399	0.884
chi-square, equation (7.10)	1.079	10.140	0.363	0.898

The three methods that respect the skewness of the problem agree in shape: lower limits near 1.1 to 1.5, upper limits near 9 to 10.6, an interval roughly from a third of the estimate to two and a half times it. The Wald interval is the outlier: it is symmetric, nearly touches zero, and understates the upper limit. With few failures the normal approximation is not adequate.

When no failure has been observed, only an upper limit is meaningful. With $r = 0$, Step 4 with a one-sided level α requires $\Pr(N = 0) = \mathrm{e}^{-\lambda_U T_{\mathrm{tot}}} = \alpha$, so

$$\lambda_U = \frac{-\ln\alpha}{T_{\mathrm{tot}}} = \frac{\chi^2_{2,\,1-\alpha}}{2T_{\mathrm{tot}}}, \tag{7.11}$$

since $\chi^2_{2,1-\alpha} = -2\ln\alpha$. With $\alpha = 0.05$ and 10 100 h of failure-free exposure this is $5.991/20\,200 = 2.966 \times 10^{-4}\,\mathrm{h}^{-1}$, the bound quoted in Section 7.4.6; the corresponding lower bound on the 1000 h reliability is $\mathrm{e}^{-0.2966} = 0.743$.

7.5.4 Confidence limits for reliability

Under the exponential model, $R(t) = \mathrm{e}^{-\lambda t}$ decreases as λ increases, for every fixed t. A monotone function maps an interval for λ into an interval for $R(t)$ with the same confidence, with the ends exchanged because the function is decreasing:

$$R_L(t) = \mathrm{e}^{-\lambda_U t}, \qquad R_U(t) = \mathrm{e}^{-\lambda_L t}. \tag{7.12}$$

At $t = 1000$ h, with the chi-square limits,

$$R_L(1000) = \mathrm{e}^{-1.0140} = 0.363, \qquad R_U(1000) = \mathrm{e}^{-0.1079} = 0.898.$$

The data and the model support a 1000 h reliability anywhere from about 36 % to 90 %, conditional on the chosen interval method and on the exponential model, around the point estimate of 67.3 %.

Figure 7.5 draws these limits over the observed range. The horizontal axis is operating time and the vertical axis reliability. The solid curve is the fitted reliability $\mathrm{e}^{-\hat\lambda t}$; the shaded band lies between $\mathrm{e}^{-\lambda_U t}$ and $\mathrm{e}^{-\lambda_L t}$; the three dots are the values at 1000 h. The band widens with time, from zero width at installation to 0.13 to 0.81 at 2000 h, because the uncertainty in the rate is multiplied by the time over which it acts.

A numerical point estimate, 67.3 %, can look precise even when the evidence behind it is weak. Small failure samples set severe limits on what any exponential reliability statement can claim, and the confidence limits are the honest form of the result.

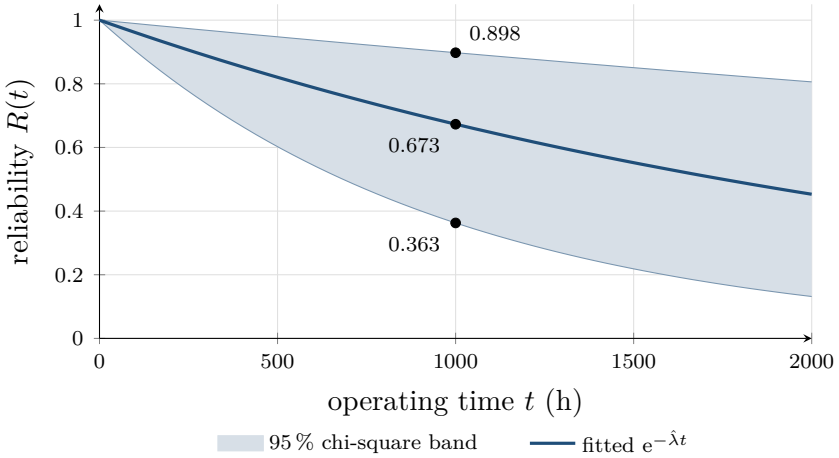


Figure 7.5: Exponential reliability of the pumps with its 95 % chi-square band: fitted curve $e^{-\hat{\lambda}t}$ with $\hat{\lambda} = 3.960 \times 10^{-4}\,\mathrm{h}^{-1}$, band between $e^{-\lambda_U t}$ and $e^{-\lambda_L t}$ with $\lambda_L = 1.079 \times 10^{-4}$ and $\lambda_U = 1.014 \times 10^{-3}\,\mathrm{h}^{-1}$ (equation (7.12)). Dots: the limits and the estimate at 1000 h.

7.5.5 Greenwood intervals for the Kaplan–Meier estimate

The Kaplan–Meier estimate has its own variance estimator, Greenwood's formula,

$$\widehat{\mathrm{Var}}[\hat{R}(t)] = \hat{R}(t)^2 \sum_{j:\, t_j \le t} \frac{d_j}{n_j(n_j - d_j)}, \tag{7.13}$$

with n_j the number at risk immediately before failure time t_j and d_j the number of failures there. It follows from the product form of the estimator in three steps. The logarithm turns the product into a sum, $\ln \hat{R}(t) = \sum_{t_j \le t} \ln \hat{p}_j$. Each factor $\hat{p}_j = 1 - d_j/n_j$ is the surviving fraction of a risk set of n_j units, a binomial proportion with variance $p_j(1-p_j)/n_j$, and the factors at different failure times are approximately uncorrelated. By the first-order rule of Section 7.5.2, $\mathrm{Var}[\ln \hat{p}_j] \approx \mathrm{Var}[\hat{p}_j]/p_j^2 = (1-p_j)/(n_j p_j)$, which with p_j replaced by $\hat{p}_j$ is $d_j/[n_j(n_j - d_j)]$; summing gives $\mathrm{Var}[\ln \hat{R}(t)]$, and a second application of the rule, $\mathrm{Var}[\hat{R}] \approx \hat{R}^2\, \mathrm{Var}[\ln \hat{R}]$, gives equation (7.13).

For the pumps at 1200 h, $\hat{R}(1200) = 0.60$ and the failures up to that time occurred at 400, 700 and 1200 h, with risk sets of 8, 7 and 5:

$$\widehat{\mathrm{Var}}[\hat{R}(1200)] = 0.60^2 \left[\frac{1}{8 \cdot 7} + \frac{1}{7 \cdot 6} + \frac{1}{5 \cdot 4} \right] = 0.36 \times 0.09167 = 0.0330,$$

and the standard error is $\sqrt{0.0330} = 0.182$, large relative to the estimate of 0.60. With only eight units, the nonparametric estimates are quite uncertain.

The symmetric interval $\hat{R} \pm 1.96\,\mathrm{SE}$ can leave the range $[0, 1]$, and transformed intervals are preferred. The *log-log* interval applies the normal approximation to $v = \ln[-\ln \hat{R}(t)]$, which ranges over the whole real line. By the first-order rule, $\mathrm{d}v/\mathrm{d}R = 1/(R \ln R)$, so

$$\sigma_v^2 = \frac{\widehat{\mathrm{Var}}[\hat{R}]}{(\hat{R} \ln \hat{R})^2} = \frac{\sum_{t_j \le t} d_j/[n_j(n_j - d_j)]}{(\ln \hat{R})^2},$$

and transforming $v \pm 1.96\,\sigma_v$ back through $R = \exp(-e^v)$ gives the limits

$$R_L(t) = \hat{R}(t)^{\exp(+1.96\,\sigma_v)}, \qquad R_U(t) = \hat{R}(t)^{\exp(-1.96\,\sigma_v)}, \tag{7.14}$$

which always lie between zero and one. Table 7.7 gives both intervals at each failure time. The columns are the failure time, the risk set, the estimate, the Greenwood sum, the variance and standard error, the symmetric interval and the log-log interval. At 1200 h, $\sigma_v =$

$\sqrt{0.09167}/|\ln 0.60| = 0.3028/0.5108 = 0.5927$, so $\mathrm{e}^{\pm 1.96\,\sigma_v} = \mathrm{e}^{\pm 1.1617}$ is 3.195 or 0.313, and the log-log interval is $0.60^{3.195} = 0.195$ to $0.60^{0.313} = 0.852$.

Table 7.7: Greenwood variance, equation (7.13), and 95 % intervals for the Kaplan–Meier estimate of the pumps. The symmetric interval exceeds one at 400 and 700 h; the log-log interval, equation (7.14), stays within $[0, 1]$. The Greenwood sum is $\sum_{t_j \le t} d_j/[n_j(n_j - d_j)]$; every $d_j = 1$.

t_j (h)	n_j	$\hat{R}(t_j)$	Greenwood sum	Variance	SE	symmetric	log-log
400	8	0.875	0.01786	0.01367	0.117	0.646 to 1.104	0.387 to 0.981
700	7	0.750	0.04167	0.02344	0.153	0.450 to 1.050	0.315 to 0.931
1200	5	0.600	0.09167	0.03300	0.182	0.244 to 0.956	0.195 to 0.852
1600	3	0.400	0.25833	0.04133	0.203	0.002 to 0.798	0.066 to 0.734

The estimate at 1000 h is the value of the step that began at 700 h, so its interval is the second row: 0.75 with a standard error of 0.153 and a log-log interval from 0.315 to 0.931. Both the exponential estimate, 0.673, and its chi-square interval lie well inside that range.

7.5.6 Weibull maximum likelihood with censoring

The likelihood construction extends to any lifetime model. For the Weibull distribution of Chapter 6,

$$R(t) = \mathrm{e}^{-(t/\eta)^{\beta}}, \qquad f(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \mathrm{e}^{-(t/\eta)^{\beta}},$$

the logarithms of the two factors of equation (7.5) are $\ln f(y) = \ln\beta - \beta\ln\eta + (\beta-1)\ln y - (y/\eta)^{\beta}$ and $\ln R(y) = -(y/\eta)^{\beta}$. Summing $\delta_i \ln f(y_i) + (1-\delta_i)\ln R(y_i)$ over the units, the term $-(y_i/\eta)^{\beta}$ appears once for every unit and the others only for the failures:

$$\ell(\beta, \eta) = r\ln\beta - r\beta\ln\eta + (\beta - 1)\sum_{i:\,\delta_i = 1} \ln y_i - \sum_{i=1}^{n}\left(\frac{y_i}{\eta}\right)^{\beta}. \tag{7.15}$$

Two features matter. The logarithmic failure-time term contains only the observed failures. The final sum contains every pump, the censored ones included: it is the total cumulative hazard accumulated by all units over their observed times, the Weibull counterpart of the total exposure λT_{tot}, to which it reduces when $\beta = 1$.

Unlike the exponential case, the maximum has no closed form, but the search reduces to one dimension. Setting $\partial\ell/\partial\eta = -r\beta/\eta + (\beta/\eta)\sum_i (y_i/\eta)^{\beta} = 0$ gives the scale for a given shape,

$$\hat{\eta}(\beta)^{\beta} = \frac{1}{r}\sum_{i=1}^{n} y_i^{\beta}, \tag{7.16}$$

which for $\beta = 1$ is T_{tot}/r, the exponential mean. Substituting it into $\partial\ell/\partial\beta = 0$ leaves one equation for the shape,

$$\frac{r}{\beta} + \sum_{i:\,\delta_i = 1} \ln y_i - r\,\frac{\sum_i y_i^{\beta}\ln y_i}{\sum_i y_i^{\beta}} = 0,$$

solved numerically. For the pumps, with $\sum_{\delta_i = 1} \ln y_i = 27.0104$, the solution is $\hat{\beta} = 1.737$ and, from equation (7.16), $\hat{\eta} = 2001$ h. The maximized log-likelihood is -34.698, against -35.336 for the exponential model. The fitted Weibull gives $\hat{R}(1000) = 0.741$ and $\hat{R}(2000) = 0.368$, and its hazard rises from $3.1 \times 10^{-4}\,\mathrm{h}^{-1}$ at 500 h to $8.7 \times 10^{-4}\,\mathrm{h}^{-1}$ at 2000 h.

The point estimate $\hat\beta = 1.74$ suggests wear-out, but the dataset contains only four failures and now has to determine two parameters. The observed information matrix, the negative matrix of second derivatives of equation (7.15) at the maximum, inverts to standard errors of 0.76 for $\hat\beta$ and 624 h for $\hat\eta$, with a correlation of -0.38 between them. A better picture comes from the *profile log-likelihood* of the shape, $\ell_p(\beta) = \ell(\beta, \hat\eta(\beta))$, the log-likelihood maximized over the scale for each fixed shape. Substituting equation (7.16) into equation (7.15), where the last sum becomes r,

$$\ell_p(\beta) = r \ln \beta - r \ln\Big(\frac{1}{r}\sum_{i=1}^{n} y_i^{\beta}\Big) + (\beta - 1)\sum_{i:\,\delta_i=1} \ln y_i - r.$$

Figure 7.6 plots it. The horizontal axis is the shape β and the vertical axis $\ell_p(\beta) - \ell_p(\hat\beta)$. The peak is at $\hat\beta = 1.74$; the dotted line at -1.92 cuts the curve at 0.61 and 3.63, the 95 % profile-likelihood interval for the shape; the marked point at $\beta = 1$, the exponential model, lies only 0.64 below the peak. The interval contains shapes below one, a decreasing hazard, as well as shapes above three, a strongly increasing one. The symmetric Wald interval, $1.74 \pm 1.96 \times 0.76 = 0.24$ to 3.23, is distorted in the same way as the exponential Wald interval.

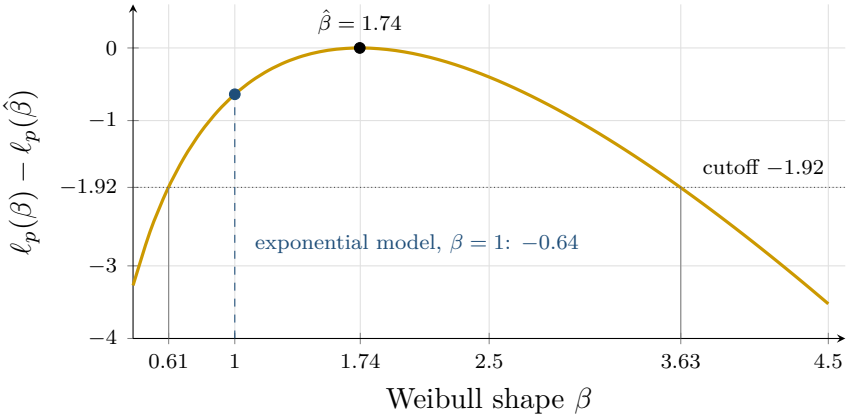


Figure 7.6: Profile log-likelihood of the Weibull shape for the eight pumps, $\ell_p(\beta) = r\ln\beta - r\ln(\sum_i y_i^{\beta}/r) + (\beta-1)\sum_{\delta_i=1}\ln y_i - r$, relative to its maximum at $\hat\beta = 1.737$ ($\hat\eta = 2001$ h), evaluated numerically at the plotted points. The 95 % profile-likelihood interval, where the curve is above -1.92, runs from 0.61 to 3.63 and contains the exponential model $\beta = 1$.

Even when a numerical optimizer returns precise-looking values, two parameters estimated from four failures are poorly constrained. In particular, the data leave considerable uncertainty about whether the hazard is decreasing, constant or increasing, which is the question a Weibull fit is usually asked to answer. This is an important warning when Weibull estimates are drawn from sparse field failure records.

7.6 Model diagnostics and engineering credibility

An estimate is useful only if its model is reasonable for the population and its operating conditions. This section states four diagnostic checks, distinguishes the uncertainty a confidence interval measures from the uncertainty of the model itself, and describes how the analysis becomes an engineering report.

7.6.1 Agreement with the nonparametric estimate

The first check, A, asks whether the fitted survival function agrees with the observations. The exponential estimate, $\hat R_{\exp}(t) = \mathrm{e}^{-0.000396t}$, is compared with the Kaplan–Meier estimate, which

is determined by the data alone. Table 7.8 lists both at each observed time, together with the log-log interval of the Kaplan–Meier estimate and the fitted Weibull reliability of Section 7.5.6. Figure 7.7 draws the same comparison: the horizontal axis is operating time and the vertical axis reliability; the step function is the Kaplan–Meier estimate, the thin dashed steps are its 95 % log-log limits, the solid curve is the exponential fit and the dashed curve the Weibull fit.

Table 7.8: Nonparametric and parametric survival estimates of the eight pumps at the observed times. Kaplan–Meier with its 95 % log-log interval; exponential with $\hat{\lambda} = 3.9604 \times 10^{-4}\,\mathrm{h}^{-1}$; Weibull with $\hat{\beta} = 1.737$, $\hat{\eta} = 2001\,\mathrm{h}$.

Time (h)	Kaplan–Meier	log-log interval	exponential	Weibull
400	0.875	0.387 to 0.981	0.8535	0.9408
700	0.750	0.315 to 0.931	0.7579	0.8510
900	0.750	0.315 to 0.931	0.7002	0.7791
1200	0.600	0.195 to 0.852	0.6217	0.6626
1500	0.600	0.195 to 0.852	0.5521	0.5453
1600	0.400	0.066 to 0.734	0.5306	0.5075
1800	0.400	0.066 to 0.734	0.4902	0.4350
2000	0.400	0.066 to 0.734	0.4529	0.3681

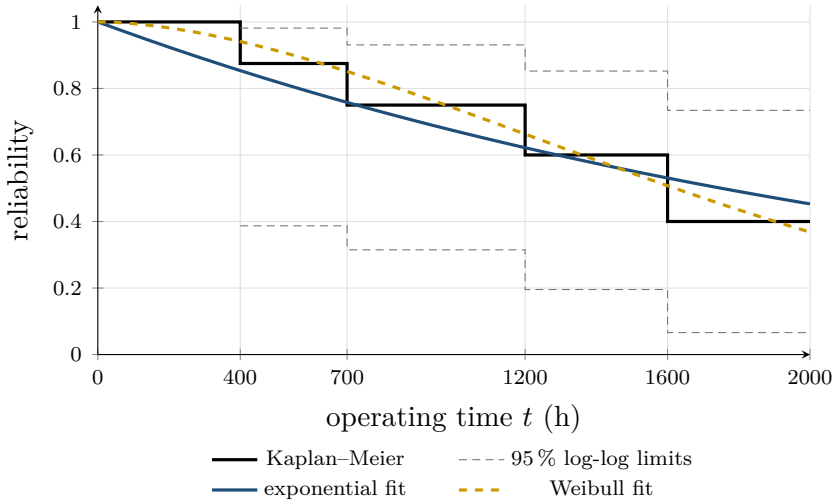


Figure 7.7: Kaplan–Meier estimate of the eight pumps with its 95 % log-log limits (equation (7.14)), against the exponential fit $\mathrm{e}^{-\hat{\lambda}t}$, $\hat{\lambda} = 3.96 \times 10^{-4}\,\mathrm{h}^{-1}$, and the Weibull fit $\mathrm{e}^{-(t/\hat{\eta})^{\hat{\beta}}}$, $\hat{\beta} = 1.737$, $\hat{\eta} = 2001\,\mathrm{h}$. Both fits stay inside the nonparametric limits over the whole observed range.

The curves differ, most visibly after the failure at 1600 h, where the Kaplan–Meier estimate drops to 0.40 and the exponential curve passes 0.53. The sample is too small for such differences to support a strong conclusion: both parametric curves lie inside the nonparametric limits at every observed time. The step function records the observed failures; the exponential curve smooths them under a constant-hazard assumption. A diagnostic comparison must consider both the differences and their statistical uncertainty.

7.6.2 Evidence of an age-dependent hazard

The second check, B, asks whether the hazard changes with age. The exponential model assumes $\lambda(t) = \lambda$; the Weibull model allows

$$\lambda(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1},$$

which increases with age when $\beta > 1$. If the fitted shape is credibly greater than one, the data support an increasing-hazard interpretation. A point estimate of β above one does not by itself prove aging; its uncertainty must be examined and the competing models compared. For the pumps the comparison is a likelihood-ratio test of the exponential model, $\beta = 1$, nested inside the Weibull. The statistic is twice the gain in maximized log-likelihood,

$$2\big[\ell(\hat{\beta}, \hat{\eta}) - \ell(\hat{\lambda})\big] = 2\,[-34.698 - (-35.336)] = 1.276,$$

which under the exponential model is approximately chi-square with one degree of freedom, the number of parameters added. It lies well below the 5 % critical value 3.841; the probability of a value at least this large under a constant hazard is 0.26. The data are consistent with a constant hazard and also with a wear-out shape near 1.7; four failures cannot separate them, as the profile interval of Figure 7.6 showed. Physical evidence, the failure mechanism found at each teardown, is what can settle the question when the data cannot.

7.6.3 Homogeneity of the population and of the failure modes

The two remaining checks concern what was pooled into one dataset.

C: comparability of the observations. Suppose four of the pumps operate under high loads and the other four under low loads. Pooling their records into one lifetime distribution then produces an estimate that describes neither group, and the mixture can show a hazard shape, typically a decreasing one, that no individual pump has, as the mixed population of Chapter 5 did. Before pooling, the records should be checked for differences in operating conditions and duty cycles, equipment manufacturer and design revisions, maintenance practices, installation environment, failure modes and causes, and changes in operating conditions over time. A single pooled distribution is inappropriate when the population contains meaningful subgroups; the subgroups are then estimated separately, or the conditions enter the hazard as covariates, as in the proportional hazards model of Chapter 5.

D: mixing of failure modes. A pump can fail by bearing seizure, by motor insulation breakdown or by seal leakage, and these failure modes follow different mechanisms with different age dependencies (Chapter 4). If all failures are combined into one dataset, the aggregate hazard can obscure those distinctions: a wear-out mode and a random mode together may look nearly constant. Depending on the engineering question, cause-specific reliability models, or a *competing-risks* analysis, are needed. In a competing-risks analysis each mode has its own lifetime, the unit fails at the earliest of them, and a failure by one mode is a censored observation for every other mode: the likelihood of equation (7.5) is formed mode by mode, with the other modes' failures entering as censored records, under the assumption that the modes are independent.

7.6.4 Confidence intervals and model uncertainty

At least three sources of uncertainty affect a reliability study. Table 7.9 lists them, each with an example from the pumps and the means of addressing it. Sampling uncertainty arises because

only a finite set of units was observed; parameter uncertainty is the resulting uncertainty in the estimated parameters; model uncertainty concerns whether the lifetime model, or the way the data were pooled, is right at all. All three are epistemic uncertainty in the sense of the notation chapter: lack of knowledge, which more data or better physics can reduce, as distinct from the aleatory variability of the lifetimes themselves.

Table 7.9: Sources of uncertainty in a reliability estimate, with an example from the pump analysis and the means of addressing each.

Source	Example	How it is addressed
Sampling uncertainty	only four observed failures	confidence intervals, likelihood profiling, additional data
Parameter uncertainty	exact value of the Weibull shape β unknown	parameter intervals, covariance matrix, Bayesian posterior
Model uncertainty	exponential against Weibull against competing risks	model comparison, sensitivity analysis, physics-of-failure evidence

Parameter uncertainty and sampling uncertainty are closely connected: sampling variability is what produces the uncertainty in the estimated parameters. Model uncertainty is different in kind. A confidence interval is computed under the model; a very narrow interval under the wrong lifetime distribution still produces an unreliable engineering prediction.

The difference becomes decisive when a model is used beyond its observation window. The exponential and Weibull fits of the pumps agree reasonably within the observed 2000 h (0.673 against 0.741 at 1000 h, 0.453 against 0.368 at 2000 h), and the likelihood-ratio test cannot choose between them. At 5000 h the same two fits predict a reliability of $\mathrm{e}^{-1.980} = 0.138$ and $\mathrm{e}^{-(5000/2001)^{1.737}} = 0.0074$, a factor of nineteen apart, and at 10 000 h the gap is many orders of magnitude. Estimating reliability at 20 000 h from observations that end at 2000 h requires assumptions about how the failure behavior evolves outside the observed range, and a smooth extrapolated curve is not evidence that those assumptions are correct.

Common mistake

Reading a narrow confidence interval as confirmation of the model. A confidence interval quantifies sampling uncertainty under the model's assumptions; it says nothing about whether the exponential or any other distribution is adequate, and nothing about behavior outside the observed range of ages.

Principle 7.4. *Statistical precision is not model credibility. A reliability prediction also depends on the definition of failure and of the system, on operating conditions representative of the intended use, and on failure mechanisms that make the chosen model physically plausible.*

7.6.5 From statistical analysis to an engineering report

Suppose the analysis is to be reported as a reliability assessment of the pump fleet for the manufacturer. Table 7.10 gives a defensible summary. Its rows are the data (pumps observed, failures, total exposure), the method, the estimates with their confidence limits, and the limitations. The confidence limits and the Kaplan–Meier row are additions of this chapter to the minimal summary of four numbers and a method.

The limits on the MTTF are $1/\lambda_U$ and $1/\lambda_L$, by the same monotone transformation as equation (7.12). Beyond the summary, a report should include the underlying data, the

Table 7.10: Illustrative reliability assessment of the cooling pump fleet. Limits are the 95 % chi-square reference limits of Section 7.5.3 and the log-log limits of Section 7.5.5.

Item	Value
Pumps observed	8
Observed failures	4
Total exposure	10 100 pump-hours
Statistical method	exponential maximum likelihood estimation using the observed failures and the right-censored exposure
Estimated failure rate	$\hat{\lambda} = 3.96 \times 10^{-4}\,\mathrm{h}^{-1}$ (limits 1.08×10^{-4} to $1.01 \times 10^{-3}\,\mathrm{h}^{-1}$)
Estimated MTTF	2525 h (limits 986 to 9268 h)
Estimated reliability at 1000 h	67.3 % (limits 36.3 % to 89.8 %)
Kaplan–Meier reliability at 1000 h	75.0 % (log-log limits 31.5 % to 93.1 %)
Important limitations	small sample; constant hazard unverified (Weibull shape interval 0.61 to 3.63); unknown heterogeneity of operating conditions; uncertain extrapolation beyond the observed ages of 2000 h

definition of failure, the censoring rules, the treatment of maintenance actions and removals, the confidence-interval method and the model diagnostics. It should not present a fitted failure rate without explaining how it was obtained, because every number in the summary depends on choices that the reader must be able to check.

All the methods of this chapter treat the parameters as fixed but unknown constants. The next chapter, Chapter 8, treats them as uncertain quantities with probability distributions updated by Bayes' rule, $p(\boldsymbol{\theta} \mid D) \propto p(D \mid \boldsymbol{\theta})\, p(\boldsymbol{\theta})$, in which the censored likelihood of this chapter is the factor $p(D \mid \boldsymbol{\theta})$; that approach also answers the question of how to estimate reliability when failures are extremely sparse or none has been observed at all.

7.7 What to retain

Definition 7.1 (Estimation from censored reliability data)**.** Reliability is estimated from records of failed and still-operating units by treating each surviving unit as the evidence $T_i > y_i$: nonparametrically through the risk sets of the Kaplan–Meier estimator, and parametrically through the likelihood $\prod_i f(y_i)^{\delta_i} R(y_i)^{1-\delta_i}$, which for the exponential model gives $\hat{\lambda} = r/T_{\mathrm{tot}}$; every such estimate is reported with its statistical uncertainty and with the checks that support its model.

Table 7.11 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 7.11: The terms of reliability data analysis and where each is defined. Units assume time in hours.

Term	Expression	Units	Meaning	Defined in
Right-censored record	$T_i > y_i$	h	lifetime bounded from below	7.2.1
Observed record	$y_i = \min(T_i, C_i)$, δ_i	h, –	time and failure indicator	7.2.1
Interval, left censoring	$a < T \le b$; $T \le c$	h	failure bracketed by inspections	7.2.3
Left truncation	entry at age a; factor $/R(a)$	h	delayed entry, not censoring	7.2.3
Risk set	n_j	–	units at risk just before t_j	7.3.2
Kaplan–Meier estimate	$\hat{R}(t) = \prod_{t_j \le t}(1 - d_j/n_j)$	–	nonparametric reliability	7.3.2
Noninformative censoring	censoring independent of T	–	censored units like those at risk	7.3.5
Restricted mean survival	$\int_0^\tau R(t)\,\mathrm{d}t$	h	mean life capped at τ	7.3.6
Likelihood	$L(\theta) = \prod_i f(y_i)^{\delta_i} R(y_i)^{1-\delta_i}$	–	probability of the records given θ	7.4.2
Total time on test	$T_{\mathrm{tot}} = \sum_i y_i$	h	exposure of failed and censored units	7.4.3
Exponential MLE	$\hat{\lambda} = r/T_{\mathrm{tot}}$	h^{-1}	failures per unit exposure	7.4.4
Fisher information	$\mathcal{I}(\lambda) = \mathbb{E}[r]/\lambda^2$	h^2	curvature of ℓ; inverse variance	7.5.2
Standard error	$\hat{\lambda}/\sqrt{r}$	h^{-1}	relative error $1/\sqrt{r}$	7.5.2
Chi-square limits	$\chi^2_{2r,\alpha/2}/2T_{\mathrm{tot}}$, $\chi^2_{2r+2,1-\alpha/2}/2T_{\mathrm{tot}}$	h^{-1}	small-sample interval for λ	7.5.3
Reliability limits	$\mathrm{e}^{-\lambda_U t}$, $\mathrm{e}^{-\lambda_L t}$	–	interval for $R(t)$	7.5.4
Greenwood variance	$\hat{R}^2 \sum d_j/[n_j(n_j - d_j)]$	–	variance of the Kaplan–Meier estimate	7.5.5
Profile likelihood	$\ell_p(\beta) = \max_\eta \ell(\beta, \eta)$	–	interval for one of two parameters	7.5.6
Model uncertainty	choice of distribution, pooling	–	not covered by confidence intervals	7.6.4

Further reading. Rausand, Barros and Høyland, *System Reliability Theory*, third edition: reliability data, censoring and lifetime estimation. Modarres, Kaminskiy and Krivtsov, *Reliability Engineering and Risk Analysis*, third edition: statistical inference and reliability data analysis. O'Connor and Kleyner, *Practical Reliability Engineering*, fifth edition: field reliability data and practical estimation. The *NIST/SEMATECH e-Handbook of Statistical Methods*: estimating life distribution parameters from censored data, and maximum likelihood estimation.

7.8 Exercises

Exercise 7.1. (a) A pump has operated 2000 h without failure when its observation ends. Should the record be treated as a failure at 2000 h, as a right-censored lifetime at 2000 h, or as an observation without useful information? (Right-censored: it establishes that the lifetime exceeds 2000 h, but not the eventual failure time.) (b) In Kaplan–Meier estimation, what happens at a censoring time at which no failure occurs? (The estimated survival is unchanged, but the

unit leaves all later risk sets.) (c) Under an exponential model with independent right-censored lifetimes, what is the maximum likelihood estimate of the failure rate? (Observed failures divided by the total operating exposure, r/T_{tot}, where the exposure includes the censored units.) (d) Two studies have the same estimated failure rate, but one observed 100 times more failures. What is generally true of their precision? (The larger study is generally more precise; under a correct model its relative standard error is ten times smaller.) (e) What does a narrow confidence interval under an exponential model establish? (Limited sampling uncertainty under the model's assumptions; it does not show that the exponential model is physically correct.)

Exercise 7.2. Six fans of a cooling plant have the illustrative records 300 h (failed), 500 h (censored), 800 h (failed), 1000 h (censored), 1400 h (failed) and 1500 h (censored). (a) Compute the Kaplan–Meier estimate at each failure time and give $\hat{R}(1000)$. (0.8333, 0.6250, 0.3125; $\hat{R}(1000) = 0.625$.) (b) Compute the restricted mean survival time over 1500 h. ($300 + 500 \times 0.8333 + 600 \times 0.625 + 100 \times 0.3125 = 1122.9$ h.) (c) Compute the Greenwood standard error at 800 h. ($0.625 \times \sqrt{1/30 + 1/12} = 0.213$.) (d) Fit the exponential model: give T_{tot}, $\hat{\lambda}$, $\widehat{\mathrm{MTTF}}$ and $\hat{R}(1000)$, and compare the MTTF with the average of the failure times. (5500 h, $5.455 \times 10^{-4}\,\mathrm{h}^{-1}$, 1833 h, 0.580; the average of the failure times, 833 h, ignores the 3000 h of censored exposure.)

Exercise 7.3. A time-terminated test with immediate replacement accumulates $T_{\mathrm{tot}} = 50\,000$ unit-hours and $r = 10$ failures. (a) Compute $\hat{\lambda}$ and its standard error. ($2.0 \times 10^{-4}\,\mathrm{h}^{-1}$; $6.32 \times 10^{-5}\,\mathrm{h}^{-1}$.) (b) With $\chi^2_{20,0.025} = 9.591$ and $\chi^2_{22,0.975} = 36.781$, compute the 95 % chi-square limits on λ and on $R(1000)$. (0.959×10^{-4} to $3.678 \times 10^{-4}\,\mathrm{h}^{-1}$; $R(1000)$ from 0.692 to 0.909, point estimate 0.819.) (c) Compare the ratio λ_U/λ_L with that of the eight-pump dataset and explain the difference. (3.84 against 9.40: ten failures carry more information than four.) (d) If the same test had produced no failure, what one-sided 95 % upper bound on λ and lower bound on $R(1000)$ would equation (7.11) give? ($5.99 \times 10^{-5}\,\mathrm{h}^{-1}$; 0.942.)

Exercise 7.4. (a) Starting from equation (7.15), derive equation (7.16) and show that for $\beta = 1$ it gives the exponential estimate $\hat{\eta} = 1/\hat{\lambda} = T_{\mathrm{tot}}/r$. (For the pumps, $10\,100/4 = 2525$ h.) (b) With the fitted Weibull ($\hat{\beta} = 1.737$, $\hat{\eta} = 2001$ h) and the fitted exponential ($\hat{\lambda} = 3.9604 \times 10^{-4}\,\mathrm{h}^{-1}$), compute the reliability of a new pump at 1000, 2000 and 5000 h under each model. (Weibull 0.741, 0.368, 0.0074; exponential 0.673, 0.453, 0.138.) (c) From the maximized log-likelihoods -34.698 and -35.336, compute the likelihood-ratio statistic and state whether the data reject a constant hazard at the 5 % level. ($1.276 < 3.841$: they do not.) (d) Explain why a narrow confidence interval for $R(5000)$ under the exponential model would not settle the reliability at 5000 h. (The interval is conditional on the model; two models that fit the observed 2000 h equally well differ by a factor of nineteen at 5000 h.)

Chapter 8

Bayesian Reliability Analysis

8.1 Bayesian inference of a failure rate

Chapter 7 estimated the constant failure rate of a fleet of eight cooling-water pumps from their operating records by maximum likelihood. The estimate used every pump, including the four that had not failed when observation ended, and it came with a confidence interval. It did not answer several questions that recur in practice: how the uncertainty in the unknown failure rate should be represented, how earlier evidence on similar pumps should enter, what can be said when only one or two failures, or none, have been observed, and how the uncertainty in the rate should affect a prediction of future failures. The question of this chapter is

how failure probabilities are estimated when observations are sparse, how previous engineering knowledge is combined with new evidence, and how a reliability estimate is updated as further evidence arrives.

The chapter answers it in five steps. It sets the Bayesian treatment of an unknown parameter against the frequentist one used so far, and returns to the eight-pump data. It derives the gamma–exponential conjugate update, in which a gamma prior on the failure rate and a censored exponential likelihood give a gamma posterior in closed form, and works it for the pumps. It summarizes the posterior by credible intervals and by the posterior probabilities of engineering hypotheses. It derives the posterior predictive reliability of a new item, shows that it differs from the reliability evaluated at an estimated rate, and applies it to a test with zero observed failures. Finally, it treats sequential updating, populations of related equipment, and the limits that the model itself places on any Bayesian result.

Jargon: frequentist and Bayesian inference, prior, likelihood, posterior, marginal likelihood, conjugate prior, credible interval, posterior predictive distribution, population variability distribution

In *frequentist inference* an unknown parameter, such as the constant failure rate λ, is a fixed number; it is estimated from data, and the uncertainty of the estimator is described by its sampling distribution over hypothetical repetitions of the observation, for example by a *confidence interval*. In *Bayesian inference* the unknown parameter is given a probability distribution that expresses the state of knowledge about it. The *prior distribution* $p(\lambda)$ expresses that knowledge before the current evidence D; the *likelihood* $p(D \mid \lambda)$ is the probability (or density) of the observed evidence as a function of the parameter; the *posterior distribution* $p(\lambda \mid D)$ is the state of knowledge after the evidence, obtained by Bayes' theorem. The *marginal likelihood*, also called the *model evidence*, is the probability of the evidence averaged over the prior, the normalizing constant of Bayes' theorem. A prior is *conjugate* to a likelihood when the posterior belongs to the same family as the prior. A *credible interval* is an interval that contains the parameter with a stated posterior probability. The *posterior predictive distribution* is the distribution of a future observation, such as the lifetime of a new pump, averaged over the posterior of the parameter. A *population variability distribution* describes how a parameter varies among related groups of equipment, and is the first level of a *hierarchical* Bayesian model.

Key idea

Bayesian reliability analysis represents an unknown failure rate by a probability distribution that expresses the state of knowledge, not a change of the equipment. With a gamma prior $\mathrm{Gamma}(a, b)$ and censored exponential data containing r failures in a total exposure T_{tot}, the posterior is $\mathrm{Gamma}(a + r, b + T_{\mathrm{tot}})$: failures update the shape, exposure updates the rate, and predictions are obtained by averaging over that posterior.

This section recalls the point estimate from which the chapter starts, states the difference between the frequentist and the Bayesian treatment of an unknown parameter, writes Bayes' theorem for a failure rate, and restates the eight-pump data and their likelihood.

8.1.1 Questions left by a point estimate

The eight pumps of Chapter 7 accumulated a total exposure of $T_{\mathrm{tot}} = 10\,100$ pump-hours, in which $r = 4$ failures were observed. Under the exponential lifetime model, with the constant failure rate λ, the maximum likelihood estimate (MLE) is the ratio of failures to exposure,

$$\hat{\lambda}_{\mathrm{MLE}} = \frac{r}{T_{\mathrm{tot}}} = \frac{4}{10\,100\,\mathrm{h}} = 3.9604 \times 10^{-4}\,\mathrm{h}^{-1}, \tag{8.1}$$

or 0.396 failures per 1000 pump-hours, with an estimated mean time to failure of $1/\hat{\lambda}_{\mathrm{MLE}} = 2525\,\mathrm{h}$. The estimate is useful, and it leaves five questions open. How should the uncertainty in the unknown failure rate itself be represented? If similar pumps have been studied before, how should that evidence enter? What can be said if only one or two failures have been observed? What if a new fleet has operated for thousands of hours with no failure at all, so that the MLE is zero? And how should the uncertainty in the rate affect the predicted reliability of the next pump? Bayesian reliability analysis answers all five within one probabilistic model.

8.1.2 Frequentist and Bayesian inference

In the frequentist approach of Chapter 7, the failure rate λ of the pump population is a fixed but unknown number. It is estimated from the data, and the uncertainty is a property of

the estimator: if the observation were repeated many times under the same conditions, the estimates would scatter, and a 95 % confidence procedure produces intervals that contain the fixed true value in 95 % of such repetitions.

In the Bayesian approach, the unknown λ is assigned a probability distribution that represents uncertainty about its value. Before the current data are seen, that distribution is the prior $p(\lambda)$; after the data D are observed, it is the posterior $p(\lambda \mid D)$. The probability is a statement about knowledge of the parameter, not about repetitions of the experiment. In the vocabulary of risk analysis, the lifetime of a pump with a known rate λ carries *aleatory* uncertainty, the variability inherent in the failure process, and the distribution over λ carries *epistemic* uncertainty, the lack of knowledge about the parameter. The distinction between the physical system, the model of it and the evidence is explicit here: the pumps are the system, the exponential lifetime with rate λ is the model, and the operating records are the evidence that updates the distribution over the model's parameter.

8.1.3 Bayes' theorem for a failure rate

For a continuous parameter, Bayes' theorem follows from writing the joint density of the parameter and the evidence in two ways, $p(D, \lambda) = p(D \mid \lambda)\, p(\lambda) = p(\lambda \mid D)\, p(D)$, and solving for the posterior:

$$p(\lambda \mid D) = \frac{p(D \mid \lambda)\, p(\lambda)}{p(D)} = \frac{p(D \mid \lambda)\, p(\lambda)}{\int_0^\infty p(D \mid u)\, p(u)\, \mathrm{d}u}, \tag{8.2}$$

where u is a variable of integration over the possible values of the rate. The denominator $p(D)$ is the marginal likelihood, or model evidence: the probability of the observed evidence averaged over the prior. It does not depend on λ, and its only role in equation (8.2) is to make the posterior integrate to one. For that reason the theorem is often written as a proportionality,

$$p(\lambda \mid D) \propto p(D \mid \lambda)\, p(\lambda), \tag{8.3}$$

posterior proportional to likelihood times prior, with the constant recovered at the end by normalization. Table 8.1 lists the three elements: the column on the left names each, the middle column gives its notation, and the right column states what it represents.

Table 8.1: The three elements of Bayes' theorem for a failure rate λ and evidence D.

Element	Notation	Meaning
Prior	$p(\lambda)$	information or assumptions about λ before the current evidence
Likelihood	$p(D \mid \lambda)$	how the probability of the observed evidence depends on λ
Posterior	$p(\lambda \mid D)$	the updated uncertainty about λ after the evidence

Figure 8.1 draws the update as a flow. The three boxes are the prior, the likelihood and the posterior; the multiplication sign between the first two is equation (8.3), and the arrow to the third is the normalization. The second line of each box gives its form for the pump example worked in Section 8.2; the notes beneath name the role of each element. Previous knowledge and current evidence combine into an updated statement of uncertainty about the parameter.

Bayesian inference does not assume that the physical failure rate changes when data arrive. The pumps behave exactly as before. What changes is the distribution that expresses what is known about their failure behavior, and that distinction is the foundation of everything that follows.

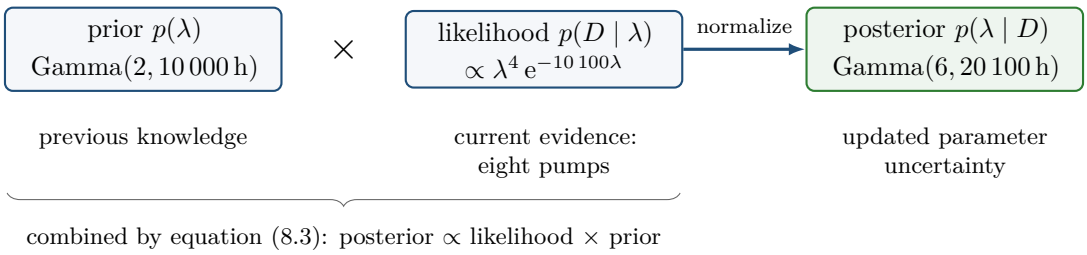


Figure 8.1: Bayes' theorem as an update of knowledge about a failure rate. The prior and the likelihood multiply; normalization by the marginal likelihood gives the posterior. Second lines: the gamma prior, the censored exponential likelihood and the gamma posterior of the eight-pump example (Section 8.2.3).

Principle 8.1. *Bayesian updating changes the state of knowledge about a reliability parameter, not the equipment. A posterior that moves when evidence arrives records new information, not a physical change of the failure process.*

8.1.4 The eight-pump data and the censored exponential likelihood

The chapter uses the same observations as Chapter 7. Table 8.2 lists them: each row is a pump, the second column its observed operating hours from installation, and the third whether the observation ended in a failure or was right-censored, the pump still operating when observation stopped. The data are synthetic and illustrative.

Table 8.2: The eight-pump dataset of Chapter 7 (illustrative). Right-censored: still operating when observation ended.

Pump	Observed operating hours	Status
P1	400	failed
P2	700	failed
P3	900	right-censored
P4	1200	failed
P5	1500	right-censored
P6	1600	failed
P7	1800	right-censored
P8	2000	right-censored

Assume that the first-failure lifetimes are independent and exponential with a common constant rate λ, and that censoring is noninformative: the time at which observation of a pump stops carries no information about its failure time beyond the fact that it had not yet failed. Let t_i be the observed hours of pump i and δ_i its failure indicator, one for a failure and zero for a censored observation. A failure at t_i contributes the density $f(t_i \mid \lambda) = \lambda\, \mathrm{e}^{-\lambda t_i}$ to the likelihood; a pump censored at t_i contributes the probability of having survived that long, $R(t_i \mid \lambda) = \mathrm{e}^{-\lambda t_i}$. The likelihood $L(\lambda)$ of the eight observations is the product,

$$L(\lambda) = p(D \mid \lambda) \propto \prod_{i=1}^{8} \left(\lambda\, \mathrm{e}^{-\lambda t_i}\right)^{\delta_i} \left(\mathrm{e}^{-\lambda t_i}\right)^{1-\delta_i} = \lambda^{r}\, \mathrm{e}^{-\lambda T_{\mathrm{tot}}}, \tag{8.4}$$

with

$$r = \sum_i \delta_i = 4, \qquad T_{\mathrm{tot}} = \sum_i t_i = 400+700+900+1200+1500+1600+1800+2000 = 10\,100\,\mathrm{h}.$$

The proportionality omits factors that do not depend on λ, such as the probabilities of the censoring times themselves, which drop out of equation (8.2) because they appear in the numerator and the denominator alike. Under the exponential model the data reduce to two numbers, the failure count r and the total exposure T_{tot}, which includes the hours of the censored pumps. The frequentist analysis maximized equation (8.4); the Bayesian analysis starts from the same likelihood and combines it with a prior.

8.2 Gamma–exponential conjugacy

A prior for the failure rate must be a distribution on the positive reals. This section introduces the gamma distribution, derives the posterior it produces with the censored exponential likelihood, works the update for the eight pumps, and reads the result as a weighted combination of prior and data.

8.2.1 The gamma prior

The failure rate is positive, $\lambda > 0$. A convenient and flexible prior is the gamma distribution in the shape–rate parameterization, $\lambda \sim \text{Gamma}(a, b)$, with density

$$p(\lambda) = \frac{b^a}{\Gamma(a)}\, \lambda^{a-1}\, \mathrm{e}^{-b\lambda}, \qquad \lambda > 0, \tag{8.5}$$

where $a > 0$ is the shape parameter (dimensionless), $b > 0$ is the rate parameter, and $\Gamma(\cdot)$ is the gamma function, $\Gamma(c) = \int_0^\infty x^{c-1}\mathrm{e}^{-x}\,\mathrm{d}x$, with $\Gamma(c+1) = c\,\Gamma(c)$ and $\Gamma(n) = (n-1)!$ for a positive integer n. Because λ is in h^{-1} and the exponent $b\lambda$ must be dimensionless, the gamma rate b is in hours. The book writes the gamma parameters a and b so that β keeps its meaning as the Weibull shape. One integral, obtained from the definition of Γ by the substitution $x = s\lambda$, gives every moment of the gamma distribution:

$$\int_0^\infty \lambda^{c-1}\, \mathrm{e}^{-s\lambda}\, \mathrm{d}\lambda = \frac{\Gamma(c)}{s^c}, \qquad c > 0,\ s > 0. \tag{8.6}$$

With $c = a$ and $s = b$ it confirms that equation (8.5) integrates to one.

Mean, variance and coefficient of variation. With $c = a + 1$ in equation (8.6),

$$\mathbb{E}[\lambda] = \frac{b^a}{\Gamma(a)} \int_0^\infty \lambda^{a}\, \mathrm{e}^{-b\lambda}\, \mathrm{d}\lambda = \frac{b^a}{\Gamma(a)}\, \frac{\Gamma(a+1)}{b^{a+1}} = \frac{a}{b},$$

and with $c = a + 2$, $\mathbb{E}[\lambda^2] = a(a+1)/b^2$, so that

$$\mathbb{E}[\lambda] = \frac{a}{b}, \qquad \mathrm{Var}(\lambda) = \frac{a(a+1)}{b^2} - \frac{a^2}{b^2} = \frac{a}{b^2}, \qquad \sigma_\lambda = \frac{\sqrt{a}}{b}, \qquad \frac{\sigma_\lambda}{\mathbb{E}[\lambda]} = \frac{1}{\sqrt{a}}, \tag{8.7}$$

where σ_λ is the standard deviation of λ. The ratio a/b sets the mean failure rate; the shape a alone sets the coefficient of variation, the spread relative to the mean. Holding the mean fixed, a larger a (with b scaled in proportion) gives a more concentrated prior.

Example prior. Suppose previous engineering evidence on similar pumps suggests a typical rate of $\lambda_0 = 2 \times 10^{-4}\,\mathrm{h}^{-1}$, which would correspond to a mean time to failure of 5000 h if that rate were known exactly. The illustrative prior $\text{Gamma}(2, 10\,000\,\mathrm{h})$ has

$$\mathbb{E}[\lambda] = \frac{2}{10\,000\,\mathrm{h}} = 2 \times 10^{-4}\,\mathrm{h}^{-1}, \qquad \sigma_\lambda = \frac{\sqrt{2}}{10\,000\,\mathrm{h}} = 1.414 \times 10^{-4}\,\mathrm{h}^{-1},$$

a coefficient of variation of $1/\sqrt{2} = 0.71$: a moderately broad prior relative to its mean. Its central 95 % range, from the quantiles derived in Section 8.3.2, runs from 0.24×10^{-4} to $5.57 \times 10^{-4}\,\mathrm{h}^{-1}$.

The mean of the reciprocal. Because λ is uncertain, the mean of $1/\lambda$ is not the reciprocal of the mean of λ. With $c = a - 1$ in equation (8.6), valid for $a > 1$,

$$\mathbb{E}\left[\frac{1}{\lambda}\right] = \frac{b^a}{\Gamma(a)} \frac{\Gamma(a-1)}{b^{a-1}} = \frac{b}{a-1}. \tag{8.8}$$

For the example prior, the mean of the conditional mean time to failure $1/\lambda$ is $10\,000\,\mathrm{h}/(2-1) = 10\,000\,\mathrm{h}$, twice the 5000 h obtained as the reciprocal of the mean rate. The function $1/\lambda$ is convex, so averaging it over the uncertain rate gives more than evaluating it at the average rate. The same effect reappears in the predictive reliability of Section 8.4.

8.2.2 The posterior distribution

With the prior of equation (8.5) and the likelihood of equation (8.4), equation (8.3) gives

$$p(\lambda \mid D) \propto \lambda^r \, \mathrm{e}^{-\lambda T_{\mathrm{tot}}} \cdot \lambda^{a-1} \, \mathrm{e}^{-b\lambda},$$

where the constant $b^a/\Gamma(a)$ has been absorbed into the proportionality. Collecting the powers of λ and the exponents,

$$p(\lambda \mid D) \propto \lambda^{a+r-1} \, \mathrm{e}^{-(b+T_{\mathrm{tot}})\lambda}.$$

This is the kernel of a gamma density with shape $a + r$ and rate $b + T_{\mathrm{tot}}$, and the normalizing constant follows from equation (8.6). Therefore

$$\lambda \mid D \sim \mathrm{Gamma}(a', b'), \qquad a' = a + r, \qquad b' = b + T_{\mathrm{tot}}. \tag{8.9}$$

This is *gamma–exponential conjugacy*: a gamma prior combined with a censored exponential likelihood produces a gamma posterior. The failure count adds to the shape and the exposure adds to the rate. Because the posterior has a known form, its mean, variance, quantiles and predictive consequences are available analytically, without numerical integration or Monte Carlo simulation.

The same integral gives the marginal likelihood in closed form, which Section 8.6.4 uses for model comparison:

$$p(D) = \int_0^\infty \lambda^r \, \mathrm{e}^{-\lambda T_{\mathrm{tot}}} \, \frac{b^a}{\Gamma(a)} \, \lambda^{a-1} \, \mathrm{e}^{-b\lambda} \, \mathrm{d}\lambda = \frac{b^a}{\Gamma(a)} \, \frac{\Gamma(a+r)}{(b + T_{\mathrm{tot}})^{a+r}}, \tag{8.10}$$

up to the λ-free factors omitted from the likelihood.

8.2.3 Worked update for the eight pumps

With the prior $\mathrm{Gamma}(2, 10\,000\,\mathrm{h})$ and the evidence $r = 4$, $T_{\mathrm{tot}} = 10\,100\,\mathrm{h}$, equation (8.9) gives

$$a' = 2 + 4 = 6, \qquad b' = 10\,000 + 10\,100 = 20\,100\,\mathrm{h}, \qquad \lambda \mid D \sim \mathrm{Gamma}(6, 20\,100\,\mathrm{h}).$$

The posterior mean is

$$\mathbb{E}[\lambda \mid D] = \frac{a'}{b'} = \frac{6}{20\,100\,\mathrm{h}} = 2.9851 \times 10^{-4}\,\mathrm{h}^{-1}.$$

Table 8.3 compares it with the prior mean and the MLE. The posterior mean lies between them: the prior indicates a lower rate than the data alone, and the posterior combines both sources of evidence. Chapter 7 showed that the MLE uses the censored pumps' exposure correctly; the posterior uses the same exposure and adds the prior information.

Figure 8.2 draws the three distributions that enter the update. The horizontal axis is the failure rate in failures per 1000 h, so that 0.2 means $2 \times 10^{-4}\,\mathrm{h}^{-1}$; the vertical axis is the

Table 8.3: Point estimates of the pump failure rate. Prior Gamma$(2, 10\,000\,\mathrm{h})$; data $r = 4$, $T_{\mathrm{tot}} = 10\,100\,\mathrm{h}$; posterior Gamma$(6, 20\,100\,\mathrm{h})$.

Estimate	Expression	Rate (h^{-1})
Prior mean	a/b	0.0002000
Maximum likelihood	r/T_{tot}	0.0003960
Posterior mean	$(a+r)/(b+T_{\mathrm{tot}})$	0.0002985

probability density per unit of that axis. The blue curve is the prior; the dashed gold curve is the likelihood of equation (8.4) rescaled to unit area, which is the Gamma$(5, 10\,100\,\mathrm{h})$ density and the posterior that a flat prior would give; the green curve is the posterior. The dotted vertical lines mark the prior mean, the posterior mean and the MLE. The posterior sits between the prior and the likelihood and is narrower than either: two sources of information about the same parameter concentrate it more than each one alone. Its peak, the posterior mode $(a'-1)/b' = 2.49 \times 10^{-4}\,\mathrm{h}^{-1}$, lies below its mean because the gamma density is skewed to the right.

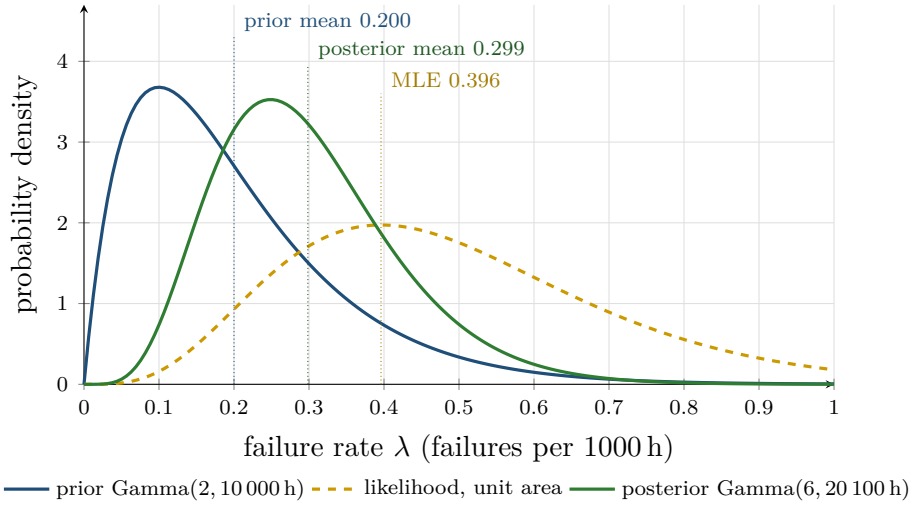


Figure 8.2: Prior, normalized likelihood and posterior of the pump failure rate, in absolute density units with $x = 1000\lambda$ in failures per 1000 h: prior $100\,x\,\mathrm{e}^{-10x}$, likelihood rescaled to unit area $(10.1^5/4!)\,x^4\mathrm{e}^{-10.1x}$, posterior $(20.1^6/5!)\,x^5\mathrm{e}^{-20.1x}$, all from equation (8.5). The posterior is concentrated between the prior and the evidence of four failures in 10 100 pump-hours.

8.2.4 The posterior mean as a weighted combination

The posterior mean $(a+r)/(b+T_{\mathrm{tot}})$ can be rearranged into a weighted average of the prior mean and the MLE. Split the numerator and multiply each part by a factor equal to one:

$$\mathbb{E}[\lambda \mid D] = \frac{a}{b+T_{\mathrm{tot}}} + \frac{r}{b+T_{\mathrm{tot}}} = \underbrace{\frac{b}{b+T_{\mathrm{tot}}}}_{w}\,\frac{a}{b} + \underbrace{\frac{T_{\mathrm{tot}}}{b+T_{\mathrm{tot}}}}_{1-w}\,\frac{r}{T_{\mathrm{tot}}}. \tag{8.11}$$

The weight $w = b/(b+T_{\mathrm{tot}})$ on the prior mean and the weight $1-w$ on the MLE add to one, and they depend only on how the exposure observed compares with the rate parameter of the prior. For the pumps,

$$w = \frac{10\,000}{20\,100} = 0.4975, \qquad 1-w = \frac{10\,100}{20\,100} = 0.5025,$$

and $0.4975 \times 2.000 \times 10^{-4} + 0.5025 \times 3.960 \times 10^{-4} = 2.985 \times 10^{-4}\,\mathrm{h}^{-1}$, the posterior mean of Table 8.3. The prior and the current data receive almost equal weight in this example. As exposure accumulates, $1 - w \to 1$ and the posterior mean approaches the MLE; with little exposure, $w \to 1$ and it stays near the prior mean. The weighting is a consequence of the conjugate algebra. It does not establish that the prior is equivalent to a genuine historical experiment.

8.2.5 Prior information as pseudo-exposure

Equation (8.9) adds the observed failures to a and the observed exposure to b. Algebraically, therefore, the prior acts as if it contributed a failures in b hours of exposure: for the example prior, two failure-count units and 10 000 hour units. The reading is useful for judging the strength of a prior: a prior with $b = 10\,000\,\mathrm{h}$ carries about as much weight as 10 000 pump-hours of new data, which is why it and the 10 100 pump-hours of the fleet receive almost equal weight.

The interpretation is exact only as a feature of the algebra. A Gamma$(2, 10\,000\,\mathrm{h})$ prior is not two observed failures unless it was explicitly constructed from such data, and a proper gamma prior with a noninteger shape, such as $a = 0.5$, corresponds to no count of failures at all.

Common mistake

Reading the prior's pseudo-counts as real evidence. The prior Gamma$(2, 10\,000\,\mathrm{h})$ behaves in the update like two failures in 10 000 h, but it is an expression of judgment or of other data, and its weight in the posterior is a modeling choice that must be justified and disclosed, not a record of observed failures.

8.3 Posterior uncertainty and credible intervals

A posterior distribution is more than a point estimate. This section summarizes the spread of the pump posterior, derives its equal-tailed credible interval, contrasts that interval with a confidence interval, and computes the posterior probability that the failure rate exceeds a specified value.

8.3.1 The posterior as a distribution

The posterior Gamma$(6, 20\,100\,\mathrm{h})$ describes a range of plausible values of the failure rate, not one value. By equation (8.7) its standard deviation is $\sqrt{6}/20\,100\,\mathrm{h} = 1.219 \times 10^{-4}\,\mathrm{h}^{-1}$ and its coefficient of variation is $1/\sqrt{6} = 0.41$, down from 0.71 for the prior: the six effective failures of the posterior, two from the prior and four observed, pin the rate down relative to its mean more tightly than the two of the prior. Three point summaries of the posterior are in use: the mean $a'/b' = 2.985 \times 10^{-4}\,\mathrm{h}^{-1}$, which minimizes the expected squared error and is the usual Bayes estimate; the median $2.821 \times 10^{-4}\,\mathrm{h}^{-1}$; and the mode $(a' - 1)/b' = 2.488 \times 10^{-4}\,\mathrm{h}^{-1}$. They differ because the distribution is skewed, and none of them conveys the spread. The posterior carries both the estimate and its uncertainty, and the intervals and probabilities below are read from it directly.

8.3.2 Equal-tailed credible intervals

A 95 % equal-tailed credible interval is the pair of values λ_L and λ_U that leave 2.5 % of posterior probability in each tail:

$$\Pr(\lambda < \lambda_L \mid D) = 0.025, \qquad \Pr(\lambda > \lambda_U \mid D) = 0.025. \tag{8.12}$$

They are the 0.025 and 0.975 quantiles of the gamma posterior. For an integer shape they are obtained from chi-square quantiles: if $\lambda \sim \mathrm{Gamma}(a', b')$, then $2b'\lambda$ has the gamma distribution

with shape a' and rate $1/2$, which is the chi-square distribution with $2a'$ degrees of freedom. Writing $\chi^2_{\nu,p}$ for the p quantile of the chi-square distribution with ν degrees of freedom,

$$\lambda_L = \frac{\chi^2_{2a',0.025}}{2b'}, \qquad \lambda_U = \frac{\chi^2_{2a',0.975}}{2b'}.$$

For the pumps, $2a' = 12$, $\chi^2_{12,0.025} = 4.4038$, $\chi^2_{12,0.975} = 23.3367$ and $2b' = 40\,200$ h, so

$$\lambda_L = \frac{4.4038}{40\,200\,\text{h}} = 1.0955 \times 10^{-4}\,\text{h}^{-1}, \qquad \lambda_U = \frac{23.3367}{40\,200\,\text{h}} = 5.8051 \times 10^{-4}\,\text{h}^{-1}.$$

Under the prior and the exponential likelihood model, there is a 95 % posterior probability that the pump failure rate lies between 0.0001096 and 0.0005805 h^{-1}. Equal tails are a convention; the shortest interval with 95 % probability, the highest posterior density interval, is another, and for a skewed posterior it lies slightly to the left of the equal-tailed one.

Figure 8.3(a) shows the interval. The horizontal axis is again the rate in failures per 1000 h and the vertical axis the posterior density. The shaded band under the curve between the two dashed lines has area 0.95; each white tail outside it has area 0.025. The interval is not symmetric about the mean, 2.985, which lies 1.89×10^{-4} above the lower limit and 2.82×10^{-4} below the upper one: the right tail of the gamma posterior is longer than the left.

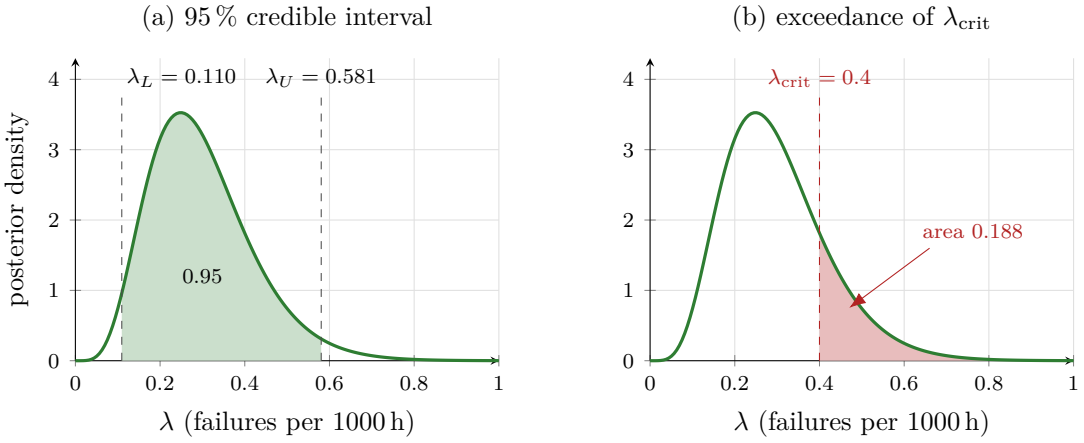


Figure 8.3: Reading the posterior Gamma(6, 20 100 h) of the pump failure rate, density $(20.1^6/5!)\,x^5 e^{-20.1x}$ with $x = 1000\lambda$. (a) Equal-tailed 95 % credible interval, 1.0955 to $5.8051 \times 10^{-4}\,\text{h}^{-1}$ (equation (8.12)); shaded area 0.95, each unshaded tail 0.025. (b) Posterior probability that the rate exceeds $4 \times 10^{-4}\,\text{h}^{-1}$, 0.188 (equation (8.14)).

8.3.3 Credible and confidence intervals

The two kinds of interval answer different questions. A 95 % frequentist confidence procedure is designed so that, over repeated samples under the model, its intervals cover the fixed true parameter in 95 % of the repetitions; a single computed interval either contains the parameter or does not. A 95 % Bayesian credible interval assigns 95 % posterior probability to the parameter lying inside the reported interval, conditional on the data, the prior and the model. The second is the statement an engineer usually wants to make, and it is available only because a prior has been stated.

The numbers differ as well. Chapter 7 computed for the same data the chi-square reference interval $\chi^2_{2r,0.025}/(2T_{\text{tot}})$ to $\chi^2_{2(r+1),0.975}/(2T_{\text{tot}})$, which gives 1.079×10^{-4} to $10.14 \times 10^{-4}\,\text{h}^{-1}$. Its limits are themselves quantiles of gamma distributions with rate T_{tot} and shapes r and $r + 1$, that is, of posteriors from nearly flat priors. The credible interval, 1.096×10^{-4} to

$5.805 \times 10^{-4}\,\mathrm{h}^{-1}$, has almost the same lower limit and an upper limit lower by 43 %, because the prior adds the equivalent of 10 000 pump-hours with a mean rate below the MLE. The difference is the prior's information, and it is only as good as that information. Neither interval establishes that the exponential model is physically correct; both are conditional on it.

Common mistake

Reading a confidence interval as a probability statement about the parameter, or a credible interval as a property of repeated sampling. "There is a 95 % probability that λ lies in this interval" is a Bayesian statement that requires a prior; a confidence interval supports only the statement that the procedure covers the true value in 95 % of repetitions.

8.3.4 Posterior probabilities of engineering hypotheses

A posterior answers directly questions such as the probability that the pump failure rate exceeds $4 \times 10^{-4}\,\mathrm{h}^{-1}$. For a threshold λ_{crit} and a gamma posterior,

$$\Pr(\lambda > \lambda_{\mathrm{crit}} \mid D) = 1 - G(\lambda_{\mathrm{crit}}; a', b'), \tag{8.13}$$

where $G(x; a', b')$ is the cumulative distribution function of $\mathrm{Gamma}(a', b')$. For an integer shape the tail has a closed form. Substituting $u = b'\lambda$ in the tail integral,

$$\Pr(\lambda > x \mid D) = \int_{b'x}^{\infty} \frac{u^{a'-1}\,\mathrm{e}^{-u}}{(a'-1)!}\,\mathrm{d}u,$$

and integrating by parts once lowers the power of u by one and produces the boundary term $\mathrm{e}^{-b'x}(b'x)^{a'-1}/(a'-1)!$; repeating down to the power zero, whose integral is $\mathrm{e}^{-b'x}$, gives

$$\Pr(\lambda > x \mid D) = \sum_{k=0}^{a'-1} \mathrm{e}^{-b'x}\,\frac{(b'x)^k}{k!}. \tag{8.14}$$

The right side is the probability that a Poisson variable with mean $b'x$ is smaller than a'. For the pumps and $\lambda_{\mathrm{crit}} = 4 \times 10^{-4}\,\mathrm{h}^{-1}$, $b'x = 20\,100 \times 0.0004 = 8.04$, and the six terms are 0.00032, 0.00259, 0.01042, 0.02792, 0.05612 and 0.09023, which add to

$$\Pr(\lambda > 4 \times 10^{-4}\,\mathrm{h}^{-1} \mid D) = 0.1876.$$

Table 8.4 gives the same probability for five thresholds, under the prior and under the posterior. The first column is the threshold, the second the argument $b'x$ of equation (8.14), and the last two the exceedance probabilities. At every threshold the evidence raises the exceedance probability, because the observed rate of $3.96 \times 10^{-4}\,\mathrm{h}^{-1}$ lies above the prior mean and the four failures make low rates less plausible; at $4 \times 10^{-4}\,\mathrm{h}^{-1}$, close to the observed rate, the probability doubles from 0.092 to 0.188. Figure 8.3(b) draws the $4 \times 10^{-4}\,\mathrm{h}^{-1}$ case: the red area under the posterior to the right of the threshold line is 0.188.

Probabilities of this kind are useful when a specification or a decision is tied to a threshold. A procurement requirement may state that a component's failure rate under specified conditions must be below a given value; instead of reporting only a posterior mean, the analysis can report the posterior probability that the requirement is met, here $1 - 0.188 = 0.812$ for a limit of $4 \times 10^{-4}\,\mathrm{h}^{-1}$. A contractual demonstration test, however, may prescribe a frequentist acceptance or confidence procedure, and a Bayesian probability does not automatically replace that requirement.

Table 8.4: Probability that the pump failure rate exceeds a threshold, before and after the evidence. Prior Gamma(2, 10 000 h); posterior Gamma(6, 20 100 h); equation (8.14).

Threshold λ_{crit} ($10^{-4}\,\text{h}^{-1}$)	$b'\lambda_{\text{crit}}$	$\Pr(\lambda > \lambda_{\text{crit}})$, prior	$\Pr(\lambda > \lambda_{\text{crit}} \mid D)$
1	2.01	0.7358	0.9831
2	4.02	0.4060	0.7820
3	6.03	0.1991	0.4409
4	8.04	0.0916	0.1876
5	10.05	0.0404	0.0652

8.4 Posterior predictive reliability

The quantity a decision usually needs is not the failure rate but the probability that a particular pump survives a particular mission. This section derives that probability when the rate is uncertain, works it for a 1000 h mission, and shows that the resulting lifetime distribution is not exponential.

8.4.1 Reliability with an uncertain failure rate

Suppose a new pump will operate for $t = 1000\,\text{h}$. If the rate λ were known, its reliability would be the exponential $R(t \mid \lambda) = \mathrm{e}^{-\lambda t}$. The rate is not known; its posterior $p(\lambda \mid D)$ is. By the law of total probability, the probability that the new pump survives to t is the conditional reliability averaged over the posterior:

$$R_{\text{pred}}(t \mid D) = \int_0^\infty R(t \mid \lambda)\, p(\lambda \mid D)\, \mathrm{d}\lambda = \int_0^\infty \mathrm{e}^{-\lambda t}\, p(\lambda \mid D)\, \mathrm{d}\lambda. \tag{8.15}$$

This is the *posterior predictive reliability*. It assumes that the new pump belongs to the same population and operates under the same conditions as the pumps observed, so that its lifetime given λ is exponential with that same λ and is independent of the past data given λ. It accounts for both kinds of uncertainty: the aleatory randomness of the failure time for a given rate, and the epistemic uncertainty about the rate.

8.4.2 The predictive reliability under a gamma posterior

With the gamma posterior of equation (8.9), equation (8.15) becomes

$$R_{\text{pred}}(t \mid D) = \frac{b'^{\,a'}}{\Gamma(a')} \int_0^\infty \lambda^{a'-1}\, \mathrm{e}^{-b'\lambda}\, \mathrm{e}^{-\lambda t}\, \mathrm{d}\lambda = \frac{b'^{\,a'}}{\Gamma(a')} \int_0^\infty \lambda^{a'-1}\, \mathrm{e}^{-(b'+t)\lambda}\, \mathrm{d}\lambda.$$

The integral is equation (8.6) with $c = a'$ and $s = b' + t$, equal to $\Gamma(a')/(b'+t)^{a'}$, and the gamma functions cancel:

$$R_{\text{pred}}(t \mid D) = \left(\frac{b'}{b'+t}\right)^{a'} = \left(1 + \frac{t}{b'}\right)^{-a'}. \tag{8.16}$$

The predictive survival probability is available analytically, without simulation. Equation (8.16) is the reliability function of the Lomax distribution, also called the Pareto distribution of the second kind, with shape a' and scale b'; its density is $f_{\text{pred}}(t \mid D) = -\mathrm{d}R_{\text{pred}}/\mathrm{d}t = a'\, b'^{\,a'}/(b'+t)^{a'+1}$.

8.4.3 Worked example: reliability over a 1000 h mission

With $a' = 6$ and $b' = 20\,100$ h,

$$R_{\text{pred}}(1000\,\text{h} \mid D) = \left(\frac{20\,100}{20\,100 + 1000}\right)^6 = \left(\frac{20\,100}{21\,100}\right)^6 = 0.952607^6 = 0.7473.$$

The posterior predictive reliability over the mission is 74.73 %. Substituting the posterior mean rate into the exponential reliability, the *plug-in* estimate, gives instead

$$R_{\text{plug}}(1000\,\text{h}) = \mathrm{e}^{-\mathbb{E}[\lambda|D]\times 1000\,\text{h}} = \mathrm{e}^{-0.298507} = 0.7419,$$

or 74.19 %. For comparison, the plug-in value at the MLE of Chapter 7 is $\mathrm{e}^{-0.39604} = 0.6730$; that difference comes from the prior, while the difference between 0.7473 and 0.7419 comes from averaging.

The two Bayesian numbers differ because averaging and exponentiation do not commute:

$$\mathbb{E}\left[\mathrm{e}^{-\lambda t} \mid D\right] \neq \mathrm{e}^{-\mathbb{E}[\lambda|D]\,t}$$

in general. For $t > 0$ the function $\lambda \mapsto \mathrm{e}^{-\lambda t}$ is convex, its second derivative $t^2\mathrm{e}^{-\lambda t}$ being positive, and Jensen's inequality, which states that the expectation of a convex function is at least the function of the expectation, gives

$$R_{\text{pred}}(t \mid D) = \mathbb{E}\left[\mathrm{e}^{-\lambda t} \mid D\right] \geq \mathrm{e}^{-\mathbb{E}[\lambda|D]\,t}. \tag{8.17}$$

The predictive reliability is therefore at least the plug-in reliability at the posterior mean rate. This does not mean that parameter uncertainty makes equipment physically more reliable. It is a mathematical consequence of averaging a convex survival function over uncertain rates: the low rates the posterior admits raise the survival probability by more than the high rates lower it.

Common mistake

Substituting a point estimate of the rate into $\mathrm{e}^{-\lambda t}$ and reporting the result as the reliability of a new item. The plug-in value ignores the uncertainty in λ; the predictive reliability of equation (8.15) includes it. The two are close for a short mission and a well-determined rate, and far apart in the tail, as Figure 8.4 shows.

8.4.4 The predictive distribution is not exponential

The predictive reliability of equation (8.16) has the cumulative hazard

$$\Lambda_{\text{pred}}(t) = -\ln R_{\text{pred}}(t \mid D) = a' \ln\left(1 + \frac{t}{b'}\right),$$

and differentiating gives the predictive hazard

$$\lambda_{\text{pred}}(t) = \frac{\mathrm{d}\Lambda_{\text{pred}}}{\mathrm{d}t} = \frac{a'}{b' + t}, \tag{8.18}$$

which decreases with age, although every conditional model in the analysis has a constant hazard. The reason is visible in the form of equation (8.18). Writing the hazard as density over reliability with equation (8.15),

$$\lambda_{\text{pred}}(t) = \frac{\int_0^\infty \lambda\,\mathrm{e}^{-\lambda t}\,p(\lambda \mid D)\,\mathrm{d}\lambda}{\int_0^\infty \mathrm{e}^{-\lambda t}\,p(\lambda \mid D)\,\mathrm{d}\lambda} = \mathbb{E}[\lambda \mid D,\ T > t],$$

the posterior mean of the rate after the further observation that the new pump has survived to t. That observation is a censored lifetime of length t, so by equation (8.9) it updates the posterior to Gamma$(a', b' + t)$, whose mean is $a'/(b' + t)$. Survival to a greater age makes lower rates more plausible, and the predictive hazard falls. This is the population-mixture, or selection, effect of the mixed population in Chapter 5: a decreasing predictive hazard does not imply that a pump improves with age.

The mean of the predictive lifetime, from $\int_0^\infty R_{\text{pred}}(t)\,\mathrm{d}t$, is $b'/(a' - 1) = 20\,100/5 = 4020\,\text{h}$ for $a' > 1$. It equals $\mathbb{E}[1/\lambda \mid D]$ of equation (8.8), as it must, since the mean lifetime given λ is $1/\lambda$; the plug-in exponential has mean $1/\mathbb{E}[\lambda \mid D] = 3350\,\text{h}$. The predictive tail is a power law, $R_{\text{pred}} \sim (b'/t)^{a'}$ for $t \gg b'$, and decays far more slowly than any exponential: at 10 000 h the predictive reliability is 0.0887 against 0.0505 for the plug-in, and at 20 000 h it is 0.0159 against 0.0026.

Figure 8.4 compares the two. In panel (a) the horizontal axis is the mission time in hours and the vertical axis the reliability; the green curve is the predictive reliability of equation (8.16) and the dashed blue curve the plug-in exponential at the posterior mean rate; the dots are the 1000 h values 0.7473 and 0.7419. The curves agree closely over the first 2000 h, where they differ by less than 0.016, and separate in the tail. In panel (b) the vertical axis is the hazard: the plug-in hazard is constant at $2.985 \times 10^{-4}\,\text{h}^{-1}$, and the predictive hazard starts at the same value and falls to $1.99 \times 10^{-4}\,\text{h}^{-1}$ at 10 000 h.

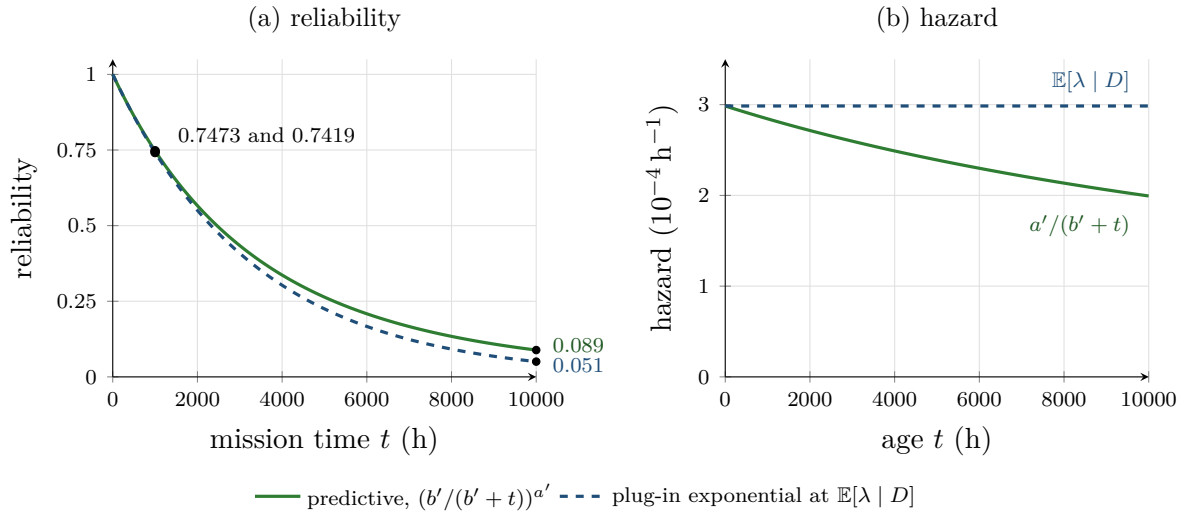


Figure 8.4: Posterior predictive against plug-in reliability for the pump posterior Gamma$(6, 20\,100\,\text{h})$. (a) Reliability: predictive $(20\,100/(20\,100 + t))^6$ (equation (8.16)) and plug-in $\mathrm{e}^{-2.985\times10^{-4}t}$; dots at 1000 h; tail values at 10 000 h. (b) Hazard: predictive $6/(20\,100 + t)$ (equation (8.18)), decreasing, and the constant plug-in $2.985 \times 10^{-4}\,\text{h}^{-1}$.

Principle 8.2. *A prediction for a new item averages the conditional model over the posterior of its parameters. The result is not the conditional model at an estimated parameter, and for an exponential lifetime with a gamma posterior it is a Lomax reliability with a decreasing hazard.*

8.5 Reliability estimation with zero observed failures

New equipment often accumulates thousands of hours before its first failure. This section shows why maximum likelihood gives no usable answer in that case, derives the Bayesian one,

computes the predictive reliability, shows how strongly the result depends on the prior, and collects the dependence in two worked tables.

8.5.1 The maximum likelihood estimate at the boundary

Suppose ten new electrical protection devices are installed, and each operates for 2000 h without failure. The evidence is

$$r = 0, \qquad T_{\text{tot}} = 10 \times 2000\,\text{h} = 20\,000\,\text{h}$$

of device exposure. The likelihood of equation (8.4) is $\mathrm{e}^{-\lambda T_{\text{tot}}}$, which decreases monotonically in λ and is largest at the boundary $\lambda = 0$; the MLE is $\hat{\lambda}_{\text{MLE}} = r/T_{\text{tot}} = 0$.

That is the maximum likelihood estimate, but it would be a serious error to read it as proof that the devices cannot fail. Zero observed failures do not imply zero failure probability; they mean that the observations have not yet revealed a failure. The likelihood still carries information: it is the survival term $\mathrm{e}^{-\lambda T_{\text{tot}}}$, which penalizes high rates more the longer the failure-free exposure. What it does not provide is a failure event that locates a positive rate.

Common mistake

Reporting $\hat{\lambda} = 0$ after a failure-free test. A rate of zero implies certain survival over any mission. The evidence supports only a statement that high rates have become implausible, which is expressed by an upper bound or by a posterior distribution.

8.5.2 The Bayesian answer

With the illustrative prior $\text{Gamma}(2, 10\,000\,\text{h})$, the update of equation (8.9) is

$$a' = 2 + 0 = 2, \qquad b' = 10\,000 + 20\,000 = 30\,000\,\text{h}, \qquad \lambda \mid D \sim \text{Gamma}(2, 30\,000\,\text{h}),$$

with posterior mean

$$\mathbb{E}[\lambda \mid D] = \frac{2}{30\,000\,\text{h}} = 6.6667 \times 10^{-5}\,\text{h}^{-1},$$

one third of the prior mean of $2 \times 10^{-4}\,\text{h}^{-1}$. The failure-free exposure has moved probability toward lower rates without moving any of it to zero. By equation (8.11) with an MLE of zero, the posterior mean is the prior mean multiplied by the prior weight $w = 10\,000/30\,000 = 1/3$. The 95 % equal-tailed credible interval, from $\chi^2_{4,0.025} = 0.4844$ and $\chi^2_{4,0.975} = 11.143$, is 0.081×10^{-4} to $1.857 \times 10^{-4}\,\text{h}^{-1}$, and the one-sided 95 % upper credible bound is $\chi^2_{4,0.95}/(2b') = 9.4877/60\,000\,\text{h} = 1.581 \times 10^{-4}\,\text{h}^{-1}$.

The classical one-sided 95 % upper confidence bound for zero failures is $\chi^2_{2,0.95}/(2T_{\text{tot}}) = -\ln(0.05)/T_{\text{tot}} = 2.996/20\,000\,\text{h} = 1.498 \times 10^{-4}\,\text{h}^{-1}$, the familiar "rule of three" $3/T_{\text{tot}}$. It equals the Bayesian upper bound under the limiting prior $\text{Gamma}(1, b)$ with $b \to 0$, for which the posterior is $\text{Gamma}(1, T_{\text{tot}})$, an exponential distribution whose 0.95 quantile is $-\ln(0.05)/T_{\text{tot}}$.

8.5.3 Predictive reliability with no failures

For a new device operating 1000 h, equation (8.16) gives

$$R_{\text{pred}}(1000\,\text{h} \mid D) = \left(\frac{30\,000}{31\,000}\right)^2 = 0.9365.$$

Under this prior and likelihood model the predicted 1000 h reliability is 93.65 %, below 100 % although no failure has been observed. Before the test, the same prior predicted $(10\,000/11\,000)^2 = 0.8264$; the test raised the prediction by eleven percentage points without making it certain. The plug-in value at the posterior mean, $\mathrm{e}^{-0.0667} = 0.9355$, is again slightly lower, as equation (8.17) requires.

8.5.4 Sensitivity to the prior

With zero failures, the data inform the rate only through the survival term $\mathrm{e}^{-\lambda T_{\mathrm{tot}}}$, and they do not identify a positive rate through any failure event. The shape of the posterior is then the shape of the prior: $a' = a$, and only the rate parameter grows. The result is therefore particularly sensitive to the prior, and most so when the exposure is short compared with the prior's pseudo-exposure b. Two priors with the same mean, Gamma$(1, 5000\,\mathrm{h})$ and Gamma$(10, 50\,000\,\mathrm{h})$, give after the same 20 000 failure-free device-hours posterior means of 0.40×10^{-4} and $1.43 \times 10^{-4}\,\mathrm{h}^{-1}$, a ratio of 3.6. A Bayesian analysis of sparse or failure-free data must therefore state and justify its prior, and the sensitivity of the conclusion to that choice should be shown.

8.5.5 Worked tables over priors and exposures

The update of equation (8.9) and the predictive reliability of equation (8.16) together form a calculator with five inputs: the prior shape a and rate b, the failures r, the exposure T_{tot} and the mission time t. Its outputs are the posterior Gamma$(a + r, b + T_{\mathrm{tot}})$, the posterior mean rate and the predictive reliability $((b + T_{\mathrm{tot}})/(b + T_{\mathrm{tot}} + t))^{a+r}$. All results assume one common constant first-failure rate, independent items and noninformative censoring; the inputs describe hypothetical datasets, not validated field evidence. Two tables evaluate it for a mission of $t = 1000\,\mathrm{h}$.

Table 8.5 holds the prior at Gamma$(2, 10\,000\,\mathrm{h})$ and the failure count at zero, and varies the failure-free exposure. Each row gives the exposure, the posterior, its mean in failures per 1000 h, the predictive reliability and the one-sided 95 % upper credible bound. The row at 20 000 h is the protection-device example. As exposure grows, the posterior mean falls in proportion to $b/(b + T_{\mathrm{tot}})$ and the predictive reliability rises toward one without reaching it.

Table 8.5: Zero observed failures under the prior Gamma$(2, 10\,000\,\mathrm{h})$: posterior, mean rate, predictive reliability over 1000 h and one-sided 95 % upper credible bound, as the failure-free exposure grows. The rows at 5000, 10 000 and 50 000 h are added to show the trend.

Exposure T_{tot} (h)	Posterior	Mean (per 1000 h)	$R_{\mathrm{pred}}(1000\,\mathrm{h})$	95 % upper bound ($10^{-4}\,\mathrm{h}^{-1}$)
0 (prior)	Gamma$(2, 10\,000)$	0.200	0.8264	4.744
5000	Gamma$(2, 15\,000)$	0.133	0.8789	3.163
10 000	Gamma$(2, 20\,000)$	0.100	0.9070	2.372
20 000	Gamma$(2, 30\,000)$	0.067	0.9365	1.581
50 000	Gamma$(2, 60\,000)$	0.033	0.9675	0.791

Table 8.6 varies the concentration of the prior while keeping its mean at $2 \times 10^{-4}\,\mathrm{h}^{-1}$, from Gamma$(1, 5000\,\mathrm{h})$ to Gamma$(20, 100\,000\,\mathrm{h})$, and applies each prior to the two datasets of the chapter: the eight pumps ($r = 4$, $T_{\mathrm{tot}} = 10\,100\,\mathrm{h}$) and the protection devices ($r = 0$, $T_{\mathrm{tot}} = 20\,000\,\mathrm{h}$). For each dataset the columns give the weight $1 - w = T_{\mathrm{tot}}/(b + T_{\mathrm{tot}})$ of the data in equation (8.11), the posterior mean in failures per 1000 h, and the predictive reliability over 1000 h. One row of the arithmetic, for the prior Gamma$(5, 25\,000\,\mathrm{h})$ and the pumps: $a' = 9$, $b' = 35\,100\,\mathrm{h}$, $1 - w = 10\,100/35\,100 = 0.288$, $\mathbb{E}[\lambda \mid D] = 9/35\,100\,\mathrm{h} = 2.564 \times 10^{-4}\,\mathrm{h}^{-1}$, and $R_{\mathrm{pred}}(1000) = (35\,100/36\,100)^9 = 0.7766$.

The trend is the one equation (8.11) predicts. A more concentrated prior, with the same mean, carries more pseudo-exposure, so the data weigh less and the posterior stays closer to the prior mean of 0.200. For the pumps, whose MLE of 0.396 lies above the prior mean, the posterior mean falls from 0.331 to 0.218 and the predictive reliability rises from 0.726 to 0.805. For the failure-free devices, whose MLE of zero lies below it, the posterior mean rises from

Table 8.6: Effect of prior concentration at a fixed prior mean of $2 \times 10^{-4}\,\mathrm{h}^{-1}$ on the posterior mean (per 1000 h) and the predictive reliability over 1000 h, for the eight pumps and for the failure-free protection devices. $1 - w$ is the weight of the data. The row Gamma$(2, 10\,000)$ is the chapter's prior; the others are added.

	Pumps: $r = 4$, $T_{\mathrm{tot}} = 10\,100$ h			Devices: $r = 0$, $T_{\mathrm{tot}} = 20\,000$ h		
Prior	$1 - w$	mean	R_{pred}	$1 - w$	mean	R_{pred}
Gamma$(1, 5000)$	0.669	0.331	0.7257	0.800	0.040	0.9615
Gamma$(2, 10\,000)$	0.502	0.299	0.7473	0.667	0.067	0.9365
Gamma$(5, 25\,000)$	0.288	0.256	0.7766	0.444	0.111	0.8959
Gamma$(10, 50\,000)$	0.168	0.233	0.7937	0.286	0.143	0.8678
Gamma$(20, 100\,000)$	0.092	0.218	0.8049	0.167	0.167	0.8471

0.040 to 0.167 and the predictive reliability falls from 0.962 to 0.847. A strong prior makes the posterior less sensitive to a given amount of new evidence, in both directions.

Figure 8.5 draws the failure-free case as a function of exposure. The horizontal axis is the failure-free exposure in hours and the vertical axis the predictive reliability over 1000 h; the three curves are three priors with the same mean and increasing concentration. All three start near 0.83, the prior predictive value, and rise with exposure, but the weak prior Gamma$(1, 5000\,\mathrm{h})$ rises fastest and the strong prior Gamma$(10, 50\,000\,\mathrm{h})$ slowest: at 20 000 h, the dot on the middle curve, they give 0.962, 0.937 and 0.868.

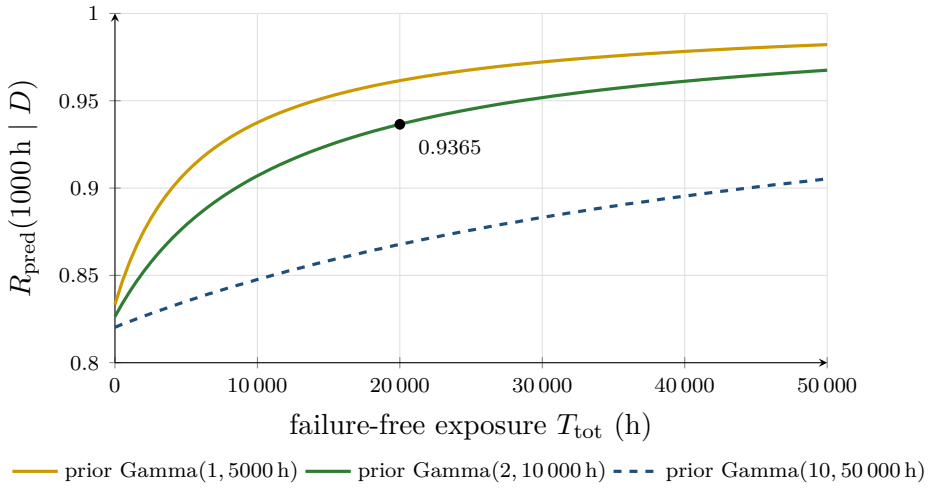


Figure 8.5: Predictive reliability over 1000 h after a failure-free test, $((b+T_{\mathrm{tot}})/(b+T_{\mathrm{tot}}+1000))^a$ (equation (8.16) with $r = 0$), for three illustrative priors with the same mean $2 \times 10^{-4}\,\mathrm{h}^{-1}$. Dot: the ten protection devices, 20 000 device-hours under Gamma$(2, 10\,000\,\mathrm{h})$.

8.6 Sequential updating, equipment populations and model credibility

Evidence arrives over time and from equipment that is related but not identical, and every Bayesian result rests on a model. This section shows that updating in stages gives the same posterior as updating once, extends the model to several related populations, separates parameter uncertainty from model uncertainty, and states when the Bayesian approach is most useful.

8.6.1 Sequential updating

Bayesian updating can be performed as the evidence arrives. The starting point is the pump prior, Gamma(2, 10 000 h). In a first monitoring period, $r_1 = 1$ failure is observed in $T_1 = 3000$ h of exposure; by equation (8.9) the posterior is Gamma(3, 13 000 h), with mean $2.308 \times 10^{-4}\,\mathrm{h}^{-1}$. That posterior becomes the prior for a second period, in which $r_2 = 3$ further failures are observed in $T_2 = 7100$ h; the posterior becomes

$$\text{Gamma}(3 + 3, 13\,000 + 7100) = \text{Gamma}(6, 20\,100\,\text{h}),$$

exactly the posterior obtained by analyzing all four failures and 10 100 pump-hours at once in Section 8.2.3. Figure 8.6 draws the two routes. The top row is the sequential route: the prior, the first posterior, the second posterior, with the evidence of each period on the arrows. The lower path is the single batch update with the pooled evidence. Both arrive at the same distribution.

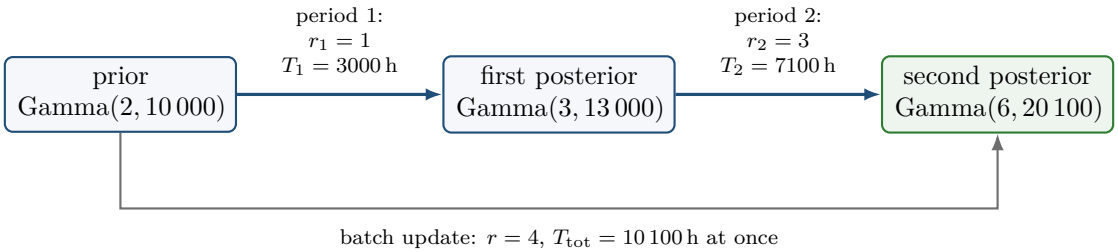


Figure 8.6: Sequential and batch updating of the pump failure rate, gamma rate parameters in hours. Each period's posterior is the next period's prior; the final posterior equals the batch posterior because the likelihood factorizes over periods (equation (8.19)).

The equality holds because the likelihood contributions factorize under the model assumptions. Given λ, the evidence of the two periods is independent, so $p(D_1, D_2 \mid \lambda) = p(D_1 \mid \lambda)\, p(D_2 \mid \lambda)$, and

$$p(\lambda \mid D_1, D_2) \propto p(D_2 \mid \lambda)\, p(D_1 \mid \lambda)\, p(\lambda) \propto p(D_2 \mid \lambda)\, p(\lambda \mid D_1)\,: \tag{8.19}$$

the posterior after the first period plays the role of the prior for the second. For the exponential model each factor is $\lambda^{r_k} \mathrm{e}^{-\lambda T_k}$, and the counts and exposures simply add. A pump that runs through the end of one period into the next contributes a censored observation to the first and continues in the second; splitting its history this way is exact for the exponential model because of the memoryless property, which makes the survival term over the second period, $\mathrm{e}^{-\lambda T}$, independent of the age at which the period starts. For an aging lifetime model the second period's contribution must instead be conditioned on the age reached.

Limits of updating one constant rate. The procedure assumes that the underlying rate is the same in every period. If equipment ages, operating conditions change, or the population changes, a time-varying or hierarchical model is required. An increase in observed failures in a later period may reflect statistical variability, equipment aging, a different operating environment, a manufacturing change, or a failure mechanism previously overlooked. Updating a single constant rate cannot distinguish these explanations: it absorbs them all into one number, whose posterior then describes none of them well.

8.6.2 Related equipment populations

Suppose failure data are available for pumps at three facilities, each operating under different thermal, mechanical and maintenance conditions. Assigning all three one common rate ignores

the differences; analyzing each alone wastes the information the others carry. The intermediate model assigns each facility j its own rate λ_j and treats the three rates as drawn from a common population variability distribution,

$$\lambda_j \sim \text{Gamma}(a, b), \qquad j = 1, 2, 3, \tag{8.20}$$

whose hyperparameters a and b are themselves estimated from the combined data. Estimating (a, b) by maximizing the product over facilities of the marginal likelihood of equation (8.10) is the *empirical Bayes* approach; giving (a, b) a prior of its own, a hyperprior, and integrating over it is the full *hierarchical Bayes* approach. Figure 8.7 draws the structure. The top box is the population distribution with its hyperparameters; the middle row holds the three facility rates; the bottom row holds each facility's evidence, its failure count r_j and exposure T_j. Arrows point from what generates to what is generated; inference runs upward against them.

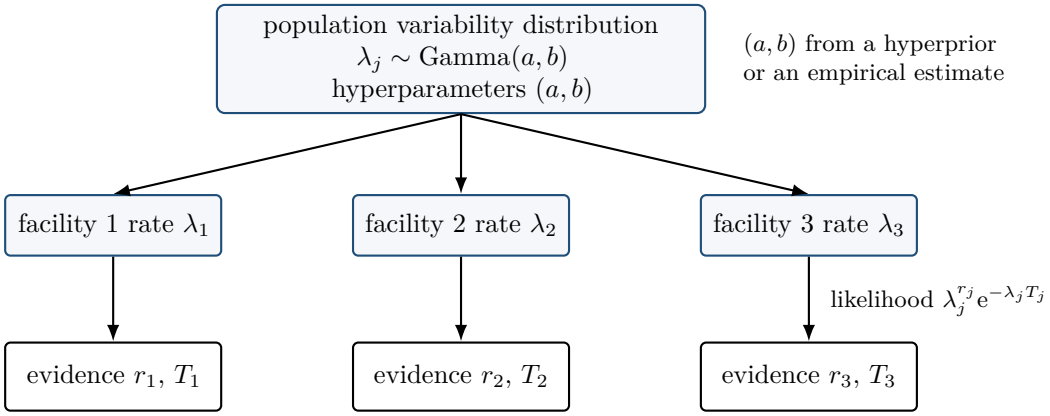


Figure 8.7: A hierarchical model for three related pump populations (illustrative). The facility rates are drawn from a common gamma population variability distribution; each facility's evidence depends on its own rate through the censored exponential likelihood.

Given the hyperparameters, each facility's posterior follows from equation (8.9), $\lambda_j \mid D_j \sim \text{Gamma}(a + r_j, b + T_j)$, and by equation (8.11) its mean is a weighted average of the population mean a/b and the facility's own MLE r_j/T_j, with weight $b/(b + T_j)$ on the population. Table 8.7 illustrates the effect with the population distribution fixed at Gamma(2, 10 000 h) for the purpose of the example. The columns give each facility's evidence, its MLE, the weight on the population, the posterior mean and the predictive reliability over 1000 h. Facility 1, with 2000 h and no failures, is held 83 % of the way to the population mean; facility 3, with 101 000 h, is driven 91 % by its own evidence. Groups with sparse data borrow information from the population; groups with extensive data are governed by their own.

Table 8.7: Partial pooling across three facilities, with the population variability distribution fixed at Gamma(2, 10 000 h) (illustrative values). Rates in $10^{-4}\,\text{h}^{-1}$.

Facility	r_j	T_j (h)	MLE	weight $b/(b + T_j)$	posterior mean	$R_{\text{pred}}(1000\,\text{h})$
1	0	2000	0	0.833	1.667	0.8521
2	4	10 100	3.960	0.498	2.985	0.7473
3	40	101 000	3.960	0.090	3.784	0.6861

This approach is useful whenever related equipment operates in different settings: equipment deployed at many utility substations, similar pumps across several industrial plants, battery modules in different thermal environments, or chillers of the same type in data centers operated

under different policies. The difficulty lies in deciding which populations are genuinely comparable and which differences, such as temperature or load, should be represented explicitly as covariates rather than absorbed into the population spread. Hierarchical modeling does not remove the need for engineering judgment about failure mechanisms and operating conditions.

8.6.3 Parameter uncertainty and model uncertainty

Bayesian updating cannot repair an inadequate reliability model. Consider the exponential model $R(t \mid \lambda) = \mathrm{e}^{-\lambda t}$. Millions of operating hours would give an extremely narrow posterior for λ, since by equation (8.7) the coefficient of variation falls as $1/\sqrt{a + r}$. If the true hazard increases with age, the constant-hazard model will nevertheless underestimate the risk of late life systematically, and the analysis will become increasingly confident in a parameter of the wrong model. The distinction between the physical system, the model and the evidence matters here more than anywhere: the posterior describes the parameter of the model given the evidence, and says nothing about whether the model represents the system.

Parameter uncertainty and model uncertainty must therefore be distinguished, together with other sources that a parameter posterior does not cover. Table 8.8 lists them: each row is a question an analyst would ask about the pump fleet, and the right column names the kind of uncertainty it concerns. Only the first row is addressed by the posterior of λ.

Table 8.8: Questions about the pump fleet and the kinds of uncertainty they concern. The posterior of a constant failure rate addresses only the first.

Question	Uncertainty
What is the pump's constant failure rate?	parameter uncertainty
Does the hazard actually remain constant?	model uncertainty
Are the observed pumps representative of the equipment assessed?	population and operating-condition uncertainty
Are failures recorded accurately?	measurement and observation uncertainty
Could a common-cause failure affect several pumps?	system structure and dependency uncertainty

The analysis should therefore evaluate whether the probability model represents the equipment, its operating conditions and its relevant failure mechanisms adequately, for example by comparing the fitted exponential reliability with the Kaplan–Meier estimate of Chapter 7 and by examining the failure mechanisms behind the recorded failures.

> **Principle 8.3.** *A credible posterior requires a credible model. Bayesian updating narrows the uncertainty about the parameters of the assumed model; it does not test or correct the assumption that the model represents the failure process.*

8.6.4 Bayesian model comparison

Uncertainty about the model structure can itself be treated probabilistically. Consider two candidate models, M_1 the exponential lifetime and M_2 the Weibull lifetime of Chapter 6, with prior model probabilities $\Pr(M_1)$ and $\Pr(M_2)$. Applying Bayes' theorem to the models rather than to a parameter,

$$\Pr(M_j \mid D) = \frac{p(D \mid M_j)\,\Pr(M_j)}{\sum_k p(D \mid M_k)\,\Pr(M_k)}, \qquad p(D \mid M_j) = \int p(D \mid \boldsymbol{\theta}_j, M_j)\, p(\boldsymbol{\theta}_j \mid M_j)\, \mathrm{d}\boldsymbol{\theta}_j, \tag{8.21}$$

where $\boldsymbol{\theta}_j$ is the parameter vector of model M_j (λ for the exponential, the shape β and scale η for the Weibull) and $p(D \mid M_j)$ is the marginal likelihood of equation (8.2) under that model. For the exponential model with a gamma prior it is available in closed form, equation (8.10); for the Weibull model it requires a two-dimensional numerical integral over (β, η), with the same censored likelihood written with the Weibull density and reliability. The ratio $p(D \mid M_2)/p(D \mid M_1)$ is the *Bayes factor* of the Weibull model against the exponential. Because a marginal likelihood averages the fit over the prior, a model with more parameters spread over implausible values is penalized automatically.

This introduces uncertainty about the model structure into the analysis. The resulting model probabilities depend on the set of candidate models and on their parameter priors; they are relative to that set, and they do not prove that the true physical mechanism has been identified. Model credibility returns in the later chapters on system reliability and dependencies.

8.6.5 When Bayesian reliability analysis is most useful

Bayesian methods are particularly attractive in the situations of Table 8.9, where each row pairs an engineering situation with what the Bayesian treatment contributes to it. They are not automatically superior to frequentist methods. When large, representative datasets exist and simple models are appropriate, classical estimators perform very well, and with abundant data the posterior is dominated by the likelihood, as equation (8.11) shows. The advantage of the Bayesian analysis is its explicit treatment of uncertainty and its ability to combine sources of information within one probabilistic model. The quality of the result still depends on the validity of the prior, the likelihood and the engineering assumptions behind both.

Table 8.9: Situations in which Bayesian reliability analysis is particularly useful, and its contribution in each.

Situation	Bayesian contribution
New equipment with few failures	incorporates defensible prior engineering information
Zero-failure demonstration tests	quantifies the remaining uncertainty without declaring a zero failure rate
Several related fleets	shares information through a hierarchical model
Sequential field monitoring	updates the estimates as observations arrive
Uncertain future failures	integrates parameter uncertainty into the predictive reliability
Reliability-based decisions	gives probabilities and expected consequences under the posterior

The chapter has treated the reliability of one type of component. The next chapter, Chapter 9, combines component reliabilities into system reliability through block diagrams, series and parallel structures and redundancy.

8.7 What to retain

Definition 8.1 (Bayesian reliability update)**.** Bayesian reliability analysis represents an unknown failure rate by a distribution of knowledge and updates it with evidence: a gamma prior Gamma(a, b) and r failures in a total exposure T_{tot}, censored exposure included, give the posterior Gamma$(a + r, b + T_{\text{tot}})$, and the reliability of a new item is the exponential reliability averaged over that posterior, $(b'/(b' + t))^{a'}$, which a credible model is required to make meaningful.

Table 8.10 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 8.10: The terms of Bayesian reliability analysis and where each is defined. Units assume time in hours.

Term	Expression	Units	Meaning	Defined in
Prior	$p(\lambda)$	h	knowledge before the evidence	8.1.2
Likelihood	$L(\lambda) \propto \lambda^{r} \mathrm{e}^{-\lambda T_{\mathrm{tot}}}$	–	evidence as a function of λ	8.1.4
Posterior	$p(\lambda \mid D) \propto L(\lambda)\, p(\lambda)$	h	knowledge after the evidence	8.1.3
Marginal likelihood	$p(D) = \int L(\lambda)\, p(\lambda)\, \mathrm{d}\lambda$	–	normalizer; model evidence	8.1.3
Total exposure	$T_{\mathrm{tot}} = \sum_i t_i$	h	hours of failed and censored items	8.1.4
Gamma prior	Gamma(a, b), mean a/b	–, h	shape a, rate b	8.2.1
Conjugate update	Gamma$(a + r, b + T_{\mathrm{tot}})$	–, h	failures to shape, exposure to rate	8.2.2
Data weight	$1 - w = T_{\mathrm{tot}}/(b + T_{\mathrm{tot}})$	–	weight of the MLE in the mean	8.2.4
Pseudo-exposure	a failures in b hours	–, h	algebraic strength of the prior	8.2.5
Credible interval	$\Pr(\lambda_L < \lambda < \lambda_U \mid D) = 0.95$	h^{-1}	posterior probability statement	8.3.2
Exceedance probability	$\Pr(\lambda > \lambda_{\mathrm{crit}} \mid D)$	–	probability a limit is violated	8.3.4
Predictive reliability	$R_{\mathrm{pred}}(t) = (b'/(b' + t))^{a'}$	–	survival averaged over the posterior	8.4.2
Predictive hazard	$\lambda_{\mathrm{pred}}(t) = a'/(b' + t)$	h^{-1}	decreasing; selection effect	8.4.4
Population variability distribution	$\lambda_j \sim$ Gamma(a, b)	–	spread of rates among groups	8.6.2
Bayes factor	$p(D \mid M_2)/p(D \mid M_1)$	–	evidence for one model over another	8.6.4

Further reading. Hamada, Wilson, Reese and Martz, *Bayesian Reliability*: Bayesian lifetime inference, prior distributions, posterior reliability and predictive analysis. Rausand, Barros and Høyland, *System Reliability Theory*, third edition: reliability estimation and its statistical foundations. Modarres, Kaminskiy and Krivtsov, *Reliability Engineering and Risk Analysis*, third edition: uncertainty quantification and probabilistic reliability analysis.

8.8 Exercises

Exercise 8.1. (a) A gamma prior Gamma(a, b) on an exponential failure rate is updated with r failures over a total exposure T_{tot}. Which is the posterior: Gamma$(a + r, b + T_{\mathrm{tot}})$, Gamma$(a + T_{\mathrm{tot}}, b + r)$ or Gamma(ar, bT_{tot})? (Gamma$(a + r, b + T_{\mathrm{tot}})$: the failure count updates the shape and the observed operating exposure updates the rate.) (b) State what a 95 % Bayesian credible interval means. (Under the specified model and prior, the parameter has 95 % posterior probability of lying in the interval; it is a probability statement about the parameter conditional on the data and the assumptions, not a coverage property over repeated samples, and not a statement that 95 % of future components survive.) (c) Can a narrow posterior

for a constant failure rate establish that the exponential model is correct? (No: a precisely estimated parameter does not establish model adequacy or remove the uncertainty about failure mechanisms and operating conditions.)

Exercise 8.2. A prior for the failure rate of a fan is Gamma$(3, 15\,000\,\mathrm{h})$. (a) Compute the prior mean, the prior standard deviation and $\mathbb{E}[1/\lambda]$. ($2 \times 10^{-4}\,\mathrm{h}^{-1}$, $1.155 \times 10^{-4}\,\mathrm{h}^{-1}$, 7500 h.) (b) Field data give $r = 2$ failures in $T_{\mathrm{tot}} = 8000\,\mathrm{h}$. Compute the posterior, its mean, and the weights of the prior mean and the MLE in equation (8.11). (Gamma$(5, 23\,000\,\mathrm{h})$; $2.174 \times 10^{-4}\,\mathrm{h}^{-1}$; 0.652 on the prior mean 2×10^{-4} and 0.348 on the MLE $2.5 \times 10^{-4}\,\mathrm{h}^{-1}$.) (c) Compute the 95 % equal-tailed credible interval from $\chi^2_{10,0.025} = 3.247$ and $\chi^2_{10,0.975} = 20.483$, and the posterior probability that the rate exceeds $3 \times 10^{-4}\,\mathrm{h}^{-1}$ with equation (8.14). (0.706×10^{-4} to $4.453 \times 10^{-4}\,\mathrm{h}^{-1}$; $b'x = 6.9$ and the five Poisson terms add to 0.182.) (d) Compute the predictive and the plug-in reliability over 500 h, and explain why they differ. ($(23\,000/23\,500)^5 = 0.8980$ against $\mathrm{e}^{-0.1087} = 0.8970$; the exponential survival function is nonlinear, convex in λ, so its posterior average differs from its value at the posterior mean, and by equation (8.17) exceeds it.)

Exercise 8.3. Twelve new protective relays each operate 1500 h without failure. (a) Compute T_{tot} and the MLE, and state what can be concluded from zero failures over this exposure. (18 000 h and 0; the survival observations favor lower rates but do not establish a zero failure risk.) (b) With the prior Gamma$(1, 5000\,\mathrm{h})$, compute the posterior and its mean. (Gamma$(1, 23\,000\,\mathrm{h})$, an exponential distribution; $4.348 \times 10^{-5}\,\mathrm{h}^{-1}$.) (c) Compute the predictive reliability of a relay over 1000 h and the posterior probability that its rate exceeds $10^{-4}\,\mathrm{h}^{-1}$. ($23\,000/24\,000 = 0.9583$; $\mathrm{e}^{-2.3} = 0.1003$.) (d) Compute the one-sided 95 % upper credible bound, and compare it with the classical zero-failure bound $-\ln(0.05)/T_{\mathrm{tot}}$. ($-\ln(0.05)/23\,000\,\mathrm{h} = 1.302 \times 10^{-4}\,\mathrm{h}^{-1}$ against $1.664 \times 10^{-4}\,\mathrm{h}^{-1}$; the prior's 5000 h of pseudo-exposure lowers the bound.)

Exercise 8.4. The pump prior Gamma$(2, 10\,000\,\mathrm{h})$ is updated over three monitoring periods: no failure in 4000 h; two failures in 6000 h; one failure in 5000 h. (a) Give the posterior after each period, and show that the last equals the batch update with the pooled evidence. (Gamma$(2, 14\,000)$, Gamma$(4, 20\,000)$, Gamma$(5, 25\,000\,\mathrm{h})$; the batch update with $r = 3$ and $T_{\mathrm{tot}} = 15\,000\,\mathrm{h}$ gives the same.) (b) Compute the posterior mean, the predictive reliability over 2000 h and the plug-in reliability at the posterior mean. ($2 \times 10^{-4}\,\mathrm{h}^{-1}$; $(25\,000/27\,000)^5 = 0.6806$; $\mathrm{e}^{-0.4} = 0.6703$.) (c) Compute the predictive hazard at ages 0 and 2000 h, and the predictive mean lifetime. (2.000×10^{-4} and $1.852 \times 10^{-4}\,\mathrm{h}^{-1}$; $25\,000/4 = 6250\,\mathrm{h}$, against $1/\mathbb{E}[\lambda \mid D] = 5000\,\mathrm{h}$.) (d) The failure in the third period occurred on a pump moved to a hotter building. State what the pooled update assumes and what model would represent the change. (It assumes one constant rate common to all periods and pumps; a covariate-dependent hazard or a hierarchical model with a separate rate for the hotter location would represent it.)

Part III

System Reliability and Fault Trees

Chapter 9

Reliability Block Diagrams and Redundancy

9.1 System reliability from component reliability

The preceding chapters treated one item at a time. They described the lifetime of a component by its reliability function, hazard function and lifetime distribution, estimated those functions from failure and survival data with censoring, and expressed the uncertainty in their parameters with Bayesian methods. For a component i they produce a reliability function

$$R_i(t) = \Pr(T_i > t),$$

where T_i is the random time to failure of component i. An industrial system rarely depends on one component. A data center may have three chillers, four pumps, redundant power feeds and numerous shared controllers; a transmission network contains lines, transformers, breakers and protection systems whose failures can interact. The step from component reliability to system reliability is the subject of this chapter, which opens the part of the book on system reliability and fault trees. Its question is

how the reliability of a system is calculated from the reliabilities of its components, and under what circumstances redundancy improves system reliability, or fails to improve it.

The chapter answers it in six steps. It first defines the reliability block diagram and the structure function that the diagram represents. It then derives the reliability of series systems, of parallel systems and of k-out-of-n systems, and compares redundancy in which all units run with redundancy in which a spare waits to be switched in. It shows how dependence between component failures, and common-cause failure in particular, defeats the multiplication rules on which those results rest. It closes with structures that combine series and parallel elements, with a first measure of which component matters most, and with the limitation of block diagrams whose components are either working or failed.

Jargon: reliability block diagram, structure function, redundancy, active and standby redundancy, k-out-of-n

A *reliability block diagram* (RBD) is a graph of blocks, each the success event of a component or subsystem, connected between an input and an output so that the system functions when a path of functioning blocks joins the two. The *structure function* ϕ is the Boolean function that the diagram represents: it maps the states of the components, each 1 (functioning) or 0 (failed), to the state of the system. *Redundancy* is the provision of more means of performing a function than the function requires, so that it survives the loss of some of them. In *active redundancy* the redundant units operate together and share the load; in *standby redundancy* a spare is held out of operation and switched in when the operating unit fails, and *cold standby* is the case in which the spare cannot fail while it waits. A k-out-of-n system has n units and functions when at least k of them function.

Key idea

A reliability block diagram separates which combinations of functioning components make a system work, the structure function, from how likely each combination is, the probability model. Series arrangements multiply reliabilities and parallel arrangements multiply failure probabilities, but redundancy improves a system only when each redundant unit can carry the required function, the units do not share a cause of failure, and any switching they need succeeds.

9.1.1 Reliability block diagrams

A reliability block diagram is a graphical representation of the combinations of functioning components that are required for a system to perform its required function. Each block represents a component or a subsystem, and stands for the event that this item performs its required function over the mission. Blocks are joined by lines between an input terminal on the left and an output terminal on the right. A path through the diagram, from input to output through a chain of blocks, represents a combination of components whose successful operation is sufficient for system success, and the system functions when at least one path consists entirely of functioning blocks.

The diagram is a diagram of functional success logic. It is not necessarily a picture of the physical flow of electricity, water, heat or information, and two diagrams of the same plant drawn for two different required functions can differ. Consider a cooling system containing two pumps. If the required flow needs both pumps, the pumps are arranged in series in the block diagram, whatever the piping looks like; if either pump alone can satisfy the cooling requirement, they are arranged in parallel. Chapter 1 computed both cases for two pumps of reliability 0.95: 0.9025 in series and 0.9975 in parallel. The reverse also occurs. Two shut-off valves installed one after the other in a pipe are physically in series, but if the required function is to stop the flow, either valve closing suffices, and in the block diagram for that function the two valves are in parallel. The block diagram is therefore drawn from the definition of the required function and the system boundary, treated in Chapter 2, not from the piping and wiring drawings.

9.1.2 Boolean success events and the structure function

The block diagram is made precise with one binary variable per component. Let the system have n components and a common mission of duration t. For component i, let the state variable

$$S_i(t) = \begin{cases} 1 & \text{if component } i \text{ performs its required function throughout the mission,} \\ 0 & \text{otherwise,} \end{cases} \tag{9.1}$$

so that $\{S_i(t) = 1\}$ is the success event of component i and $\{S_i(t) = 0\}$ its failure event. For a nonrepairable component in a binary-state model, success throughout the mission is survival beyond its end,

$$S_i(t) = \mathbf{1}\{T_i > t\}, \tag{9.2}$$

and $\Pr(S_i(t) = 1) = R_i(t)$. The vector $S(t) = (S_1(t), \ldots, S_n(t))$ collects the component states. The system's state is a function of it,

$$\phi : \{0,1\}^n \to \{0,1\}, \qquad \text{system functions} \iff \phi\big(S(t)\big) = 1, \tag{9.3}$$

and ϕ is the structure function. The system reliability over the mission is the probability of the system success event,

$$R_{\text{system}}(t) = \Pr\big(\phi(S(t)) = 1\big) = \mathbb{E}\big[\phi(S(t))\big], \tag{9.4}$$

where the second form holds because ϕ takes only the values 0 and 1, and the expected value of an indicator is the probability of its event.

Equation (9.4) has two ingredients, and they come from different places. The structure function describes *which combinations of component states succeed*; it is fixed by the required function, the system boundary and the arrangement of the components, and it contains no probability. The probability model describes *how likely those combinations are*; it comes from the component reliabilities of the earlier chapters and from assumptions about how the component states depend on one another. When the component states are independent, the system reliability is a function of the component reliabilities alone,

$$R_{\text{system}} = h(R_1, \ldots, R_n), \tag{9.5}$$

called the reliability function of the structure, and the rest of the chapter derives h for the common structures. A set of components whose functioning ensures that the system functions, whatever the other components do, is called a path set; a path set with no superfluous member is a minimal path set, written P_j. A path through the block diagram is a path set. The separation of the logic from the probability recurs throughout system reliability theory, in fault trees and minimal cut sets as much as in block diagrams.

9.2 Series systems

The simplest structure requires every component. This section derives the reliability of a series system, shows how it falls as components are added, specializes it to components with constant failure rates, and works an example.

9.2.1 The series structure

Consider a cooling-water system consisting of an electric motor, a pump and a heat exchanger, all three necessary to deliver the required cooling function. Figure 9.1 draws its reliability block diagram. The three blocks lie on one line between the input on the left and the output on the right, and each is labeled with its reliability over the mission: 0.98 for the motor, 0.97 for the pump and 0.99 for the heat exchanger (illustrative values). There is one path, and it passes through all three blocks.

The system succeeds if and only if every required component succeeds. Its success event is the intersection of the component success events,

$$\{\phi(S) = 1\} = \{S_1 = 1\} \cap \{S_2 = 1\} \cap \cdots \cap \{S_n = 1\}, \tag{9.6}$$

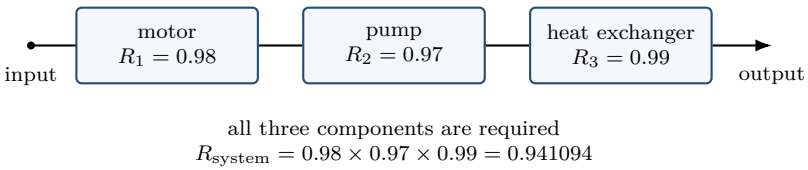


Figure 9.1: Reliability block diagram of a series cooling-water system: a motor, a pump and a heat exchanger, each required, with illustrative mission reliabilities and independent failures. The single path passes through every block, so the failure of any one block fails the system.

and its structure function is the product of the states,

$$\phi(S) = \prod_{i=1}^{n} S_i = \min_i S_i, \tag{9.7}$$

which is 1 exactly when no factor is 0. Before any assumption about dependence, the intersection in equation (9.6) is contained in each of its events, so $R_{\text{system}} \leq \min_i R_i$: a series system is never more reliable than its least reliable component.

To compute the probability of the intersection, apply the multiplication rule of probability,

$$\Pr(E_1 \cap E_2 \cap \cdots \cap E_n) = \Pr(E_1)\, \Pr(E_2 \mid E_1) \cdots \Pr(E_n \mid E_1 \cap \cdots \cap E_{n-1}),$$

with $E_i = \{S_i = 1\}$. If the component success events are independent, every conditional probability equals the unconditional one, $\Pr(E_i \mid E_1 \cap \cdots \cap E_{i-1}) = R_i$, and

$$R_{\text{series}}(t) = \prod_{i=1}^{n} R_i(t). \tag{9.8}$$

For the cooling-water system,

$$R_{\text{system}} = 0.98 \times 0.97 \times 0.99 = 0.941094,$$

a mission reliability of about 94.11 %. All three components have a reliability of 97 % or more, and the system has less than any of them. Its failure probability, $1 - 0.941094 = 0.058906$, is close to the sum of the component failure probabilities, $0.02 + 0.03 + 0.01 = 0.06$. The sum is always an upper bound, since the probability of a union of failure events cannot exceed the sum of their probabilities, and it is a good approximation when the component failure probabilities are small: in a series system, failure probabilities approximately add.

9.2.2 Reliability and the number of required components

Suppose every component has the same reliability R over the mission. For n independent components in series, equation (9.8) becomes

$$R_{\text{series}} = R^n. \tag{9.9}$$

Figure 9.2 draws this for $R = 0.99$. The horizontal axis is the number n of required components, from 1 to 100, and the vertical axis is the system reliability. The curve is equation (9.9) and the dots mark the values at $n = 1, 5, 10, 20, 30, 40, 50, 75$ and 100; the dashed horizontal line marks a reliability of one half.

The figure shows the reliability falling steadily with the number of required components: 0.951 for five, 0.904 for ten, 0.818 for twenty, 0.605 for fifty and 0.366 for one hundred. The system becomes more likely to fail than to survive when $0.99^n < 0.5$, that is, when $n > \ln 0.5 / \ln 0.99 = 68.97$, so from $n = 69$. For small total failure probability, $1 - R^n \approx n(1 - R)$:

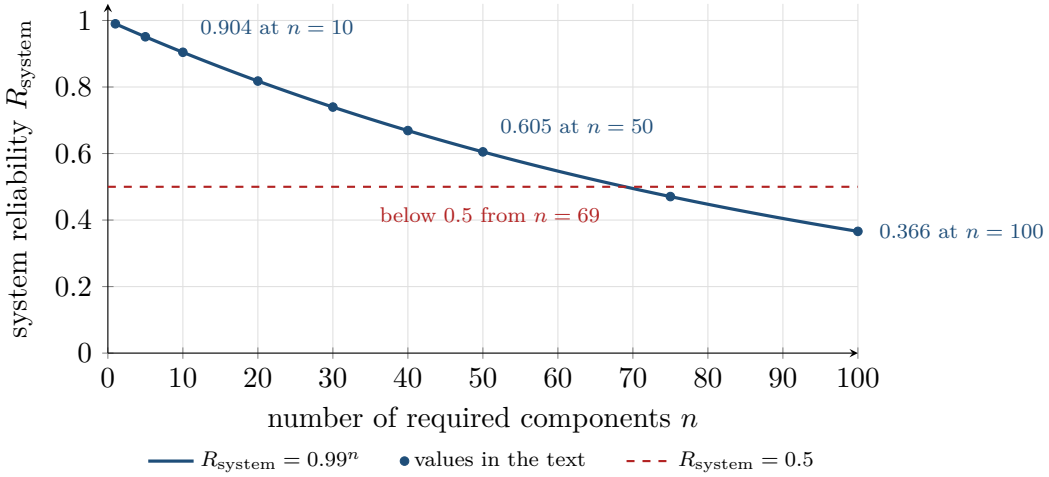


Figure 9.2: Reliability of a series system of n independent components, each of reliability 0.99 over the same mission, $R_{\text{system}} = 0.99^n$ (equation (9.9)). The system reliability falls below one half when 69 or more components are required.

ten components give 0.0956 against the approximation 0.10. The approximation fails as $n(1-R)$ approaches one, and at $n = 100$ it would predict a failure probability of 1 against the exact 0.634. Individually reliable components can therefore produce a system with low mission reliability when many of them must all succeed.

This does not mean that engineering complexity always reduces reliability. The decrease holds for components that are all *required*. Redundancy, fault tolerance and alternative functional pathways change the structure function, and the sections that follow show that additional components arranged in parallel raise the reliability instead.

9.2.3 Exponential components and the sum of failure rates

Suppose component i has an exponential lifetime with constant failure rate λ_i, the standard model of Chapter 6,

$$R_i(t) = \mathrm{e}^{-\lambda_i t}.$$

For independent components in series, equation (9.8) gives

$$R_{\text{series}}(t) = \prod_{i=1}^{n} \mathrm{e}^{-\lambda_i t} = \exp\Big(-t \sum_{i=1}^{n} \lambda_i\Big), \tag{9.10}$$

because a product of exponentials is the exponential of the sum of their exponents. The system reliability is again exponential, with failure rate

$$\lambda_{\text{sys}} = \sum_{i=1}^{n} \lambda_i \qquad \text{and} \qquad \text{MTTF}_{\text{sys}} = \frac{1}{\lambda_{\text{sys}}} = \frac{1}{\sum_i \lambda_i}. \tag{9.11}$$

The same result reads as a statement about lifetimes. A series system fails at the first component failure, so its lifetime is $T_{\text{sys}} = \min_i T_i$, and $\Pr(\min_i T_i > t) = \Pr(T_1 > t, \ldots, T_n > t)$. The minimum of independent exponentially distributed lifetimes is therefore itself exponentially distributed, with rate equal to the sum of the rates. A more general statement holds for any lifetime distributions. Since the hazard function is $\lambda(t) = -\mathrm{d}\ln R(t)/\mathrm{d}t$ (Chapter 5), taking the logarithm of equation (9.8) and differentiating gives $\lambda_{\text{sys}}(t) = \sum_i \lambda_i(t)$: the hazard of a series system of independent components is the sum of the component hazards, and the exponential case is the one in which every term is constant.

9.2.4 Worked example: three constant failure rates

Suppose three required components have constant failure rates

$$\lambda_1 = 0.0001\,\mathrm{h}^{-1}, \qquad \lambda_2 = 0.0002\,\mathrm{h}^{-1}, \qquad \lambda_3 = 0.00005\,\mathrm{h}^{-1}.$$

By equation (9.11) the system failure rate is

$$\lambda_{\mathrm{sys}} = 0.0001 + 0.0002 + 0.00005 = 0.00035\,\mathrm{h}^{-1},$$

and the mean time to failure of the system is $1/0.00035 = 2857\,\mathrm{h}$, against 10 000, 5000 and 20 000 h for the three components. For a mission of 1000 h,

$$R_{\mathrm{system}}(1000\,\mathrm{h}) = \mathrm{e}^{-0.00035\times 1000} = \mathrm{e}^{-0.35} = 0.7047.$$

Under these assumptions the system has a probability of about 70.47 % of completing the mission without a failure of a required component.

Two further numbers follow from the same model. First, the probability that a given component is the one that ends the mission is its share of the total rate, $\lambda_i/\lambda_{\mathrm{sys}}$, a standard result for independent exponential lifetimes: $2/7 = 0.286$ for component 1, $4/7 = 0.571$ for component 2 and $1/7 = 0.143$ for component 3. Component 2 causes more than half of the system failures, and it is the first candidate for improvement. Second, the rare-event approximation $1 - \mathrm{e}^{-\lambda t} \approx \lambda t$ gives a failure probability of 0.35, against the exact $1 - 0.7047 = 0.2953$; at $\lambda_{\mathrm{sys}} t = 0.35$ the mission is long enough that the approximation overstates the failure probability by about a fifth, and the exact exponential should be used.

9.3 Parallel and k-out-of-n systems

Redundancy enters the block diagram as parallel paths. This section derives the reliability of a parallel system, shows what additional redundant units buy and what supporting equipment costs, and generalizes the result to systems that need k of n units, for identical and for unequal components.

9.3.1 The parallel structure

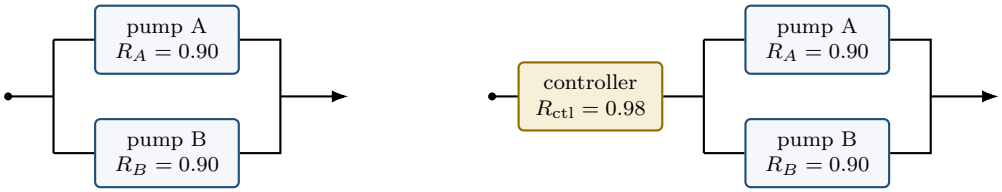
Suppose two independent pumps are installed in parallel, and either pump can supply the full required cooling-water flow. Panel (a) of Figure 9.3 draws the reliability block diagram: the line from the input splits into two branches, one through each pump, and the branches join before the output. Each pump has a mission reliability of 0.90 (illustrative). There are two paths, and the system functions if either is intact. The parallel model applies only if each pump alone is capable of providing the complete required function; two pumps that each deliver half the required flow are a series system in the block diagram, although they are piped in parallel.

The system succeeds if at least one pump succeeds, so its success event is the union $\{S_A = 1\} \cup \{S_B = 1\}$. Equivalently, it fails only if both pumps fail. Write the component failure events as $F_A = \{S_A = 0\}$ and $F_B = \{S_B = 0\}$, the complements of the success events. The system failure event is their intersection, $F_{\mathrm{sys}} = F_A \cap F_B$, and

$$R_{\mathrm{system}} = 1 - \Pr(F_A \cap F_B). \tag{9.12}$$

If the failure events are independent, $\Pr(F_A \cap F_B) = \Pr(F_A)\,\Pr(F_B)$, and since $\Pr(F_i) = 1 - R_i$,

$$R_{\mathrm{parallel}} = 1 - (1 - R_A)(1 - R_B). \tag{9.13}$$



(a) either pump sufficient
$R_{\text{system}} = 1 - 0.10^2 = 0.99$

(b) a controller required by both pumps
$R_{\text{system}} = 0.98 \times 0.99 = 0.9702$

Figure 9.3: Reliability block diagrams of two redundant pumps with illustrative mission reliabilities and independent failures. (a) Either pump alone can provide the complete required flow, so the pumps are in parallel (equation (9.13)). (b) Both pumps need one shared controller, which sits in series with the parallel pair (equation (9.16)).

For n independent components in parallel the same argument gives

$$R_{\text{parallel}} = 1 - \prod_{i=1}^{n}(1 - R_i), \qquad \phi(S) = 1 - \prod_{i=1}^{n}(1 - S_i) = \max_i S_i, \tag{9.14}$$

where the structure function is 0 only when every state is 0. With $R_A = R_B = 0.90$,

$$R_{\text{system}} = 1 - (0.10)(0.10) = 1 - 0.01 = 0.99.$$

Two pumps with an individual reliability of 90 % achieve a system reliability of 99 %, provided that either pump alone is sufficient and that their failures are independent. Where the failure probabilities of a series system approximately add, those of a parallel system multiply.

9.3.2 Additional redundant pumps

For n identical, independent components in parallel, equation (9.14) becomes

$$R_{\text{parallel}} = 1 - (1 - R)^n. \tag{9.15}$$

Table 9.1 evaluates it for pumps of reliability 0.90. Each row is a number of pumps installed, any one of which is sufficient; the columns give the system reliability and the system failure probability $(0.10)^n$. Each added pump divides the failure probability by ten, from 10 % with one pump to 0.01 % with four.

Table 9.1: Reliability of n identical pumps in parallel, each of reliability 0.90, any one sufficient, independent failures (equation (9.15)).

Pumps n	System reliability	Failure probability
1	0.90	0.10
2	0.99	0.01
3	0.999	0.001
4	0.9999	0.0001

The improvements in the table rest on several assumptions, and each must be checked before the numbers are used:

- each pump can meet the demand alone;
- the pumps share no cause of failure;
- any switching or control function that redundancy needs succeeds;

- the system stays within its functional criteria when a pump fails, for example without a pressure transient that trips the others;
- every component starts the mission in its required condition, none out of service for maintenance.

Real engineered systems frequently violate one or more of these assumptions, and the failure probabilities of the lower rows of the table are then far smaller than anything the plant achieves. A failure probability of 10^{-4} for four pumps means that a cause shared by all four with a probability above 10^{-4} dominates the result; the rest of the chapter treats three such causes: supporting equipment, switching and common-cause failure.

9.3.3 Redundancy as an added failure opportunity

Adding a pump adds more than a pump. It may require additional valves, sensors, electrical breakers, control software and switching logic, and those supporting components can fail. Consider two redundant pumps that both require one shared controller, of reliability $R_{\text{ctl}} = 0.98$ over the mission, panel (b) of Figure 9.3. The controller is required whatever the pumps do, so in the block diagram it is in series with the parallel pair. If the controller's success is independent of the pumps', the complete system reliability is

$$R_{\text{system}} = R_{\text{ctl}} \left[1 - (1 - R_A)(1 - R_B)\right], \tag{9.16}$$

and with $R_A = R_B = 0.90$,

$$R_{\text{system}} = 0.98 \times 0.99 = 0.9702.$$

The complete system has a reliability of 97.02 %, not 99 %. Its failure probability, 0.0298, is three times that of the pumps alone, and two thirds of it, 0.02, comes from the controller. A component whose failure alone fails the system is called a single point of failure; the controller is one, and it caps the system reliability at its own 0.98 however many pumps are installed. This is the simplest form of a shared dependency. It is not yet a common-cause failure of the pumps themselves, which Section 9.5.3 takes up.

9.3.4 The two-out-of-three chiller plant

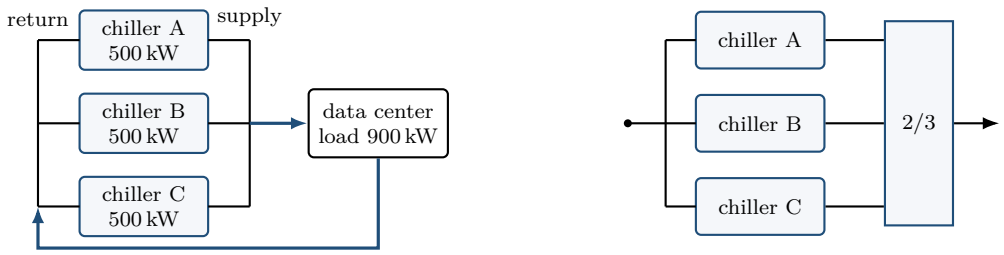
Not every redundant system succeeds when only one component survives. Consider the three chillers of the data center cooling plant. In this chapter each chiller has a capacity of 500 kW and the cooling demand is 900 kW (illustrative). One chiller alone cannot meet the demand and any two can, so the plant requires at least two of its three chillers to remain operational: it is a 2-out-of-3 system, commonly written 2oo3.

Figure 9.4 shows why the physical arrangement and the success logic must be kept apart. Panel (a) is the physical arrangement: three chillers of 500 kW connected in parallel between a supply and a return header that serve the data center load of 900 kW. Panel (b) is the reliability block diagram: the three chiller blocks feed a voting element, marked 2/3, that passes the function on when at least two of its inputs are working. The plant is physically parallel, but its success criterion is 2-out-of-3, not 1-out-of-3.

The success event is that the number of functioning chillers is at least two,

$$\{\phi(S) = 1\} = \{S_A + S_B + S_C \geq 2\}. \tag{9.17}$$

It is the union of two mutually exclusive cases: all three chillers survive, or exactly two survive. With identical, independent chillers of reliability R, the first case has probability R^3. In the second, there are $\binom{3}{2} = 3$ ways to choose the two survivors, and each way has probability



(a) physical arrangement: three units in parallel piping

(b) block diagram: any two of three required

Figure 9.4: The data center chiller plant with three chillers of 500 kW and a cooling demand of 900 kW (illustrative). (a) The physical arrangement: the chillers are piped in parallel. (b) The reliability block diagram: because one chiller cannot carry the demand and any two can, the success criterion is 2-out-of-3, drawn as a voting element, not 1-out-of-3.

$R^2(1-R)$, so the case has probability $3R^2(1-R)$. The two cases cannot occur together, and their probabilities add,

$$R_{\text{2-out-of-3}} = R^3 + 3R^2(1-R) = 3R^2 - 2R^3. \tag{9.18}$$

For $R = 0.95$,

$$R_{\text{2-out-of-3}} = 0.95^3 + 3\,(0.95)^2(0.05) = 0.857375 + 0.135375 = 0.99275,$$

a system reliability of 99.275 %, the value obtained for the same plant in Chapter 1.

9.3.5 General k-out-of-n reliability

The chiller result generalizes. Suppose a system has n identical components with independent success events, each of reliability $R(t)$ over the mission, and it requires at least k of them to succeed. The number of successful components, $N_{\text{s}} = \sum_{i=1}^{n} S_i(t)$, counts the successes in n independent trials with the same success probability, so it is a binomial random variable,

$$N_{\text{s}} \sim \text{Binomial}\big(n, R(t)\big), \qquad \Pr(N_{\text{s}} = j) = \binom{n}{j} R(t)^j \big[1 - R(t)\big]^{n-j},$$

where $\binom{n}{j} = n!/[j!\,(n-j)!]$ is the number of ways to choose the j survivors. The system succeeds when $N_{\text{s}} \geq k$, and the events $\{N_{\text{s}} = j\}$ for different j are mutually exclusive, so

$$R_{k\text{-out-of-}n}(t) = \sum_{j=k}^{n} \binom{n}{j} R(t)^j \big[1 - R(t)\big]^{n-j}. \tag{9.19}$$

Equation (9.18) is the case $k = 2$, $n = 3$.

Series and parallel systems are the two extreme cases. For $k = n$ only the term $j = n$ remains,

$$R_{n\text{-out-of-}n} = R^n,$$

the series system of equation (9.9). For $k = 1$ the sum runs over every $j \geq 1$, which is everything except $j = 0$,

$$R_{1\text{-out-of-}n} = 1 - (1-R)^n,$$

the parallel system of equation (9.15). Between the two, raising k removes terms from the sum, so the reliability of a k-out-of-n system with fixed components can only fall as more of them are required.

9.3.6 Components with unequal reliabilities

The binomial formula assumes that all components have the same reliability. If the components are independent but not identical, the system reliability is the sum of the probabilities of all component configurations that contain at least k functioning components. Let $J \subseteq \{1, \ldots, n\}$ denote the set of functioning components in a configuration. The probability of that configuration is the product of R_i over the members of J and of $1 - R_i$ over the others, and

$$R_{k\text{-out-of-}n} = \sum_{\substack{J \subseteq \{1,\ldots,n\} \\ |J| \geq k}} \prod_{i \in J} R_i \prod_{i \notin J} (1 - R_i). \tag{9.20}$$

As an example, let the three chillers have reliabilities $R_A = 0.98$, $R_B = 0.95$ and $R_C = 0.90$, and require any two. Table 9.2 lists all $2^3 = 8$ configurations. Each row is one configuration, written as the states (S_A, S_B, S_C); the columns give the number of functioning chillers, the product that gives its probability, the probability, and whether the plant succeeds. The eight probabilities add to one, as they must for a partition of the outcomes.

Table 9.2: The eight configurations of a 2-out-of-3 plant of chillers with reliabilities 0.98, 0.95 and 0.90 and independent failures (equation (9.20)).

(S_A, S_B, S_C)	Working	Product	Probability	Plant
$(1, 1, 1)$	3	$0.98 \times 0.95 \times 0.90$	0.8379	succeeds
$(1, 1, 0)$	2	$0.98 \times 0.95 \times 0.10$	0.0931	succeeds
$(1, 0, 1)$	2	$0.98 \times 0.05 \times 0.90$	0.0441	succeeds
$(0, 1, 1)$	2	$0.02 \times 0.95 \times 0.90$	0.0171	succeeds
$(1, 0, 0)$	1	$0.98 \times 0.05 \times 0.10$	0.0049	fails
$(0, 1, 0)$	1	$0.02 \times 0.95 \times 0.10$	0.0019	fails
$(0, 0, 1)$	1	$0.02 \times 0.05 \times 0.90$	0.0009	fails
$(0, 0, 0)$	0	$0.02 \times 0.05 \times 0.10$	0.0001	fails

The first four rows are the success configurations, and the system reliability is their sum,

$$\begin{aligned} R_{\text{system}} &= R_A R_B R_C + R_A R_B (1 - R_C) + R_A R_C (1 - R_B) + R_B R_C (1 - R_A) \\ &= 0.8379 + 0.0931 + 0.0441 + 0.0171 = 0.9922, \end{aligned}$$

a reliability of 99.22 %. The last four rows add to the failure probability, 0.0078, and their sum is a check on the first. The failure probability is dominated by the outcomes in which exactly two chillers fail. With $q_i = 1 - R_i$, it is approximately the sum of the pairwise products, $q_A q_B + q_A q_C + q_B q_C = 0.001 + 0.002 + 0.005 = 0.008$, and the pair of the two weakest chillers, B and C, contributes almost two thirds of it.

Equation (9.20) has 2^n terms, which becomes computationally expensive for large n: a k-out-of-n group of thirty units has over a billion configurations. Specialized reliability algorithms avoid the enumeration. A standard one conditions on the state of the last component: writing $R_{(k,n)}$ for the reliability of a k-out-of-n group of components $1, \ldots, n$,

$$R_{(k,n)} = R_n \, R_{(k-1,\, n-1)} + (1 - R_n) \, R_{(k,\, n-1)},$$

with $R_{(0,m)} = 1$ and $R_{(k,m)} = 0$ for $k > m$. The recursion, a form of dynamic programming, needs on the order of $k\,n$ operations instead of 2^n.

9.3.7 Worked table of k-out-of-n reliability

Table 9.3 evaluates equation (9.19) for eight structures and three component reliabilities. Each row is one structure, from two to four identical, independent components; the columns give

the system reliability for component reliabilities 0.90, 0.95 and 0.99, and the system failure probability for 0.95. No repair, switching failure or common-cause event is modeled. As one row of the arithmetic, the 3-out-of-4 system with $R = 0.95$ succeeds with three or four working components,

$$R_{\text{3-out-of-4}} = \binom{4}{3}(0.95)^3(0.05) + (0.95)^4 = 4 \times 0.857375 \times 0.05 + 0.814506$$
$$= 0.171475 + 0.814506 = 0.985981.$$

Table 9.3: Reliability of k-out-of-n systems of identical components with independent failures (equation (9.19)), for three component reliabilities R. The last column is the system failure probability at $R = 0.95$.

Structure	Equivalent	$R = 0.90$	$R = 0.95$	$R = 0.99$	$1 - R_{\text{system}}$ at 0.95
1-out-of-2	parallel pair	0.990000	0.997500	0.999900	0.002500
2-out-of-2	series pair	0.810000	0.902500	0.980100	0.097500
1-out-of-3	parallel	0.999000	0.999875	0.999999	0.000125
2-out-of-3	chiller plant	0.972000	0.992750	0.999702	0.007250
3-out-of-3	series	0.729000	0.857375	0.970299	0.142625
2-out-of-4	two spares	0.996300	0.999519	0.999996	0.000481
3-out-of-4	one spare	0.947700	0.985981	0.999408	0.014019
4-out-of-4	series	0.656100	0.814506	0.960596	0.185494

The table shows three trends. Holding $n = 3$ and $R = 0.95$ fixed and raising k from 1 to 3 lowers the system reliability from 0.999875 through 0.99275 to 0.857375. The decrease is a consequence of the stricter functional requirement, not of any deterioration of the components, which are the same in the three rows. Second, at fixed $n - k$, the number of spares, a larger group is less reliable: one spare gives 0.9975 for 1-out-of-2, 0.99275 for 2-out-of-3 and 0.985981 for 3-out-of-4, because more units must survive and more of them can fail. Third, better components help redundant structures more than series ones: from $R = 0.90$ to $R = 0.99$ the failure probability of the 2-out-of-3 plant falls by a factor of 94, from 0.028 to 0.000298, while that of the 3-out-of-3 series system falls by a factor of 9, from 0.271 to 0.0297.

9.4 Active and standby redundancy

Parallel units can provide redundancy in two ways: all running at once, or one running while the other waits. This section derives the mission reliability of both arrangements for components with constant failure rates, compares them, extends the standby result to n units, and shows how an imperfect switch changes the comparison.

9.4.1 Active redundancy

In an actively redundant system, several components operate simultaneously. For example, two pumps may both run, sharing the flow, while either can support the required flow alone if the other fails. Both accumulate operating time and both are exposed to the operating environment from the start of the mission. If each pump has an independent exponential lifetime with failure rate λ, its mission reliability is $\mathrm{e}^{-\lambda t}$, and the probability that at least one pump survives a mission of duration t is, by equation (9.13),

$$R_{\text{active}}(t) = 1 - \left(1 - \mathrm{e}^{-\lambda t}\right)^2 = 2\mathrm{e}^{-\lambda t} - \mathrm{e}^{-2\lambda t}, \tag{9.21}$$

where the second form follows by expanding the square, $1-(1-2\mathrm{e}^{-\lambda t}+\mathrm{e}^{-2\lambda t})$.

For two identical pumps, each with $\lambda = 0.0002\,\mathrm{h}^{-1}$, over a mission of 1000 h, the reliability of one pump is $R = \mathrm{e}^{-0.2} = 0.81873$, and

$$R_{\text{active}}(1000\,\mathrm{h}) = 1-(1-0.81873)^2 = 1-0.03286 = 0.9671.$$

Active redundancy provides a mission reliability of about 96.71 %. The result assumes that either pump can sustain the full load after the other fails, that the load carried by the survivor does not raise its failure rate, and that the system tolerates the loss of one pump without violating its function. The mean time to failure of the pair follows by integrating the reliability function over all time, as in Chapter 5,

$$\mathrm{MTTF}_{\text{active}} = \int_0^\infty (2\mathrm{e}^{-\lambda t}-\mathrm{e}^{-2\lambda t})\,\mathrm{d}t = \frac{2}{\lambda}-\frac{1}{2\lambda} = \frac{3}{2\lambda},$$

which is 7500 h for $\lambda = 0.0002\,\mathrm{h}^{-1}$, half again the 5000 h of one pump.

9.4.2 Cold standby redundancy

Now consider an alternative arrangement of the same two pumps. One pump operates while the second remains inactive. If the operating pump, the primary, fails, the standby pump is started and takes over. This is cold standby redundancy. Figure 9.5 draws it. The left box is the primary pump, operating; the right box is the standby pump, inactive; the box between them below is the switching function that transfers the duty to the standby when the primary fails, and the arrow shows the direction of the transfer.

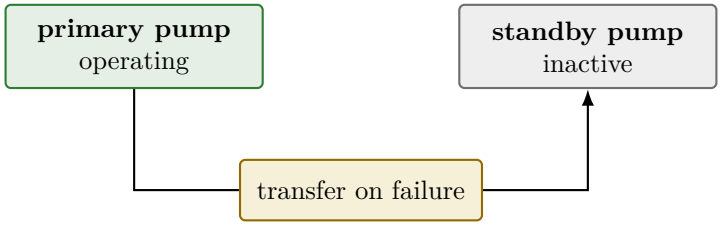


Figure 9.5: Cold standby redundancy of two pumps. The primary operates; the standby is inactive and accumulates no operating time; the switching function transfers the duty to the standby when the primary fails. The ideal model of equation (9.22) assumes that the standby cannot fail while it waits and that the transfer always succeeds.

The mission reliability of this arrangement is derived under the following idealized assumptions:

- the operating pump has a constant failure rate λ;
- the cold standby cannot fail while inactive;
- the standby becomes operational immediately when needed;
- switching always succeeds;
- the standby, once activated, has the same exponential lifetime distribution, with failure rate λ;
- the failures are independent.

The system survives the mission of duration t in two mutually exclusive ways: (1) the primary survives the entire mission; or (2) the primary fails at some time u during the mission, and the standby survives the remaining time $t-u$. The first way has probability

$$P_1 = \mathrm{e}^{-\lambda t}.$$

For the second, the primary's failure time has density $f_1(u) = \lambda \mathrm{e}^{-\lambda u}$, so the probability that it fails in the short interval $[u, u + \mathrm{d}u]$ is $\lambda \mathrm{e}^{-\lambda u}\,\mathrm{d}u$. The standby, started at u as good as new, must then survive from u to t, with probability $R_2(t-u) = \mathrm{e}^{-\lambda(t-u)}$, independent of the primary. By the law of total probability, integrating over all possible failure times of the primary within the mission,

$$P_2 = \int_0^t \lambda \mathrm{e}^{-\lambda u}\, \mathrm{e}^{-\lambda(t-u)}\,\mathrm{d}u = \int_0^t \lambda \mathrm{e}^{-\lambda t}\,\mathrm{d}u = \lambda t\, \mathrm{e}^{-\lambda t},$$

since the exponents add to $-\lambda t$ for every u and the integrand is constant. The reliability of the ideal cold standby pair is the sum,

$$R_{\mathrm{cold}}(t) = P_1 + P_2 = \mathrm{e}^{-\lambda t}(1 + \lambda t), \tag{9.22}$$

and its mean time to failure is $\int_0^\infty \mathrm{e}^{-\lambda t}(1+\lambda t)\,\mathrm{d}t = 1/\lambda + 1/\lambda = 2/\lambda$, which is 10 000 h for $\lambda = 0.0002\,\mathrm{h}^{-1}$: twice the mean life of one pump, because the pair's lifetime is the sum of two pump lifetimes.

9.4.3 Comparison of active and cold standby redundancy

With $\lambda = 0.0002\,\mathrm{h}^{-1}$ and $t = 1000\,\mathrm{h}$, $\lambda t = 0.2$ and

$$R_{\mathrm{cold}}(1000\,\mathrm{h}) = \mathrm{e}^{-0.2}(1 + 0.2) = 0.81873 \times 1.2 = 0.9825.$$

Table 9.4 compares the arrangements. Each row is one arrangement of identical pumps with $\lambda = 0.0002\,\mathrm{h}^{-1}$; the columns give the reliability over the 1000 h mission, the corresponding failure probability, and the mean time to failure. The first three rows are the comparison of this subsection; the rows with a switch reliability R_{sw} and the rows with three pumps are derived in the two subsections that follow.

Table 9.4: Mission reliability over 1000 h and mean time to failure of redundant arrangements of identical pumps with constant failure rate $\lambda = 0.0002\,\mathrm{h}^{-1}$, independent failures and no repair. Standby units cannot fail while inactive.

Arrangement	$R_{\mathrm{system}}(1000\,\mathrm{h})$	Failure probability	MTTF (h)
One pump	0.818731	0.181269	5000
Two pumps, active	0.967141	0.032859	7500
Two pumps, ideal cold standby	0.982477	0.017523	10 000
Two pumps, cold standby, $R_{\mathrm{sw}} = 0.95$	0.974290	0.025710	9750
Two pumps, cold standby, $R_{\mathrm{sw}} = 0.80$	0.949728	0.050272	9000
Three pumps, active	0.994044	0.005956	9167
Three pumps, ideal cold standby	0.998852	0.001148	15 000

Under the idealized assumptions, cold standby provides the higher mission reliability: 0.9825 against 0.9671, a failure probability of 0.0175 against 0.0329, almost half. The reason is that the standby component does not accumulate operating exposure until it is needed. The same conclusion holds for every realization, not only on average: the active pair fails at $\max(T_1, T_2)$, when the second of two simultaneously running pumps fails, while the ideal standby pair fails at $T_1 + T_2$, when the second pump has run its own full life after the first; and $T_1 + T_2 \geq \max(T_1, T_2)$ always. The ideal cold standby therefore dominates active redundancy for every mission length.

Figure 9.6 draws the reliability of the arrangements as functions of the mission length. The horizontal axis is the dimensionless mission length λt, from 0 to 2, so that $\lambda t = 1$ is a mission as

long as the mean life of one pump; the vertical axis is the mission reliability. The curves are one pump, $e^{-\lambda t}$; two active pumps, equation (9.21); the ideal cold standby pair, equation (9.22); and a cold standby pair whose switch succeeds with probability 0.80, equation (9.25) below. The dotted vertical line marks the 1000 h mission of the table, $\lambda t = 0.2$. The ideal standby curve lies above the active one everywhere, and the gap widens with the mission length: at $\lambda t = 1$ the values are 0.736 and 0.600.

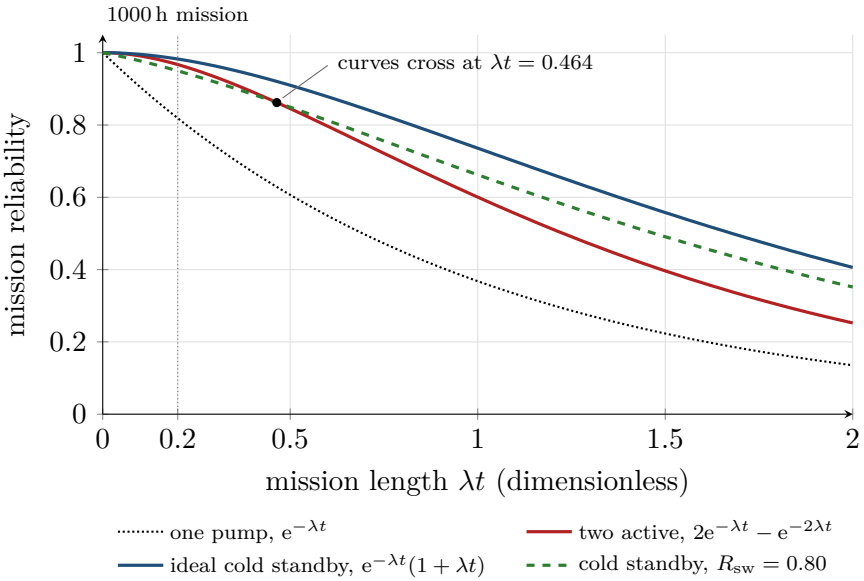


Figure 9.6: Mission reliability of redundant pairs of identical pumps with constant failure rate λ and independent failures, against the dimensionless mission length λt: one pump; two active pumps, equation (9.21); an ideal cold standby pair, equation (9.22); and a cold standby pair with switch reliability 0.80, equation (9.25). The dot marks where the last crosses the active pair.

The advantage of cold standby depends heavily on the idealized switching and dormant-failure assumptions. The dashed curve already shows the effect of a switch that fails one time in five: for short missions it lies below the active pair, which needs no switch, and it rises above it only for missions longer than $\lambda t = 0.464$. The two subsections that follow extend the ideal model to more units and then relax the switching assumption.

9.4.4 Cold standby with n units

Consider n identical units in cold standby: one operates, and when it fails the next is switched in, until all n have failed. With independent exponential lifetimes of the same operating failure rate λ, instantaneous perfect switching and no deterioration in standby, the system lifetime is the sum of the n operating lifetimes, $T_{\text{sys}} = T_1 + \cdots + T_n$. Its reliability is

$$R_{\text{cold},n}(t) = e^{-\lambda t} \sum_{j=0}^{n-1} \frac{(\lambda t)^j}{j!}, \tag{9.23}$$

which for $n = 2$ is equation (9.22).

Equation (9.23) follows by induction on n with the argument of Section 9.4.2. Suppose it holds for n units, and let $f_n(u) = \lambda^n u^{n-1} e^{-\lambda u}/(n-1)!$ be the density of the time at which the n-th unit fails, which is the negative derivative of $R_{\text{cold},n}(u)$ (the terms of the derivative telescope to this single one). A system of $n+1$ units survives to t if the first n units last beyond

t, or if the n-th of them fails at some $u < t$ and the last unit survives from u to t:

$$R_{\mathrm{cold},n+1}(t) = R_{\mathrm{cold},n}(t) + \int_0^t \frac{\lambda^n u^{n-1} \mathrm{e}^{-\lambda u}}{(n-1)!}\, \mathrm{e}^{-\lambda(t-u)}\, \mathrm{d}u = R_{\mathrm{cold},n}(t) + \mathrm{e}^{-\lambda t}\, \frac{(\lambda t)^n}{n!},$$

because the exponentials again combine to $\mathrm{e}^{-\lambda t}$ and $\int_0^t u^{n-1}\,\mathrm{d}u = t^n/n$. The new term is the next term of the sum, which proves the formula for $n+1$; the case $n = 1$ is the exponential reliability $\mathrm{e}^{-\lambda t}$.

Two interpretations of equation (9.23) are standard. The sum of n independent exponential lifetimes with rate λ has the Erlang distribution with shape n and rate λ, a gamma distribution with integer shape, and equation (9.23) is its survival function. Equivalently, while units remain, failures arrive as a Poisson process of rate λ, since exactly one unit operates at a time; the number of failures by time t has the Poisson distribution with mean λt, and the system survives if fewer than n failures occur by time t, which is the sum of the Poisson probabilities for $j = 0, \ldots, n-1$. The mean time to failure is n/λ. For three pumps with $\lambda t = 0.2$,

$$R_{\mathrm{cold},3}(1000\,\mathrm{h}) = \mathrm{e}^{-0.2}\Big(1 + 0.2 + \frac{0.2^2}{2}\Big) = 0.81873 \times 1.22 = 0.998852,$$

against $1 - (1 - 0.81873)^3 = 0.994044$ for three active pumps, Table 9.4. The mean time to failure of n active units is $(1/\lambda)\sum_{j=1}^{n} 1/j$, which grows only as the logarithm of n: 9167 h for three pumps against 15 000 h in cold standby.

9.4.5 Imperfect switching

Real standby systems can fail when called upon: a transfer switch does not close, a start signal is lost, a standby motor does not start. Let the switch reliability be

$$R_{\mathrm{sw}} = \Pr(\text{successful standby activation} \mid \text{primary fails}), \tag{9.24}$$

the probability that the transfer succeeds on demand, and assume that it is constant, independent of the primary's failure time, and that the standby still cannot fail while dormant. In the derivation of Section 9.4.2, the first way of surviving needs no switch, and the second needs the switch to succeed at the moment u, an event independent of everything else with probability R_{sw}. So the second term is multiplied by R_{sw},

$$R_{\mathrm{standby}}(t) = \mathrm{e}^{-\lambda t} + R_{\mathrm{sw}}\, \lambda t\, \mathrm{e}^{-\lambda t},$$

that is,

$$R_{\mathrm{standby}}(t) = \mathrm{e}^{-\lambda t}(1 + R_{\mathrm{sw}}\, \lambda t). \tag{9.25}$$

For $R_{\mathrm{sw}} = 0.95$ and the mission of the previous subsections,

$$R_{\mathrm{standby}}(1000\,\mathrm{h}) = \mathrm{e}^{-0.2}(1 + 0.95 \times 0.2) = 0.81873 \times 1.19 = 0.9743.$$

This remains higher than the active result, 0.9671, but the advantage is smaller: the failure probability has risen from 0.0175 to 0.0257. For $R_{\mathrm{sw}} = 0.80$,

$$R_{\mathrm{standby}}(1000\,\mathrm{h}) = \mathrm{e}^{-0.2}(1.16) = 0.9497,$$

and active redundancy now performs better.

The switch reliability at which the two arrangements are equally reliable follows by setting equation (9.25) equal to equation (9.21) and dividing by $\mathrm{e}^{-\lambda t}$: $1 + R_{\mathrm{sw}}\lambda t = 2 - \mathrm{e}^{-\lambda t}$, so

$$R_{\mathrm{sw}}^{*} = \frac{1 - \mathrm{e}^{-\lambda t}}{\lambda t}. \tag{9.26}$$

For $\lambda t = 0.2$ the break-even switch reliability is $0.18127/0.2 = 0.9063$; a switch less reliable than that makes the standby arrangement worse than two running pumps over this mission. The break-even value falls as the mission lengthens, which is why the curves of Figure 9.6 cross at $\lambda t = 0.464$, the mission length at which $R_{\text{sw}}^{*} = 0.80$. The mean time to failure of the imperfect standby pair is $(1 + R_{\text{sw}})/\lambda$, listed in Table 9.4.

Switching is only the first of the effects that the ideal model omits. A standby unit that can fail while dormant, at a lower rate than in operation, is called a warm standby; a switching delay, imperfect detection of the primary's failure, and failure rates that depend on the load or the condition of the units are further effects. Each requires a more detailed state-based or dynamic reliability model, and later chapters on dependent failures and on Markov models of repairable systems treat them.

9.5 Independence and common-cause failure

Every reliability formula derived so far multiplied probabilities, and multiplication requires independence. This section writes the parallel formula without that assumption, quantifies the effect of a shared supply, and defines the common-cause failures that most often defeat redundancy.

9.5.1 Parallel reliability without independence

The parallel formula $R_{\text{parallel}} = 1 - (1 - R_A)(1 - R_B)$ assumes independence. Without it, only equation (9.12) remains,

$$R_{\text{parallel}} = 1 - \Pr(F_A \cap F_B) = 1 - \Pr(F_A)\ \Pr(F_B \mid F_A),$$

and the difficult quantity is the joint failure probability $\Pr(F_A \cap F_B)$, or equivalently the conditional probability that B fails given that A has failed. Independence sets the conditional probability equal to $\Pr(F_B)$. Two identical pumps with failure probability 0.05 each show the range: if their failures are independent the joint probability is $0.05^2 = 0.0025$; if a failure of one always accompanies a failure of the other, the conditional probability is one and the joint probability is 0.05, so the pair is no more reliable than one pump. Every degree of dependence between these extremes is possible, and component reliabilities alone do not say which holds.

Suppose two cooling pumps share a common electrical feeder. A feeder failure disables both pumps at once. The pumps, considered as providers of the cooling service, then fail together more often than independence predicts, and their failures are no longer adequately modeled as independent outcomes for the complete cooling service. The remedy is not to guess the conditional probability but to model the shared element explicitly, as the next subsection does.

9.5.2 A shared electrical supply

Suppose each pump has a mission reliability of 0.95, the shared electrical supply has a reliability $R_E = 0.98$, the pumps have independent success events conditional on the supply functioning, and the supply's failures are independent of the pump-specific failures. The structure is the one of panel (b) of Figure 9.3, with the supply in place of the controller: the supply is in series with the parallel pair. Then

$$R_{\text{system}} = R_E\left[1 - (1 - R_A)(1 - R_B)\right] = 0.98\left[1 - (0.05)^2\right] = 0.98 \times 0.9975 = 0.97755.$$

Instead of 99.75 %, the system reliability is 97.755 %. The system failure probability, 0.02245, is nine times that of the pump pair, and 0.02 of it, about 89 %, comes from the supply.

The example also shows the error that the shared supply invites. The pumping service of each pump, pump and supply together, has reliability $0.98 \times 0.95 = 0.931$, a failure probability

of 0.069. Treating the two services as independent gives $1 - 0.069^2 = 0.995239$, almost five times too optimistic in failure probability. In truth the probability that B's service is lost given that A's is lost is $0.02245/0.069 = 0.325$, not 0.069: the services are strongly dependent through the supply they share, although the pumps themselves fail independently. Redundancy must therefore be evaluated at the functional system level, with every shared element in the block diagram, and not by multiplying the reliabilities of services that overlap.

Common mistake

Multiplying the failure probabilities of redundant trains whose reliabilities already include a shared supply, controller or cooling source. The product treats the shared element as if each train had its own; the correct model places the shared element once, in series with the redundant group.

9.5.3 Common-cause failure

A common-cause failure occurs when a shared cause or mechanism leads to the failure of several components, within a short time or at once. Examples include:

- two pumps damaged by contaminated cooling water;
- several electrical devices affected by one voltage surge;
- redundant sensors exposed to the same calibration error;
- several servers affected by the same cooling-system outage;
- several transmission lines affected by the same extreme-weather event.

In each case the redundant items fail together for a reason that no item-by-item reliability captures, and common-cause failures can undermine most of the benefit of redundancy: Table 9.1 promised a failure probability of 10^{-4} for four pumps, which a contaminant that damages all four with a probability of 10^{-3} per mission would exceed tenfold.

A simple conceptual model separates the independent failures of each item from shared-cause events, each of which fails several items at once. The shared supply of Section 9.5.2 is a shared-cause event that happens to have a name and a block of its own. More formal methods quantify shared causes that cannot be listed one by one: the beta-factor model, the alpha-factor model, the multiple Greek letter model, and explicit shared-event models. Chapter 13 treats common-cause and dependent failures, and Chapter 14 develops the beta-factor and alpha-factor models.

One modeling warning applies from the start. Assuming a numerical correlation between component failures is not equivalent to constructing a physical common-cause failure model. Correlation describes statistical dependence; identifying a shared cause, a feeder, a water source, a calibration procedure, provides actionable engineering information, such as which barrier, separation or diversity of design would remove the dependence, and it tells the analyst which evidence would measure it.

9.6 Combined structures, component importance and capacity

Real block diagrams mix series and parallel elements. This section reduces such a diagram to its system reliability, asks which component should be improved first and answers with the Birnbaum importance, shows where the binary-state block diagram stops describing a system with unequal capacities, and lists the common mistakes in block diagrams.

9.6.1 Series–parallel structures

Consider a cooling system containing one electrical supply, two redundant pumps, one heat exchanger and one controller. The pumps are in parallel, while the other components are each

required. Figure 9.7 draws the block diagram. From the input on the left, the line passes through the electrical supply (reliability 0.99), splits into the two pump branches (each 0.95), joins again, and passes through the heat exchanger (0.98) and the controller (0.97) to the output. There is one success path through each required stage, and two paths through the pumping stage. The first three stages are the cooling service of Chapter 2; the controller is added here.

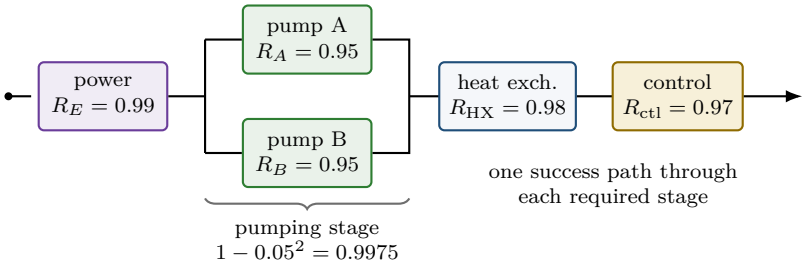


Figure 9.7: Reliability block diagram of a cooling system with one electrical supply, two redundant pumps, one heat exchanger and one controller, with illustrative mission reliabilities and independent failures. The pumps form a parallel stage; every other block is in series. System reliability 0.938741 (equation (9.27)).

A series–parallel diagram is reduced stage by stage. The parallel pump stage is replaced by one block of reliability $1 - (1 - R_A)(1 - R_B)$, and the resulting chain of four blocks is a series system. Assuming independent success events for all blocks,

$$R_{\mathrm{system}} = R_E \left[1 - (1 - R_A)(1 - R_B)\right] R_{\mathrm{HX}}\, R_{\mathrm{ctl}}, \tag{9.27}$$

and substituting,

$$R_{\mathrm{system}} = 0.99 \times 0.9975 \times 0.98 \times 0.97 = 0.938741265.$$

The system reliability is about 93.87 %. Without the controller the product is 0.9677745, the reliability of the complete cooling service of Chapter 2, and the controller multiplies it by 0.97.

Even though pump redundancy provides a reliability of 99.75 % at the pumping stage, the complete system reliability is only about 93.87 %. Apportioning the failure probability among the stages with the rare-event approximation, $1 - R_{\mathrm{system}} \approx \sum(1 - R_{\mathrm{stage}})$, gives $0.01 + 0.0025 + 0.02 + 0.03 = 0.0625$, against the exact 0.0613: the controller contributes about half, the heat exchanger a third, the supply a sixth, and the redundant pumps one twenty-fifth. This is a common engineering observation: improving one highly reliable subsystem provides relatively little benefit when other required components remain significant sources of failure.

9.6.2 Birnbaum importance

The question of which component to improve first is answered by an importance measure. With independent components, the system reliability is the function h of equation (9.5). The Birnbaum importance of component i is its partial derivative,

$$I_i^{\mathrm{B}} = \frac{\partial R_{\mathrm{system}}}{\partial R_i} = \frac{\partial h(R_1, \ldots, R_n)}{\partial R_i}. \tag{9.28}$$

It measures the local sensitivity of the system reliability to a change in the reliability of component i, holding the others fixed. It has a second reading. Conditioning on the state of component i gives $R_{\mathrm{system}} = R_i\, h(1_i) + (1 - R_i)\, h(0_i)$, where $h(1_i)$ and $h(0_i)$ are the system reliabilities with component i working and failed and the other components unchanged. The function h is therefore linear in each R_i separately, and

$$I_i^{\mathrm{B}} = h(1_i) - h(0_i), \tag{9.29}$$

the probability that component i is critical, that is, that the system works with component i working and fails with it failed. Because h is linear in R_i, the derivative is exact for finite changes: raising R_i by Δ raises R_{system} by exactly $I_i^{\text{B}}\Delta$.

For the system of equation (9.27), write the pump stage as $R_A + R_B - R_A R_B$, so that

$$R_{\text{system}} = R_E \, R_{\text{HX}} \, R_{\text{ctl}} \, (R_A + R_B - R_A R_B).$$

The importance of pump A is

$$I_A^{\text{B}} = \frac{\partial R_{\text{system}}}{\partial R_A} = R_E \, R_{\text{HX}} \, R_{\text{ctl}} \, (1 - R_B) = 0.99 \times 0.98 \times 0.97 \times 0.05 = 0.0471,$$

and that of the electrical supply is

$$I_E^{\text{B}} = \frac{\partial R_{\text{system}}}{\partial R_E} = R_{\text{HX}} \, R_{\text{ctl}} \, (R_A + R_B - R_A R_B) = 0.98 \times 0.97 \times 0.9975 = 0.9482.$$

For any component in series with the rest, the derivative is the product of all the other factors, $I_i^{\text{B}} = R_{\text{system}}/R_i$. Table 9.5 lists all five components. Each row is one component; the columns give its reliability, its failure probability $q_i = 1 - R_i$, its Birnbaum importance, the system reliability if the component were made perfect, and the gain that this would bring, $I_i^{\text{B}} q_i$ by the linearity of h.

Table 9.5: Birnbaum importance (equation (9.28)) of the components of Figure 9.7, $R_{\text{system}} = 0.938741$. The gain is the increase of R_{system} if the component were made perfect, $I_i^{\text{B}} q_i$.

Component	R_i	q_i	I_i^{B}	R_{system} if perfect	Gain	Rank
Controller	0.97	0.03	0.9678	0.967775	0.029033	1
Heat exchanger	0.98	0.02	0.9579	0.957899	0.019158	2
Electrical supply	0.99	0.01	0.9482	0.948224	0.009482	3
Pump A	0.95	0.05	0.0471	0.941094	0.002353	4
Pump B	0.95	0.05	0.0471	0.941094	0.002353	4

The electrical supply has a local importance twenty times that of a pump, 0.9482 against 0.0471. Improving one pump has limited benefit because the other pump already provides redundancy: a pump matters only in the outcomes in which the other pump has failed, which have probability 0.05. Improving the single required power supply directly improves the reliability of the entire system. The table also shows that the supply is not the most important component. Among the series components, $I_i^{\text{B}} = R_{\text{system}}/R_i$ is largest for the least reliable one, and the controller, at 0.97, ranks first, followed by the heat exchanger; making the controller perfect would raise the system reliability by 0.029, twelve times what a perfect pump would give.

9.6.3 Importance and failure probability

The pumps are the components most likely to fail, with $q_i = 0.05$, and they are the least important. A component can fail relatively frequently yet have low system importance because another component provides backup. Conversely, a highly reliable component can have high importance because there is no alternative functional pathway: the electrical supply, the most reliable block in the diagram with $q_E = 0.01$, has twenty times the Birnbaum importance of a pump. Component reliability and contribution to system reliability are different quantities, and ranking components by their failure probabilities alone ranks the pumps first and the controller second, the wrong order for deciding where to spend.

The Birnbaum importance is local: it describes the effect of small or linear changes of one component with the rest fixed, and it does not depend on the reliability of the component itself. Other measures answer other questions: the Fussell–Vesely importance I_i^{FV}, the fraction of the system failure probability that involves the component; the risk achievement worth RAW_i, how much worse the system becomes if the component is certainly failed; and the risk reduction worth RRW_i, how much better it becomes if the component is made perfect. Chapter 12 develops the four measures systematically.

Principle 9.1. *Component importance depends on system structure: the component most likely to fail is not necessarily the one whose improvement gives the greatest system benefit, and a single required element of high reliability can outrank redundant elements of low reliability.*

9.6.4 Capacity and the binary-state block diagram

A block diagram treats each component as working or failed and the system as succeeding when the right combination works. That is enough when each surviving combination delivers the same performance, and not otherwise. Consider three chillers of unequal capacity,

$$C_1 = 500\,\mathrm{kW}, \qquad C_2 = 500\,\mathrm{kW}, \qquad C_3 = 300\,\mathrm{kW},$$

and a required cooling demand $L = 800\,\mathrm{kW}$ (illustrative). A conventional 2-out-of-3 interpretation is not quite sufficient, because two functioning chillers do not always provide the same total capacity: chillers 1 and 2 provide 1000 kW, chillers 1 and 3 provide 800 kW, and chillers 2 and 3 provide 800 kW. In this example any two are just sufficient, and the 2-out-of-3 diagram happens to be right. If the demand increased to 900 kW, only the combination of chillers 1 and 2 would satisfy the requirement, and the system would be a series system of those two chillers, with chiller 3 contributing nothing.

The success event is in general a capacity condition,

$$\{\phi(S) = 1\} = \Big\{\sum_{i=1}^{3} C_i\, S_i \geq L\Big\}, \tag{9.30}$$

which is the available capacity $C_{\mathrm{available}}$ of Chapter 1 compared with the load. This is a capacity-based reliability model. Table 9.6 shows how the success logic it implies changes with the demand. Each row is a range of demand L; the columns give the minimal combinations of working chillers that meet it, the block-diagram structure those combinations form, and the system reliability with each chiller of reliability 0.95 and independent failures (illustrative). For a demand between 800 and 1000 kW, for example, chillers 1 and 2 must both work, and $R_{\mathrm{system}} = 0.95^2 = 0.9025$.

The table shows that a 100 kW increase in demand, from 800 to 900 kW, changes the structure from 2-out-of-3 to a series pair and multiplies the failure probability by more than thirteen, from 0.00725 to 0.0975, with no change in any component. If the capacities C_i and the load L are uncertain or vary with time, as the chiller capacities and the cooling load do with the outdoor temperature, the model generalizes to

$$\Pr\Big(\sum_{i} C_i(t)\, S_i(t) \geq L(t)\Big) \tag{9.31}$$

for adequacy at one instant, a snapshot, or to the requirement that the inequality hold throughout a specified mission. This connects directly to the performance-margin and limit-state formulation of Chapters 1 and 2, with margin $g = C_{\mathrm{available}}(t) - L(t)$. It is especially important in power systems, data centers, cooling plants and battery systems, where units of different sizes share a load that changes, and a later chapter on power-system adequacy builds on it.

Table 9.6: Success logic of three chillers of capacities 500, 500 and 300 kW against the cooling demand L (equation (9.30)), and the system reliability with each chiller of reliability 0.95 and independent failures.

Demand L (kW)	Minimal working combinations	Equivalent structure	R_{system}
$L \leq 300$	any one chiller	1-out-of-3	0.999875
$300 < L \leq 500$	chiller 1 or chiller 2	1 and 2 in parallel	0.997500
$500 < L \leq 800$	any two chillers	2-out-of-3	0.992750
$800 < L \leq 1000$	chillers 1 and 2	1 and 2 in series	0.902500
$1000 < L \leq 1300$	all three	3-out-of-3	0.857375
$L > 1300$	none	never succeeds	0

9.6.5 Common mistakes in reliability block diagrams

The assumptions behind the formulas of this chapter are easy to lose in practice, and seven errors recur. Table 9.7 lists them. Each row is one mistake; the middle column states why it matters, and the right column the subsection that treats the point.

Table 9.7: Common mistakes in reliability block diagrams and why each matters.

Mistake	Why it matters	See
Drawing an RBD from physical piping or wiring alone	an RBD represents functional success logic, not physical connections	9.1.1, 9.3.4
Multiplying reliabilities without checking independence	shared causes and dependencies invalidate simple products	9.5.1, 9.5.2
Assuming parallel equipment always provides full redundancy	each alternative must satisfy the required function alone	9.3.1
Ignoring controllers, switching and supporting equipment	shared required elements can dominate the system failure probability	9.3.3, 9.6.1
Treating standby and active redundancy as equivalent	standby exposure and switching behavior change the reliability	9.4.3, 9.4.5
Using k-out-of-n logic for units of unequal capacity without checking the requirement	the correct success event may depend on the total available capacity	9.6.4
Assuming high component reliability guarantees high system reliability	system architecture and critical dependencies matter	9.2.2, 9.6.2

The first and the last rows bracket the chapter. A block diagram is a statement about the required function, and a system reliability computed from it is as credible as that statement and as the independence assumptions under it; a component reliability, however high, says little about the system until the structure and the dependencies are known. Reliability block diagrams describe the combinations of components that make a system succeed; the next chapter, Chapter 10, approaches the same problem from the opposite direction, starting from an undesired system event, such as the loss of data center cooling, and constructing the combinations of component failures and shared dependencies that cause it with AND and OR gates.

9.7 What to retain

Definition 9.1 (System reliability from a block diagram)**.** The reliability of a system is the probability that its structure function equals one, $R_{\text{system}}(t) = \Pr(\phi(S(t)) = 1)$: the structure function, drawn as a reliability block diagram from the required function, says which combinations of working components succeed, and the probability model says how likely they are; redundancy improves the result only when each redundant unit can carry the function, the units fail independently, and any switching they need succeeds.

Table 9.8 lists the chapter's terms with their meaning, their units and the subsection that defines each.

Table 9.8: The terms of the chapter and where each is defined. Formulas assume independent components; exponential formulas assume a constant failure rate λ.

Term	Meaning	Units	Defined in
Reliability block diagram	blocks of success events between input and output; the system works if a path of working blocks joins them	–	9.1.1
State variable $S_i(t)$	1 if component i performs its function throughout the mission, else 0; $\mathbf{1}\{T_i > t\}$	–	9.1.2
Structure function ϕ	Boolean map from component states to the system state	–	9.1.2
Path set, minimal path set P_j	components whose functioning ensures system functioning; with no superfluous member	–	9.1.2
Series system	all components required; $R_{\text{system}} = \prod_i R_i$, $\lambda_{\text{sys}} = \sum_i \lambda_i$	–, h^{-1}	9.2.1
Parallel system	any one component suffices; $R_{\text{system}} = 1 - \prod_i (1 - R_i)$	–	9.3.1
Single point of failure	component whose failure alone fails the system	–	9.3.3
k-out-of-n system	at least k of n required; $\sum_{j=k}^{n} \binom{n}{j} R^j (1-R)^{n-j}$	–	9.3.5
Active redundancy	redundant units run together; two units: $2\text{e}^{-\lambda t} - \text{e}^{-2\lambda t}$	–	9.4.1
Cold standby redundancy	spare inactive until switched in; two units: $\text{e}^{-\lambda t}(1 + \lambda t)$; n units: Erlang survival function	–	9.4.2, 9.4.4
Switch reliability R_{sw}	probability that the standby is activated on demand; two units: $\text{e}^{-\lambda t}(1 + R_{\text{sw}}\lambda t)$	–	9.4.5
Warm standby	standby that can fail while dormant, at a reduced rate	–	9.4.5
Common-cause failure	failure of several components from one shared cause	–	9.5.3
Birnbaum importance I_i^{B}	$\partial R_{\text{system}} / \partial R_i = h(1_i) - h(0_i)$	–	9.6.2
Capacity-based success criterion	$\sum_i C_i S_i \geq L$ in place of a count of working units	kW	9.6.4

Further reading. Rausand, Barros and Høyland, *System Reliability Theory*, 3rd ed., for reliability block diagrams, coherent systems, k-out-of-n systems and component importance; O'Connor and Kleyner, *Practical Reliability Engineering*, 5th ed., for redundancy, reliability configurations and engineering applications; Modarres, Kaminskiy and Krivtsov, *Reliability Engineering and Risk Analysis*, 3rd ed., for system reliability modeling and dependent failures; and Billinton and Allan, *Reliability Evaluation of Power Systems*, 2nd ed., for network reliability, capacity-based system success and power-system applications.

9.8 Exercises

Exercise 9.1. (a) Three independent components each have reliability 0.90, and all three are required. Compute the system reliability. (b) Two independent pumps each have reliability 0.95, and either can meet the full demand. Compute the system reliability. (c) In arrangement (b), both pumps need one controller of reliability 0.98; compute the system reliability and the fraction of its failure probability that the controller causes. (Answers: (a) $0.9^3 = 0.729$, since a series system multiplies the reliabilities; (b) $1 - 0.05^2 = 0.9975$, since a parallel system fails only if both fail; (c) $0.98 \times 0.9975 = 0.97755$, and the controller causes $0.02/0.02245 = 89\,\%$ of the failure probability.)

Exercise 9.2. A system has four identical, independent components and requires at least three to work. (a) Which expression gives its reliability: R^4, $4R^3(1-R) + R^4$, or $1 - (1-R)^4$? (b) Evaluate it for $R = 0.90$ and compare with a 2-out-of-3 system of the same components. (Answers: (a) $4R^3(1-R) + R^4$, because the 3-out-of-4 system succeeds when exactly three or all four components succeed; (b) $0.2916 + 0.6561 = 0.9477$, against 0.972 for 2-out-of-3: with one spare in each, the larger group is less reliable.)

Exercise 9.3. (a) Which statement about cold standby is correct: it always outperforms active redundancy; its reliability depends on dormant failures and switching success; or the standby must have the same operating exposure as the primary? (b) Two identical pumps have constant failure rate $\lambda = 0.0005\,\mathrm{h}^{-1}$, and the mission lasts 1000 h. Compute the mission reliability of one pump, of the two in active redundancy, of the two in ideal cold standby, and the switch reliability below which the active pair is the more reliable. (Answers: (a) its reliability depends on dormant failures and switching success, since standby performance depends on activation, dormant deterioration and the failure processes; (b) with $\lambda t = 0.5$: $\mathrm{e}^{-0.5} = 0.6065$; $2(0.6065) - \mathrm{e}^{-1} = 0.8452$; $0.6065 \times 1.5 = 0.9098$; and, by equation (9.26), $(1 - 0.6065)/0.5 = 0.787$.)

Exercise 9.4. (a) What does the Birnbaum importance measure: the local sensitivity of system reliability to a component's reliability, the total number of failures in a system, or the cost of maintaining a component? (b) A cooling plant consists of an electrical supply of reliability 0.99 in series with a 2-out-of-3 group of chillers A, B and C, each of reliability 0.95. Compute the system reliability, the Birnbaum importance of the supply and of chiller A, and say which to improve first. (Answers: (a) the local sensitivity of system reliability to the component's reliability, the partial derivative of equation (9.28) under the specified model; (b) $R_{\mathrm{system}} = 0.99 \times 0.99275 = 0.98282$; $I_E^{\mathrm{B}} = 0.99275$; writing the 2-out-of-3 group as $R_AR_B + R_AR_C + R_BR_C - 2R_AR_BR_C$, $I_A^{\mathrm{B}} = 0.99\,(R_B + R_C - 2R_BR_C) = 0.99 \times 0.095 = 0.0941$, so the supply, more than ten times as important although five times less likely to fail, comes first.)

Chapter 10

Fault Tree Construction and Logic

10.1 Fault tree analysis and the top event

Chapter 9 represented a system by its success logic: functioning components were arranged in series, in parallel and in k-out-of-n structures, and the combinations of functioning components that deliver the required function determined the system reliability. Fault tree analysis reverses that perspective. It starts from one undesired outcome of the system, the top event, and works backward to the events that can cause it, then to the events that can cause those, until it reaches events that are not decomposed further. This chapter asks

how the combinations of component failures, operational events and shared dependencies that cause an engineered system to fail can be identified systematically, and how the probability of the failure follows from that logic.

When a data center loses cooling, the possible causes include failed chillers, a loss of pumping, unavailable electrical power and incorrect control. A fault tree connects such causes through logical relations and, when the probabilities of the elementary events are known, gives the probability of the undesired outcome. The chapter first defines the top event and the symbols of a fault tree; it then states the Boolean logic of the AND and OR gates and separates that logic from the probability model used to quantify it; builds and quantifies a complete fault tree of the cooling-water system of a data center; reduces fault-tree expressions with Boolean algebra; distinguishes coherent from noncoherent logic and introduces further gates; and ends with the practice of construction and an application to a transmission network.

Jargon: fault tree analysis, top event, intermediate event, basic event, undeveloped event, transfer, gate, Boolean expression, single point of failure, coherent fault tree, static and dynamic fault trees

Fault tree analysis (FTA) is a deductive, top-down method that represents the combinations of events producing one defined undesired system event as a logic diagram, the *fault tree*, and quantifies its probability. The *top event* is that undesired event, stated precisely. An *intermediate event* is an event in the tree that is explained further by a gate below it. A *basic event* is an event at which the decomposition stops, because the chosen resolution of the analysis has been reached and its probability can be estimated; an *undeveloped event* is one that could be decomposed but is not, for lack of information or because it is judged unimportant. A *transfer* symbol indicates that a branch of the tree is continued on another page or in another tree. A *gate* is a logical relation that produces its output event from its input events: an AND gate when all inputs are required, an OR gate when any one input suffices. The *Boolean expression* of the top event writes the tree as a formula in the basic events. A *single point of failure* is a basic event that by itself causes the top event. A *coherent* fault tree is one in which further failures can never remove the top event once it has occurred. A *static* fault tree describes the combination of event states, without their order or timing; a *dynamic* fault tree adds gates that depend on the sequence of events.

Key idea

A fault tree states the structure of failure: which combinations of events produce the top event. The probability model states how likely those events are and how they depend on one another. The two are distinct modeling responsibilities, and a correct calculation on the second cannot compensate for an error in the first.

10.1.1 The purpose of fault tree analysis

Consider a data center whose cooling-water system has two redundant pumps. Each pump has a high reliability, and either pump alone can supply the required flow, yet the cooling system can still fail for several reasons:

- both pumps become unavailable;
- the common electrical feeder of the pumps fails;
- the discharge manifold shared by the pumps becomes blocked;
- the shared pump controller fails.

A reliability block diagram can represent the success logic of this system, but a fault tree is often better suited to identifying the undesired events explicitly and to listing the combinations in which they act. The fault tree is drawn in the language of failures, which is also the language of the evidence: incident reports, maintenance records and alarms describe what failed, not what succeeded.

Fault trees are especially useful for the analysis of rare events that result from several interacting failures, for which there is little or no direct operating experience of the system event itself. Typical top events are the loss of electrical supply to a load, the failure of emergency cooling, the loss of containment of a hazardous material, the failure of a protective system on demand, and unserved electrical demand in a power network. The method was developed for such problems in the aerospace and nuclear industries and is documented in their standard fault-tree handbooks, named in the further reading at the end of the chapter. A fault tree models the failure logic of a defined system for one defined top event; it is not a catalogue of every failure of the system, and a different top event, such as a spurious start of a pump instead of a loss of flow, needs a tree of its own.

10.1.2 Definition of the top event

Every fault tree begins with a precisely defined top event, because every gate below it is a statement about what causes that event and no other. Consider again the cooling system of a data center. A vague top event would be "cooling system failure". That statement leaves open which equipment is included, what level of cooling counts as sufficient, how long a shortfall must last to count, and over which period the event is assessed. A better definition is: *the cooling system cannot deliver at least* 800 kW *of heat removal for longer than five consecutive minutes during the specified 24-hour mission.* Table 10.1 lists the elements that this definition fixes. Each row names one element and its value; together they are the system boundary and the required function of Chapter 2, written for one undesired event.

Table 10.1: Elements of a precisely defined top event for the cooling system of a data center (illustrative values).

Element	Definition
System	Data center cooling infrastructure
Required function	Remove at least 800 kW of heat
Failure threshold	Cooling capacity below 800 kW
Allowable interruption	Five consecutive minutes
Mission	24 hours
Top event	The requirement is violated

The time criterion matters. With the available cooling capacity $C_{\text{available}}(t)$ of Chapter 1, the sum of the capacities of the available units, the top event of the definition is

$$F_{\text{top}} = \big\{\exists\, t_0 \in [0, H-\tau] :\ C_{\text{available}}(t) < 800\,\text{kW} \text{ for all } t \in [t_0, t_0+\tau]\big\}, \tag{10.1}$$

where F_{top} denotes the top event, $H = 24\,\text{h}$ the mission and $\tau = 5\,\text{min}$ the allowable interruption. Whether this event occurs depends on when units fail, how long switching to a standby unit takes, how quickly a failed unit is repaired, and how the thermal inertia of the chilled-water loop and the room delays the effect of a capacity shortfall. A static fault tree, which combines equipment states without their timing, may not be sufficient to represent it.

The first examples of this chapter therefore simplify the top event to a static snapshot condition,

$$F_{\text{top}} = \{C_{\text{available}} < 800\,\text{kW}\}, \tag{10.2}$$

evaluated on the states of the equipment at one time, or on the states reached by the end of the mission for items that are not repaired during it. For the chiller plant of Chapter 1, whose units deliver about 430 kW each, the snapshot event occurs when fewer than two of the three chillers are available. Later chapters on repairable systems and Markov models treat the full temporal condition of equation (10.1). The probability of the top event is written $Q_{\text{top}} = \Pr(F_{\text{top}})$.

Principle 10.1. *A fault tree is no more precise than its top event. The top event must state the system, the required function, the failure threshold, the allowable interruption and the mission before any gate is drawn.*

10.1.3 Events, gates and their symbols

A conventional fault tree uses graphical symbols to describe events and their logical relations, and the symbols separate two kinds of object. An event describes something that may happen: a pump fails, a feeder is de-energized, the cooling capacity falls below the requirement. A gate

describes how its input events combine to produce another event, its output. The distinction is essential, because the same basic events can produce very different system probabilities depending on the gates that combine them: two pump failures joined by an AND gate give a rare system event, the same two joined by an OR gate a frequent one.

Figure 10.1 shows the standard symbols. The upper row holds the event symbols: a rectangle for the top event and for every intermediate event, an event that a gate below it explains; a circle for a basic event, not developed further in the analysis; and a diamond for an undeveloped event, which could be developed but is not, for example the loss of the utility supply when the utility is outside the system boundary. The lower row holds the two basic gates and the transfer symbols. The AND gate has a flat input side and a rounded output; the OR gate has a curved input side and a pointed output. In this book the output of a gate points up, toward the event it produces, and its inputs enter from below. The triangles are transfers: a transfer-in triangle, with its line at the apex, attaches a subtree drawn elsewhere, and a transfer-out triangle, with its line at the side, marks a subtree that is used elsewhere. A large tree uses transfers to split itself over pages and to reuse a subtree, such as the loss of a power supply shared by several systems, without drawing it twice.

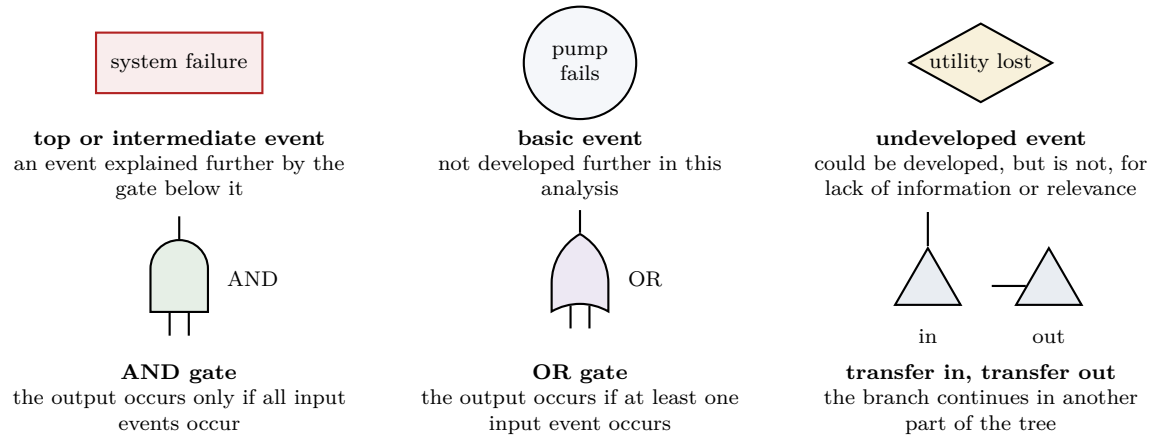


Figure 10.1: Standard fault-tree symbols. Upper row: events (top or intermediate event, basic event, undeveloped event). Lower row: the AND and OR gates, drawn with their output upward and their inputs below, and the transfer-in and transfer-out triangles.

10.2 The logic of AND and OR gates

The AND gate and the OR gate are sufficient for most static fault trees. This section states the event each gate produces, its truth table and its probability for independent inputs, and then separates the logic of a gate, which holds always, from the probability model used to quantify it, which holds only under its assumptions.

10.2.1 The AND gate

Suppose that two redundant pumps must both fail for the system to lose its cooling-water flow. Let

$$A = \{\text{pump A fails}\}, \qquad B = \{\text{pump B fails}\},$$

both over the same mission. Throughout this chapter, capital letters A, B, C, ... name events; in particular C and S are events here, not the consequence, capacity or unit status of the notation chapter. The top event is the intersection of the two events,

$$F_{\text{top}} = A \cap B, \tag{10.3}$$

which Boolean notation writes as a product,

$$F_{\text{top}} = AB. \tag{10.4}$$

This is an AND relation. Figure 10.2 draws it as a fault tree: the rectangle at the top is the top event, the loss of the pumping function; the AND gate below it has two inputs; and the two circles are the basic events "pump A fails" and "pump B fails". The tree is read from the top down as "the loss of pumping occurs if pump A fails and pump B fails". For nonrepairable pumps over a mission, $A \cap B$ means that both have failed by the end of the mission; with repair it would mean only that both failed at some time, which is not the same as both being failed at the same time, a distinction that the static tree cannot make and that the later chapters on repairable systems resolve.

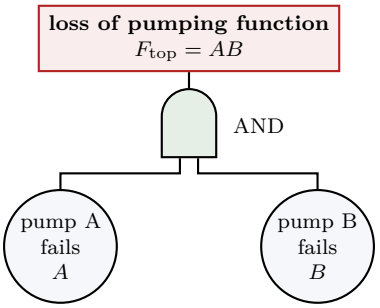


Figure 10.2: An AND gate: the loss of the pumping function occurs only if both pump A and pump B fail, $F_{\text{top}} = A \cap B = AB$.

The logic of a gate can be listed exhaustively in a truth table. Let the indicator x_A equal 1 if event A occurs and 0 if it does not, and similarly x_B and x_{top}. Table 10.2 lists the four combinations of the two inputs and the output of the AND gate for each: the system fails only when both pumps fail. In indicator form the AND gate is the product, $x_{\text{top}} = x_A x_B = \min(x_A, x_B)$, which is why Boolean notation writes it as a product.

Table 10.2: Truth table of the AND gate of Figure 10.2 (1 = the event occurs).

Pump A fails, x_A	Pump B fails, x_B	System fails, x_{top}
0	0	0
0	1	0
1	0	0
1	1	1

Probability of the AND gate. The exact probability of the top event is the probability of the intersection, $Q_{\text{top}} = \Pr(A \cap B)$. Write $q_A = \Pr(A)$ and $q_B = \Pr(B)$ for the basic-event probabilities. If A and B are independent, the probability of their intersection is by definition the product of their probabilities,

$$Q_{\text{top}} = \Pr(A)\,\Pr(B) = q_A\, q_B, \tag{10.5}$$

and for n mutually independent inputs $Q_{\text{top}} = \prod_{i=1}^{n} q_i$. With $q_A = q_B = 0.05$ (illustrative),

$$Q_{\text{top}} = 0.05 \times 0.05 = 0.0025,$$

a top-event probability of 0.25 %. Independence is a probabilistic assumption about the events, not a property of the AND gate. The Boolean expression $A \cap B$ remains valid when the events are dependent; only equation (10.5) does not.

10.2.2 The OR gate

Suppose now that a cooling system has one pump and one electrical supply, and that the failure of either causes the loss of the cooling-water flow. Define

$$A = \{\text{pump fails}\}, \qquad B = \{\text{electrical supply fails}\}.$$

The top event is the union of the two events,

$$F_{\text{top}} = A \cup B, \tag{10.6}$$

which Boolean notation writes as a sum,

$$F_{\text{top}} = A + B, \tag{10.7}$$

where + denotes the logical OR, not arithmetic addition. Figure 10.3 draws the tree, with the same layout as Figure 10.2: the top event "loss of cooling flow", an OR gate, and the basic events "pump fails" and "power fails". It is read as "the loss of cooling flow occurs if the pump fails or the power fails". Table 10.3 is its truth table: the output is 0 only when neither input occurs. In indicator form, $x_{\text{top}} = \max(x_A, x_B) = 1 - (1 - x_A)(1 - x_B)$.

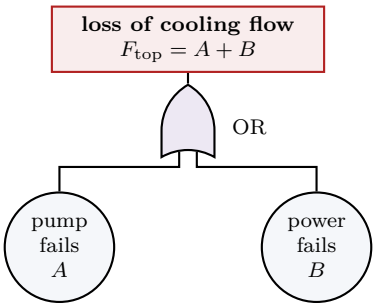


Figure 10.3: An OR gate: the loss of cooling flow occurs if the pump fails or the electrical supply fails, $F_{\text{top}} = A \cup B = A + B$.

Table 10.3: Truth table of the OR gate of Figure 10.3 (1 = the event occurs).

Pump fails, x_A	Power fails, x_B	System fails, x_{top}
0	0	0
0	1	1
1	0	1
1	1	1

Probability of the OR gate. For any two events, the union splits into two disjoint parts, $A \cup B = A \cup (B \cap \overline{A})$, where $\overline{A}$ is the complement of A, the event that A does not occur. The probability of $B \cap \overline{A}$ is $\Pr(B) - \Pr(A \cap B)$, because B is itself the disjoint union of $B \cap A$ and $B \cap \overline{A}$. Adding the probabilities of the two disjoint parts gives the inclusion–exclusion formula,

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B). \tag{10.8}$$

The intersection is subtracted because the outcomes in which both events occur would otherwise be counted twice. Under independence,

$$Q_{\text{top}} = q_A + q_B - q_A q_B = 1 - (1 - q_A)(1 - q_B). \tag{10.9}$$

The second form follows from the complement: the top event fails to occur only if neither input occurs, the complements of independent events are independent, and so $\Pr(\overline{A} \cap \overline{B}) = (1 - q_A)(1 - q_B)$. For n mutually independent inputs it generalizes to $Q_{\text{top}} = 1 - \prod_{i=1}^{n}(1 - q_i)$.

With $q_A = 0.05$ and $q_B = 0.02$ (illustrative),

$$Q_{\text{top}} = 0.05 + 0.02 - 0.05 \times 0.02 = 0.07 - 0.001 = 0.069,$$

and the complement form gives the same, $1 - 0.95 \times 0.98 = 1 - 0.931 = 0.069$. The system has a failure probability of 6.9 % over the mission common to both events. The simple sum $q_A + q_B = 0.07$ is an upper bound for the probability of any union, whatever the dependence, and is close to the exact value when the probabilities are small; this rare-event approximation is developed in Chapter 11.

10.2.3 Logic and probability

The distinction between the logic of a gate and the probability of its output is the most important point of this section. An AND gate always means $F = A \cap B$, and an OR gate always means $F = A \cup B$. These are logical relations between events, and they hold whatever the statistical relation between the events. The product

$$\Pr(A \cap B) = \Pr(A)\ \Pr(B),$$

in contrast, requires independence. For dependent events, the definition of conditional probability gives the general form

$$\Pr(A \cap B) = \Pr(A)\ \Pr(B \mid A), \tag{10.10}$$

where $\Pr(B \mid A)$ is the probability of B given that A has occurred.

The failure of pump A, for example, may increase the load on pump B, which then runs alone at a higher duty, so that a subsequent failure of B becomes more likely:

$$\Pr(B \mid A) > \Pr(B).$$

A calculation that assumes independence then underestimates the probability of losing both pumps. If, for illustration, $\Pr(B \mid A) = 0.30$ instead of $q_B = 0.05$, the probability of losing both pumps is $0.05 \times 0.30 = 0.015$, six times the independent value of 0.0025. The same dependence lowers the probability of an OR gate over the same events, since by equation (10.8) a larger intersection is subtracted: $0.05 + 0.05 - 0.015 = 0.085$ against 0.0975 for independent pumps. Whatever the dependence, the joint probability must lie within the bounds

$$\max(0,\ q_A + q_B - 1) \le \Pr(A \cap B) \le \min(q_A,\ q_B), \tag{10.11}$$

since $A \cap B$ is contained in each of A and B, and $\Pr(A \cap B) = q_A + q_B - \Pr(A \cup B) \ge q_A + q_B - 1$; a specified conditional probability that violates them describes no possible pair of events.

Common mistake

Reading a fault tree's gates as probability formulas. The tree can have correct Boolean logic while its probability model is wrong: an AND gate over two pumps that share a load, a feeder or a maintenance crew is still $A \cap B$, but its probability is $\Pr(A)\Pr(B \mid A)$, not $\Pr(A)\Pr(B)$. Independence must be justified separately for every gate, and it is one of the most important limitations to recognize in a quantitative fault-tree analysis.

Shared causes of this kind, in which one condition raises the probability of several failures at once, are the subject of the later chapters on dependent and common-cause failures, Chapters 13 and 14.

10.3 Worked example: loss of cooling-water flow

This section builds a complete fault tree for a simplified cooling-water system of a data center, from the system description to the Boolean expression of the top event, the tree and its quantification, and shows what the structure of the tree reveals before any probability is assigned.

10.3.1 The system and its Boolean expression

The cooling-water system contains

- two redundant pumps, A and B, either of which can provide the required flow;
- one common electrical feeder that supplies both pumps;
- one shared pump controller;
- one common discharge manifold through which both pumps deliver.

The top event is

$$F_{\text{top}} = \{\text{loss of the required cooling-water flow}\},$$

and the basic events are

$$\begin{aligned} A &: \text{pump A unavailable},\\ B &: \text{pump B unavailable},\\ E &: \text{common electrical feeder unavailable},\\ C &: \text{shared controller fails},\\ M &: \text{discharge manifold blocked}. \end{aligned}$$

The pump-specific events A and B are defined to exclude the shared causes: pump A is unavailable through its own motor, bearings, impeller or local breaker, not because the common feeder or the controller has failed. Defining the events this way avoids counting a shared cause twice, once inside a pump event and once as an event of its own, and it is what makes the events A, B, E, C and M distinct.

The cooling function is lost if

1. both pumps are unavailable, or
2. the common electrical feeder fails, or
3. the controller fails, or
4. the discharge manifold becomes blocked.

The top event is therefore

$$F_{\text{top}} = (A \cap B) \cup E \cup C \cup M, \tag{10.12}$$

or, in Boolean notation,

$$F_{\text{top}} = AB + E + C + M. \tag{10.13}$$

10.3.2 The structure of the tree

Figure 10.4 draws equation (10.13) as a fault tree. The top event is the rectangle at the top. Directly below it, an OR gate with four inputs combines the four ways in which the function is lost. The leftmost input is the intermediate event "both pumps unavailable" (a rectangle), which an AND gate below it explains by the basic events A and B. The other three inputs are the basic events E, C and M, connected to the OR gate directly. The braces below the basic events mark the two kinds of contributor.

The tree contains two different kinds of vulnerability. The AND branch represents a failure that requires several component events to occur together: the pumps are redundant, and the loss of one is absorbed by the other. The OR branches represent individual failures sufficient to cause the top event. The common electrical feeder, the controller and the manifold are

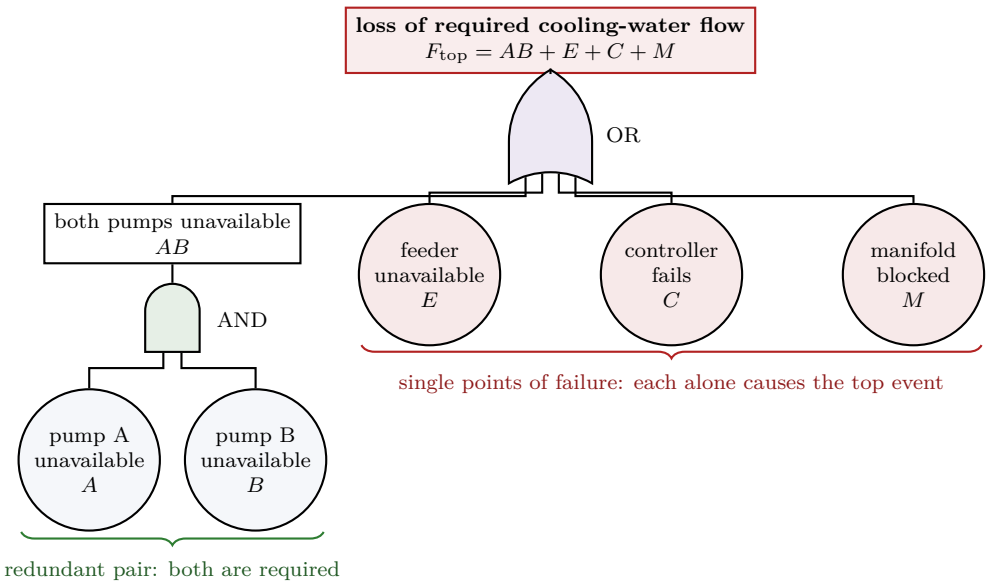


Figure 10.4: Fault tree of the loss of required cooling-water flow, $F_{\text{top}} = AB + E + C + M$: an OR gate over the AND branch of the two redundant pumps and the three shared items (feeder, controller, manifold). Rectangles are the top and intermediate events, circles basic events. The diagram is simplified; an engineering fault tree follows the diagramming conventions of its standard consistently.

therefore single points of failure under this system definition; the redundant pumps are not. In the vocabulary of the next chapter, $\{E\}$, $\{C\}$ and $\{M\}$ are minimal cut sets of order one and $\{A, B\}$ is a minimal cut set of order two.

This distinction is available from the structure of the tree alone, before any failure probability is assigned, and it is often the most useful result of the analysis: it shows where redundancy has been provided and where a shared item defeats it. The same shared items appeared in Chapter 4 as the shared dependencies that one-failure-at-a-time FMEA tends to miss; the fault tree represents them explicitly as separate inputs of the top OR gate.

10.3.3 Quantification step by step

Table 10.4 gives illustrative probabilities of the five basic events over one 24-hour mission. Each row is one basic event and its probability q_i of occurring during the mission. The five events are assumed independent and defined consistently over the same mission.

Table 10.4: Basic-event probabilities of the cooling-water fault tree over one 24-hour mission (illustrative).

Basic event	Symbol	Probability q_i
Pump A unavailable	A	0.050
Pump B unavailable	B	0.050
Electrical feeder unavailable	E	0.010
Controller failure	C	0.020
Manifold blockage	M	0.005

Step 1: the AND branch. The probability that both pumps are unavailable is, by equation (10.5),

$$\Pr(A \cap B) = q_A q_B = 0.05^2 = 0.0025.$$

Step 2: independence of the OR inputs. The top event is the union of four contributors, AB, E, C and M. No basic event appears in more than one of them, and events defined on disjoint sets of independent basic events are themselves independent, so the four contributors are independent and the n-input form of equation (10.9) applies. This step is what fails when a basic event is repeated in several branches, the case of Section 10.4.3.

Step 3: the complement. It is easier to compute the probability that none of the four contributors occurs and subtract it from one:

$$\begin{aligned}\Pr(\overline{F_{\text{top}}}) &= (1 - 0.0025)(1 - 0.01)(1 - 0.02)(1 - 0.005) \\ &= 0.9975 \times 0.99 \times 0.98 \times 0.995. \end{aligned} \tag{10.14}$$

Multiplying in order, $0.9975 \times 0.99 = 0.987525$, then $\times 0.98 = 0.967775$, then $\times 0.995 = 0.962936$. Therefore

$$Q_{\text{top}} = 1 - 0.962936 = 0.037064 \approx 3.71\,\%, \tag{10.15}$$

the probability of the loss of cooling-water flow under the assumed independent basic-event probabilities for the 24-hour mission. The sum of the four contributor probabilities, $0.0025 + 0.01 + 0.02 + 0.005 = 0.0375$, exceeds the exact value by 1.2 %, as the rare-event bound of Section 10.2.2 requires.

10.3.4 Dominance of the single points of failure

The probability that both pumps fail is only 0.25 %; the probability of the system event is about 3.71 %, nearly fifteen times larger. The three single points of failure alone, without the pumps, give $1 - 0.99 \times 0.98 \times 0.995 = 1 - 0.965349 = 0.034651$, which is 93.5 % of Q_{top}; the controller alone, at 0.02, is more than half of it. The single points of failure dominate this example.

Table 10.5 repeats the calculation of Section 10.3.3 for several sets of basic-event probabilities. Each row is one case: the pump probability (equal for both pumps), the probabilities of the three shared items, the probability of the AND branch, the exact top-event probability and the rare-event sum. Doubling the pump failure probability, from $q = 0.05$ to 0.10, raises Q_{top} only from 0.0371 to 0.0443; improving the pumps fivefold, to $q = 0.01$, lowers it only to 0.0347. Dividing the three shared probabilities by ten lowers it to 0.0060, and removing them entirely leaves the 0.0025 of the pump pair. The pumps are well protected by their redundancy; the cooling function is not, because the redundancy covers only the pumps.

> **Principle 10.2.** *Component redundancy must be analyzed in the context of the complete functional system. A redundant pair protects the system only against the failures of the pair; a shared item in series with it caps the reliability of the function, however reliable the pair.*

10.4 Boolean algebra and fault-tree reduction

Real fault trees contain hundreds or thousands of events, and the same component failure may appear in several branches. This section states the laws of Boolean algebra used to simplify a top-event expression before it is quantified, works a simplification, and shows why multiplying the probabilities of branches that share an event gives a wrong answer.

Table 10.5: Top-event probability of the cooling-water fault tree, $Q_{\text{top}} = 1 - (1 - q_A q_B)(1 - q_E)(1 - q_C)(1 - q_M)$, for several sets of basic-event probabilities (illustrative; the first row is Table 10.4).

Case	$q_A = q_B$	q_E	q_C	q_M	$q_A q_B$	Q_{top}
Baseline	0.05	0.010	0.020	0.0050	0.0025	0.037064
Pumps twice as likely to fail	0.10	0.010	0.020	0.0050	0.0100	0.044304
Pumps four times as likely	0.20	0.010	0.020	0.0050	0.0400	0.073265
Pumps five times better	0.01	0.010	0.020	0.0050	0.0001	0.034748
Shared items ten times better	0.05	0.001	0.002	0.0005	0.0025	0.005988
Shared items perfect	0.05	0	0	0	0.0025	0.002500

10.4.1 The laws of Boolean algebra

When a component failure appears in more than one branch, it is a *repeated basic event*, and the probability calculation must not count it more than once. Boolean algebra provides the rules for simplifying an expression so that each combination of events appears in it once, in its smallest form. Table 10.6 lists the laws. Each row gives the name of a rule, its expression with + for OR (union), juxtaposition for AND (intersection) and an overbar for the complement, and its meaning in fault-tree terms. The rows marked "added" complete the set for use in this chapter; Ω is the certain event and $\emptyset$ the impossible event, written 1 and 0 in Boolean shorthand.

Table 10.6: Laws of Boolean algebra for fault-tree events. They hold for events as sets and require no probabilistic independence.

Rule	Boolean expression	Meaning
Commutative (added)	$A + B = B + A$; $AB = BA$	The order of the inputs of a gate does not matter
Associative (added)	$A + (B + C) = (A + B) + C$; $A(BC) = (AB)C$	Nested gates of one type merge into one gate
Idempotence	$A + A = A$	A repeated occurrence of the same event is still one event
AND idempotence	$AA = A$	An event AND itself does not create two failures
Absorption	$A + AB = A$	If A alone causes failure, adding B is irrelevant to that branch
Absorption, dual (added)	$A(A + B) = A$	Requiring A, and A or B, requires only A
Distributive	$A(B + C) = AB + AC$	Expand combinations
Distributive, second form (added)	$A + BC = (A + B)(A + C)$	No counterpart in ordinary algebra
De Morgan	$\overline{A + B} = \overline{A}\,\overline{B}$	The complement of OR is the AND of the complements
De Morgan	$\overline{AB} = \overline{A} + \overline{B}$	The complement of AND is the OR of the complements
Complement (added)	$A\overline{A} = \emptyset$; $A + \overline{A} = \Omega$	An event and its complement cannot both occur; one of them occurs
Identities (added)	$A + \Omega = \Omega$; $A\Omega = A$; $A + \emptyset = A$; $A\emptyset = \emptyset$	In shorthand $A + 1 = 1$, $A \cdot 1 = A$, $A + 0 = A$, $A \cdot 0 = 0$

The laws apply to events as sets and do not require probabilistic independence; they are

identities of logic, valid for any probability model. Several of them differ from the rules of ordinary arithmetic, which is the usual source of error. In ordinary algebra $a + a = 2a$ and $a \cdot a = a^2$; in Boolean algebra $A + A = A$ and $AA = A$, because a pump that fails "twice" in two branches of a tree has failed once. Absorption follows from the identities: $A + AB = A\Omega + AB = A(\Omega + B) = A\Omega = A$. The second distributive law has no counterpart in ordinary algebra, and it can be checked through the first: $(A+B)(A+C) = AA+AC+BA+BC = A+AC+AB+BC = A+BC$, by idempotence and two absorptions. The De Morgan laws turn the complement of an OR gate into an AND gate of complements, which is the step used in equation (10.9) to compute the probability of an OR gate from the product of the probabilities that its inputs do not occur.

10.4.2 A worked simplification

Suppose an initial analysis produces the top-event expression

$$F_{\text{top}} = AB + AC + ABC + D. \tag{10.16}$$

The term ABC is redundant: if A, B and C all occur, then AB has already occurred, so the term adds no outcome that AB does not already include. By absorption,

$$AB + ABC = AB(\Omega + C) = AB\,\Omega = AB,$$

and therefore

$$F_{\text{top}} = AB + AC + D. \tag{10.17}$$

Factoring by the distributive law,

$$F_{\text{top}} = A(B + C) + D, \tag{10.18}$$

a more concise representation. It states that the system fails if A occurs together with either B or C, or if D occurs alone. The three critical failure combinations are

$$\{A, B\}, \qquad \{A, C\}, \qquad \{D\};$$

none of them can be made smaller and still cause the top event. They are the minimal cut sets of the tree, which Chapter 11 defines rigorously and computes for large trees.

Figure 10.5 draws the two forms. Panel (a) is the tree of equation (10.16) as derived: an OR gate over three AND gates and the basic event D; the basic event A (shaded) appears three times, and the third AND gate, over A, B and C, is the redundant term. Panel (b) is the tree of equation (10.18): A appears once, joined by an AND gate to an OR gate over B and C, and D enters the top OR gate directly. Both trees have the same truth table over the sixteen combinations of A, B, C and D; the reduction changes the representation, not the logic.

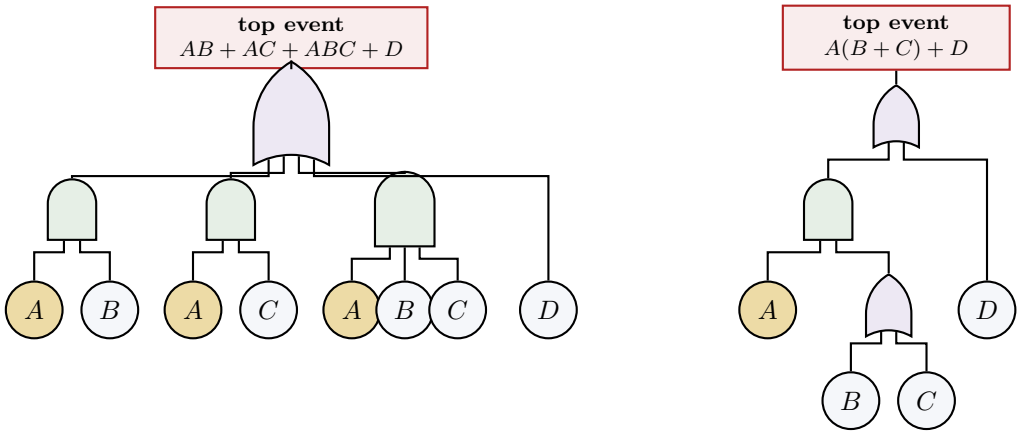
10.4.3 Why naive multiplication fails

Suppose

$$F_{\text{top}} = AB + AC,$$

with A, B and C mutually independent basic events of probabilities q_A, q_B and q_C. It would be incorrect to treat the terms AB and AC as independent events, because both contain A: their intersection is $AB \cap AC = ABC$, with probability $q_A q_B q_C$, while the product of their probabilities is $q_A^2 q_B q_C$. The exact probability follows from inclusion–exclusion, equation (10.8), applied to the two terms:

$$Q_{\text{top}} = \Pr(AB \cup AC) = \Pr(AB) + \Pr(AC) - \Pr(ABC) = q_A q_B + q_A q_C - q_A q_B q_C. \tag{10.19}$$



(a) as derived: A appears three times

(b) reduced: A appears once

Figure 10.5: Boolean reduction of a fault tree. (a) The tree of $AB + AC + ABC + D$, with the repeated basic event A shaded; the third AND gate is absorbed by the first. (b) The equivalent tree of $A(B + C) + D$. For compactness, gates are joined directly; in a documented tree each gate output carries an event rectangle.

Equivalently, factoring the Boolean expression first, $F_{\text{top}} = A(B + C)$, the two factors A and $B + C$ involve disjoint sets of events and are independent, so

$$Q_{\text{top}} = q_A \left[q_B + q_C - q_B q_C \right], \tag{10.20}$$

and multiplying out the bracket gives equation (10.19) term by term. The two results are identical.

With $q_A = 0.10$, $q_B = 0.20$ and $q_C = 0.30$ (illustrative),

$$Q_{\text{top}} = 0.10\,(0.20 + 0.30 - 0.06) = 0.10 \times 0.44 = 0.044,$$

a top-event probability of 4.4 %. Simply adding the probabilities of the two terms,

$$\Pr(AB) + \Pr(AC) = 0.02 + 0.03 = 0.05,$$

gives 5 % and overestimates the true probability, because the overlap ABC, of probability $0.10 \times 0.20 \times 0.30 = 0.006$, is counted twice. Treating the two terms as independent inputs of an OR gate, $1 - (1 - 0.02)(1 - 0.03) = 1 - 0.9506 = 0.0494$, is also wrong: it assigns the overlap the probability $0.02 \times 0.03 = 0.0006$ instead of 0.006. Only the calculation that respects the shared event A gives 0.044. The step of Section 10.3.3 that multiplied the complements of the four branches was valid because no basic event appeared in two branches; with repeated basic events, the distinction becomes increasingly important as the tree grows, and Chapter 11 gives the general methods.

> **Principle 10.3.** *Branches of a fault tree that share a basic event are not independent, even when every basic event is. Reduce the expression with Boolean algebra, or apply inclusion–exclusion to the shared events, before multiplying probabilities.*

10.5 Coherent and noncoherent logic and other gates

The trees of the preceding sections share one property: an additional failure can never remove the top event. This section defines that property, coherence, through the structure function of the tree, shows a failure condition that violates it, and then introduces the further gates (voting, inhibit, priority AND and exclusive OR) that static and dynamic fault trees use, closing with a worked table of two-input gate probabilities.

10.5.1 Coherent fault trees

Informally, a coherent fault tree has a monotone failure structure: *if the system has already failed, additional component failures cannot restore it to success.* To state this precisely, let $\mathbf{x} = (x_1, \ldots, x_n)$ be the vector of basic-event indicators, with $x_i = 1$ if basic event i occurs (the component has failed) and $x_i = 0$ otherwise, and define the failure-oriented structure function of the tree

$$\phi_F(\mathbf{x}) = \begin{cases} 1, & \text{the top event occurs,} \\ 0, & \text{the top event does not occur.} \end{cases} \tag{10.21}$$

The structure is monotone if

$$\mathbf{x} \le \mathbf{y} \quad \Longrightarrow \quad \phi_F(\mathbf{x}) \le \phi_F(\mathbf{y}), \tag{10.22}$$

where $\mathbf{x} \le \mathbf{y}$ means $x_i \le y_i$ for every i: every failure present in $\mathbf{x}$ is also present in $\mathbf{y}$. The structure function of Chapter 9 is success-oriented, $\phi(\mathbf{s}) = 1$ when the system functions, with $s_i = 1$ when component i functions; the two are related by $\phi_F(\mathbf{x}) = 1 - \phi(\mathbf{1} - \mathbf{x})$.

Strict definitions of coherence add a second condition, component relevance: every component i must matter in at least one state of the others, that is, there must be a state $\mathbf{x}$ for which $\phi_F(1_i, \mathbf{x}) \ne \phi_F(0_i, \mathbf{x})$, where $(1_i, \mathbf{x})$ and $(0_i, \mathbf{x})$ denote $\mathbf{x}$ with x_i set to 1 and to 0. A structure is coherent when it is monotone and every component is relevant; a monotone structure that contains irrelevant components is sometimes called semicoherent.

Ordinary AND and OR combinations of positive failure events, events that state that something fails rather than that something works, produce monotone logic. The AND gate is $\prod_i x_i$ and the OR gate is $1 - \prod_i (1 - x_i)$; both are nondecreasing in every input, and a composition of nondecreasing functions is nondecreasing. Take the example

$$F_{\text{top}} = AB + C, \qquad \phi_F(\mathbf{x}) = 1 - (1 - x_A x_B)(1 - x_C).$$

If $x_C = 1$, the top event occurs regardless of A and B, and no additional failure can reverse it. Every component is relevant: A decides the outcome when $x_B = 1$ and $x_C = 0$, B when $x_A = 1$ and $x_C = 0$, and C when $x_A x_B = 0$. The logic is coherent. So is the cooling-water tree of Figure 10.4, with $\phi_F = 1 - (1 - x_A x_B)(1 - x_E)(1 - x_C)(1 - x_M)$.

Coherence has a quantitative consequence. With independent basic events, conditioning on component i gives $Q_{\text{top}} = q_i \Pr(\phi_F = 1 \mid x_i = 1) + (1 - q_i) \Pr(\phi_F = 1 \mid x_i = 0)$, which is linear in q_i with slope $\Pr(\phi_F = 1 \mid x_i = 1) - \Pr(\phi_F = 1 \mid x_i = 0)$. In a monotone structure this slope is never negative, so the top-event probability can only rise when the probability of any basic event rises. The slope is the Birnbaum importance of component i, the subject of Chapter 12.

10.5.2 Noncoherent failure logic

Not every engineered failure condition is monotone. Consider two control relays, and suppose that an undesired command is issued only when relay A fails in its energized position while relay B remains operational, because a failed relay B would block the command. Define

$$A = \{\text{relay A fails energized}\}, \qquad B = \{\text{relay B fails}\}.$$

The hazardous command occurs when A occurs but B does not,

$$F_{\text{top}} = A \cap \overline{B}, \tag{10.23}$$

written $F_{\text{top}} = A\overline{B}$ in Boolean notation. The complemented event $\overline{B}$ means that B does *not* fail. Suppose that A has failed and B is functioning: the top event occurs. If B then fails in a way that blocks the command, the top event no longer exists. An additional failure

has suppressed the undesired condition, which is nonmonotone, or noncoherent, logic: with $\mathbf{x} = (x_A, x_B) = (1, 0)$ and $\mathbf{y} = (1, 1)$, $\mathbf{x} \leq \mathbf{y}$ but $\phi_F(\mathbf{x}) = 1 > \phi_F(\mathbf{y}) = 0$.

Figure 10.6 draws the tree. The top event "hazardous command issued" is the output of an AND gate with two inputs: the basic event A, on the left, and on the right a NOT gate, the triangle with a small circle at its output, whose input is the basic event B and whose output is its complement $\overline{B}$. Table 10.7 is the truth table. It differs from the AND gate of Table 10.2 only in the last row, where both relays have failed and the hazardous command is not issued. A standard coherent fault tree cannot represent this condition with AND and OR gates over positive failure events alone, because every such tree is monotone and this function is not.

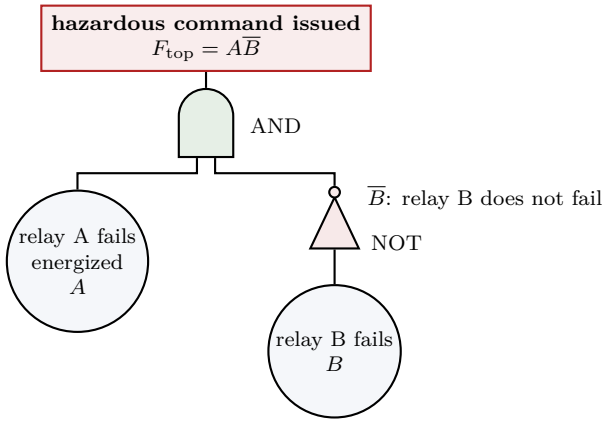


Figure 10.6: A noncoherent fault tree: the hazardous command is issued when relay A fails energized and relay B does not fail, $F_{\text{top}} = A \cap \overline{B}$. The NOT gate turns the basic event B into its complement.

Table 10.7: Truth table of the noncoherent tree of Figure 10.6 (1 = the event occurs).

Relay A fails energized, x_A	Relay B fails, x_B	Hazardous command, x_{top}
0	0	0
0	1	0
1	0	1
1	1	0

Quantitative example. Assume that A and B are independent, with $q_A = 0.10$ and $q_B = 0.20$ (illustrative). The complement of an event independent of A is also independent of A, so

$$Q_{\text{top}} = \Pr(A \cap \overline{B}) = \Pr(A)\,[1 - \Pr(B)] = 0.10 \times 0.80 = 0.08,$$

a probability of 8 %. The slope of Q_{top} with respect to q_B is $-q_A = -0.10$: a less reliable relay B makes the hazard less likely, which is the quantitative signature of noncoherence and the reverse of the monotone property of Section 10.5.1. Replacing $\overline{B}$ by the certain event, which turns the tree into the coherent $F_{\text{top}} = A$, gives 0.10, an upper bound; this conservative coherent approximation is common practice when the complemented events have probabilities close to one.

This example is intentionally unusual. In actual safety engineering, the analyst must verify that the failure of B genuinely prevents the hazardous outcome, rather than merely changing its sequence or its consequences. In particular, a failure that occurs *after* an irreversible accident cannot undo that accident, and the static expression $A\overline{B}$, which looks only at the

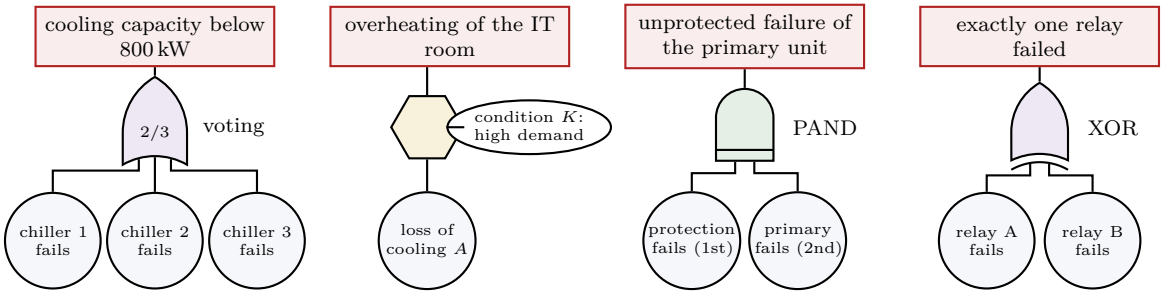
states at one time, would wrongly credit it. Modeling such systems often requires dynamic event sequencing rather than static Boolean logic, and the standard fault-tree handbooks recognize that conditional and sequence-dependent events need special treatment beyond ordinary coherent trees.

10.5.3 Voting and inhibit gates

AND and OR gates are sufficient for many static fault trees, but additional logical structures are useful. Table 10.8 lists four of them. Each row names the gate, states the condition for its output, gives an engineering example and the probability of the output when the inputs are independent (for the inhibit gate, in general). Figure 10.7 draws each gate with its example. Panel (a) is a voting gate, an OR-shaped gate marked 2/3, over the three chillers of the cooling plant; panel (b) an inhibit gate, a hexagon whose input enters from below and whose enabling condition is the oval at its side; panel (c) a priority AND gate, an AND gate with a bar across its base, whose inputs are numbered in the order required; and panel (d) an exclusive-OR gate, an OR gate with a second curved line at its input side.

Table 10.8: Further fault-tree gates, their meaning, an engineering example and the probability of the output for independent inputs (the inhibit gate in general).

Gate	Meaning	Engineering example	Output probability
Voting (k of n)	At least k of the n input events occur	Cooling failure when at least two of three chillers fail	$\sum_{j=k}^{n} \binom{n}{j} q^j (1-q)^{n-j}$ for identical inputs
Inhibit	An input event produces the output only under a specified condition	A cooling interruption causes overheating only when the thermal demand is high enough	$\Pr(A)\,\Pr(K \mid A)$
Priority AND	All inputs occur, in a specified order	The protection becomes unavailable before the primary unit fails	$\Pr(T_1 < T_2 \le H)$, eq. (10.27)
Exclusive OR	Exactly one input event occurs	A fault condition that requires one relay failed and the other functioning	$q_A + q_B - 2q_A q_B$



(a) voting gate, 2 of 3 (b) inhibit gate (c) priority AND gate (d) exclusive OR gate

Figure 10.7: Further gates with their engineering examples. (a) Voting gate: the output occurs when at least two of the three chillers fail. (b) Inhibit gate: the loss of cooling A causes overheating only under the condition K, a thermal demand above the passive cooling capability. (c) Priority AND gate: the output occurs only if the protection fails before the primary unit. (d) Exclusive OR gate: the output occurs when exactly one of the two relays has failed.

The voting gate. A k-of-n voting gate produces its output when at least k of its n input events occur. For n independent inputs with the same probability q, the number of inputs that occur is binomial, and

$$Q_{\text{top}} = \sum_{j=k}^{n} \binom{n}{j} q^j (1-q)^{n-j}. \tag{10.24}$$

The cooling plant of Chapter 1 needs two of its three chillers; each delivers about 430 kW, so the snapshot top event $C_{\text{available}} < 800\,\text{kW}$ of equation (10.2), counting chiller statuses only, occurs when at least two chillers fail. With the failure probability $q = 0.05$ of one chiller over the year used there,

$$Q_{\text{top}} = 3q^2(1-q) + q^3 = 3 \times 0.0025 \times 0.95 + 0.000125 = 0.007125 + 0.000125 = 0.00725,$$

which is $1 - 0.99275$, the complement of the 2-out-of-3 system reliability of Chapter 1. In general a k-out-of-n success structure fails when at least $n - k + 1$ units fail, so its fault tree has a voting gate with threshold $n - k + 1$; for 2-out-of-3 the two thresholds coincide at 2. A voting gate can always be expanded into AND and OR gates, here $AB + AC + BC$ over the chiller events A, B and C, but the expansion grows quickly with n, and the voting gate keeps the tree readable.

The inhibit gate. The inhibit gate is useful when a failure requires both an initiating event and an enabling condition. Its output is the intersection

$$F_{\text{top}} = A \cap K, \tag{10.25}$$

where A is the event that the cooling system is unavailable and K is the condition that the thermal demand exceeds the available passive cooling capability, so that the room overheats before cooling is restored. Logically it is an AND gate; the separate symbol marks that K is a condition rather than a failure. Its probability is

$$Q_{\text{top}} = \Pr(A)\ \Pr(K \mid A), \tag{10.26}$$

and independence should not be assumed unless it is justified. Here it is not: Chapter 1 showed that the same hot afternoons that raise the cooling load also derate the chillers, so the demand is more likely to be high when cooling is lost than on an average day. With the cooling-water result $\Pr(A) = 0.037064$ of equation (10.15) and an illustrative $\Pr(K \mid A) = 0.30$, the probability of overheating over the mission is $0.037064 \times 0.30 = 0.0111$.

10.5.4 Priority AND and exclusive OR gates

When the order of events matters, a static AND expression may be insufficient. Consider a primary unit protected by a protection function, such as an automatic transfer to a standby unit. If the primary fails first, the protection acts and a later failure of the protection is harmless for this top event; if the protection fails first, a later failure of the primary goes unprotected. The top event is "both fail, the protection first", which the priority AND gate represents. Let the times to failure T_1 of the protection and T_2 of the primary be independent and exponential, with constant failure rates λ_1 and λ_2, over a mission H. Conditioning on the failure time s of the primary, the protection must have failed before s:

$$\begin{aligned} \Pr(T_1 < T_2 \le H) &= \int_0^H \lambda_2 e^{-\lambda_2 s} \left(1 - e^{-\lambda_1 s}\right) \mathrm{d}s \\ &= \left(1 - e^{-\lambda_2 H}\right) - \frac{\lambda_2}{\lambda_1 + \lambda_2}\left(1 - e^{-(\lambda_1+\lambda_2)H}\right). \end{aligned} \tag{10.27}$$

With $\lambda_1 = 1 \times 10^{-4}\,\mathrm{h}^{-1}$, $\lambda_2 = 2 \times 10^{-4}\,\mathrm{h}^{-1}$ and $H = 1000\,\mathrm{h}$ (illustrative), $1 - e^{-0.2} = 0.181269$ and $1 - e^{-0.3} = 0.259182$, so the priority AND gives $0.181269 - (2/3) \times 0.259182 = 0.181269 - 0.172788 = 0.008481$. The static AND of the same two events, both failed by H, gives $(1 - e^{-0.1})(1 - e^{-0.2}) = 0.095163 \times 0.181269 = 0.017250$, about twice as much; the reverse order has probability 0.008769, and the two orders add up to the static AND. Because the ordered event is contained in the unordered one, a static AND always overstates the probability of an ordered event; here it doubles it, and the factor depends on the rates, approaching one when the protection is much less reliable than the primary. Priority AND gates and the other gates of dynamic fault trees represent such sequences and dependencies explicitly; the aerospace fault-tree handbook discusses these extensions and their interpretation.

The exclusive OR gate. The exclusive OR gate produces its output when exactly one of its two inputs occurs,

$$F_{\mathrm{top}} = A\overline{B} + \overline{A}B, \tag{10.28}$$

as in a fault condition that requires one relay failed and the other functioning. The two terms are disjoint, so for independent inputs $Q_{\mathrm{top}} = q_A(1 - q_B) + (1 - q_A)q_B = q_A + q_B - 2q_Aq_B$. With $q_A = 0.10$ and $q_B = 0.20$, $Q_{\mathrm{top}} = 0.10 + 0.20 - 0.04 = 0.26$, against 0.28 for the OR gate over the same events; the difference is the probability $q_Aq_B = 0.02$ that both occur. The exclusive OR contains complemented events and is noncoherent: from one failed relay, a failure of the second removes the output.

10.5.5 Top-event probability of a two-input gate

The logical gate and the statistical dependence of its inputs are distinct parts of the model, and Table 10.9 shows the effect of each on the simplest tree, one gate over two events A and B. The model is: the joint probability is $\Pr(A \cap B) = q_Aq_B$ when the events are independent and $q_A \Pr(B \mid A)$ by equation (10.10) when they are not; the AND gate gives $\Pr(A \cap B)$ and the OR gate gives $q_A + q_B - \Pr(A \cap B)$ by equation (10.8); and the inputs are consistent only if $\Pr(A \cap B)$ lies within the bounds of equation (10.11). Each row of the table is one set of inputs. For the third row, $\Pr(A \cap B) = 0.10 \times 0.50 = 0.05$, the AND gate gives 0.05 and the OR gate $0.10 + 0.20 - 0.05 = 0.25$.

Table 10.9: Top-event probability of a single AND or OR gate over two events for several input sets: $\Pr(A \cap B) = q_A \Pr(B \mid A)$, $\Pr(A \cup B) = q_A + q_B - \Pr(A \cap B)$; "indep." means $\Pr(B \mid A) = q_B$. The last two rows violate $\max(0, q_A + q_B - 1) \le \Pr(A \cap B) \le \min(q_A, q_B)$.

q_A	q_B	$\Pr(B \mid A)$	AND, $\Pr(A \cap B)$	OR, $\Pr(A \cup B)$	Remark
0.10	0.20	indep.	0.0200	0.2800	independent events
0.10	0.20	0.20	0.0200	0.2800	$\Pr(B \mid A) = q_B$ reproduces independence
0.10	0.20	0.50	0.0500	0.2500	positive dependence: AND up, OR down
0.10	0.20	1.00	0.1000	0.2000	A implies B: AND at its upper bound q_A
0.10	0.20	0	0	0.3000	mutually exclusive: OR is the sum
0.05	0.05	indep.	0.0025	0.0975	the redundant pump pair
0.05	0.05	0.30	0.0150	0.0850	load sharing raises the pair's AND sixfold
0.50	0.20	0.60	0.30	–	inconsistent: $0.30 > q_B = 0.20$
0.70	0.60	0.30	0.21	–	inconsistent: $0.21 < q_A + q_B - 1 = 0.30$

Read down the first five rows, which keep q_A and q_B fixed: as $\Pr(B \mid A)$ rises from 0 to 1, the AND probability rises from 0 to 0.10 and the OR probability falls from 0.30 to 0.20, by exactly the same amount, because the two always add up to $q_A + q_B$. With the same marginal probabilities, the AND gate can therefore give anything from 0 to q_A and the OR gate anything from 0.20 to 0.30, depending only on the dependence. The logic fixes which probability is wanted; the dependence fixes its value. The last two rows show that a conditional probability cannot be chosen freely: a joint probability above either marginal, or so small that the union would exceed one, describes no possible pair of events.

10.6 Fault-tree construction in practice

A fault tree is built by a disciplined procedure, checked against the other analyses of the same system, and interpreted within the limits of a static logic model. This section states the construction workflow, relates fault trees to FMEA and reliability block diagrams, applies the method to a transmission network, and lists the common modeling mistakes.

10.6.1 A construction workflow

A disciplined fault tree analysis proceeds through seven steps.

1. *Define the top event.* Specify exactly what undesired system condition is analyzed, with the elements of Table 10.1.
2. *Establish the boundaries.* Identify the equipment, the operating modes, the external dependencies (utility supply, control network, cooling of the electrical rooms) and the mission duration.
3. *Identify the immediate causes.* Determine which events are directly sufficient for the top event (OR) or jointly necessary for it (AND).
4. *Develop the intermediate events.* Expand each branch using documented AND, OR or other appropriate gates.
5. *Identify the basic events.* Stop decomposing when the selected resolution of the analysis is reached and the event probabilities can be supported by data or defensible judgment.
6. *Check the logic and the completeness.* Look for repeated events, omitted shared dependencies, inconsistent boundaries and incorrect gate relations.
7. *Simplify and quantify.* Use Boolean reduction and an appropriate probability method to evaluate the top event.

The standard handbooks add construction rules that keep the logic consistent. All the immediate input events of a gate are defined before any of them is developed further (the complete-the-gate rule), which prevents missing branches and inconsistent logic; every gate output is an event with its own statement, so that gates are not connected directly to gates in a documented tree (the no-gate-to-gate rule); each event is written as a fault, stating what fails and when; and the normal functioning of components is assumed where it is not itself the subject of an event, so that no "miracle" cancels a fault. For an engineering study, every branch should also be traceable to a system function, a component behavior, a failure mechanism or a defensible engineering assumption, so that a reviewer can check the tree against the system rather than against the analyst's memory.

10.6.2 Fault trees, FMEA and reliability block diagrams

FMEA, the method of Chapter 4, begins with the components and asks what effects their failures produce. Fault tree analysis begins with an undesired system event and asks which component or system failures can produce it. Table 10.10 places both, with the reliability block diagram of Chapter 9, on the same three properties: the starting point, the direction of the

analysis and the principal output. The RBD and the fault tree describe the same logic from opposite sides: for a system that fails when it stops functioning, a series arrangement of blocks becomes an OR gate over the block failures, and a parallel arrangement becomes an AND gate.

Table 10.10: FMEA, fault tree analysis and reliability block diagrams compared.

Analysis	Starting point	Direction	Principal output
FMEA	Component or function	Bottom-up	Failure modes and effects
FTA	Undesired system event	Top-down	Logical failure combinations
RBD	Successful system function	Success-oriented	Required success configurations

A practical workflow uses the findings of an FMEA to identify candidate basic events for the fault tree: the failure modes of the worksheet whose end effect is the top event, or contributes to it, become basic events or the starting points of branches. The FMEA list, however, is not automatically a complete fault tree. It examines one failure at a time, and it may not capture combinations of failures, conditional dependencies such as the load sharing of Section 10.2.3, or shared causes such as the common feeder of Figure 10.4. Those require the additional system-level analysis that the tree provides.

10.6.3 Application to power systems

Consider a transmission network with two parallel lines connecting a generating area to a load, through a common substation, and assume that either line can carry the required demand by itself. Define

$$A = \{\text{line A unavailable}\}, \qquad B = \{\text{line B unavailable}\},$$
$$S = \{\text{common substation unavailable}\}.$$

The top event, the failure of the network to provide the required transfer, is

$$F_{\text{top}} = AB + S. \tag{10.29}$$

Figure 10.8 draws the tree with the same layout as Figure 10.4: an OR gate under the top event, with the intermediate event "both lines unavailable", explained by an AND gate over A and B, as one input, and the basic event S as the other. The illustrative basic-event probabilities are printed under the circles. The substation is a single point of failure, the lines a redundant pair.

Let $q_A = 0.02$, $q_B = 0.03$ and $q_S = 0.01$. Assuming independence of the three modeled events,

$$\Pr(AB) = 0.02 \times 0.03 = 0.0006,$$

and, since AB and S share no basic event,

$$Q_{\text{top}} = 1 - (1 - 0.0006)(1 - 0.01) = 1 - 0.9994 \times 0.99 = 1 - 0.989406 = 0.010594,$$

a top-event probability of about 1.06 %, almost all of it from the substation. The independence of the two lines deserves the doubt of Section 10.2.3: lines in the same corridor share the weather, and a storm or a heat wave that reduces the rating of one reduces the rating of the other.

The model is also incomplete if the transmission network depends on

- voltage stability;
- reactive power support;
- thermal constraints under changing weather;
- the operation of the protection system;

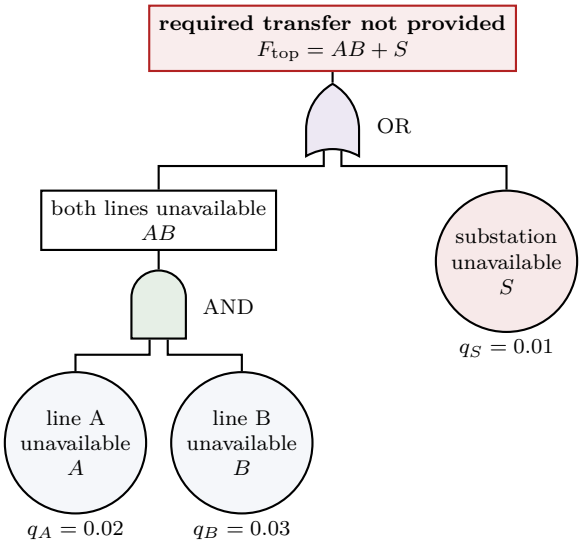


Figure 10.8: Fault tree of the failure of a two-line transmission corridor to provide the required transfer, $F_{\text{top}} = AB + S$, with illustrative basic-event probabilities. Assuming independence, $Q_{\text{top}} = 1 - (1 - 0.0006)(1 - 0.01) = 0.010594$.

- generator availability;
- load variations;
- the network topology and the redistribution of power flow.

A static binary fault tree represents combinations of outages well, but it does not by itself determine whether the resulting physical state of the network is acceptable. Whether one line can carry the transfer after the other trips depends on its rating at that moment and on how the power flow redistributes, which are outputs of a network calculation, not inputs of the tree. A more complete model connects the outage events to power-flow simulations or other physical system models, so that the top event is evaluated on the physical state each outage combination produces. The later chapters on power-system adequacy take up this distinction.

10.6.4 Common modeling mistakes

Table 10.11 collects the common mistakes of fault-tree modeling. Each row names a mistake and its consequence for the result, and each has been met in this chapter: the vague top event in Section 10.1.2, the gate read as a probability formula in Section 10.2.3, the repeated event in Section 10.4.3, the omitted shared dependency in Section 10.3.4, the time ordering in Section 10.5.4, and the static tree taken for a physical model in Section 10.6.3.

The last row summarizes the others. A mathematically correct probability calculation cannot compensate for incorrect event definitions, missing failure paths or unjustified independence assumptions. Large fault trees are reduced to their essential failure combinations, the minimal cut sets, which Chapter 11 derives, together with minimal path sets, exact and rare-event quantification, and the algorithms that handle repeated basic events.

10.7 What to retain

Definition 10.1 (Fault tree)**.** A fault tree is a deductive logic model that expresses one precisely defined undesired system event, the top event, as a Boolean function of basic events through AND, OR and other gates; the tree defines the structure of failure, and a separate probability model, with its own assumptions of independence or dependence, determines how likely that structure is to occur.

Table 10.11: Common fault-tree modeling mistakes and their consequences.

Mistake	Consequence
Defining the top event vaguely	Ambiguous reliability results
Treating AND and OR gates as probability formulas	Incorrect handling of dependent events
Assuming every repeated event is independent	Double-counted failure probabilities
Omitting shared dependencies	Overestimated benefits of redundancy
Mixing snapshot and mission events	Inconsistent probabilities
Ignoring time ordering	Incorrect representation of standby or protection sequences
Treating a static fault tree as a complete physical system model	Missing capacity, thermal, electrical or dynamic constraints
Quantifying a tree without checking its logic	Precise calculations for an incorrect failure model

Table 10.12 lists the chapter's terms with their meaning, their expression and units, and the subsection that defines each. All probabilities are dimensionless and refer to a stated mission.

Table 10.12: Terms of fault tree construction and logic.

Term	Meaning	Expression, units	Defined in
Fault tree analysis	deductive analysis of the causes of one undesired event	tree, Boolean expression	10.1.1
Top event	the undesired system event, precisely defined	F_{top}; $Q_{\text{top}} = \Pr(F_{\text{top}})$	10.1.2
Static snapshot event	top event evaluated on equipment states at one time	$\{C_{\text{available}} < 800\,\text{kW}\}$	10.1.2
Intermediate, basic, undeveloped event	explained by a gate; not decomposed; not decomposed for lack of information	rectangle; circle; diamond	10.1.3
Transfer	continuation of a branch elsewhere	triangle	10.1.3
AND gate	all inputs occur	$A \cap B = AB$; $q_A q_B$ if independent	10.2.1
OR gate	at least one input occurs	$A \cup B = A + B$; $q_A + q_B - q_A q_B$	10.2.2
Inclusion–exclusion	probability of a union	$\Pr(A) + \Pr(B) - \Pr(A \cap B)$	10.2.2
Dependent AND	joint probability through a conditional	$\Pr(A)\Pr(B \mid A)$	10.2.3
Single point of failure	basic event that alone causes the top event	minimal cut set of order one	10.3.2
Repeated basic event	event appearing in several branches	branches not independent	10.4.1
Boolean reduction	simplification by the laws of Boolean algebra	$A + AB = A$, $AA = A$, De Morgan	10.4.1, 10.4.2
Structure function	top-event indicator as a function of basic-event indicators	$\phi_F(\mathbf{x}) \in \{0, 1\}$	10.5.1
Coherent, semicoherent	monotone with every component relevant; monotone only	$\mathbf{x} \le \mathbf{y} \Rightarrow \phi_F(\mathbf{x}) \le \phi_F(\mathbf{y})$	10.5.1
Noncoherent logic	a further failure can remove the top event	complemented events, NOT gate, $A\overline{B}$	10.5.2
Voting gate	at least k of n inputs occur	eq. (10.24)	10.5.3
Inhibit gate	input causes output only under a condition	$\Pr(A)\Pr(K \mid A)$	10.5.3
Priority AND gate	all inputs occur in a specified order	$\Pr(T_1 < T_2 \le H)$	10.5.4
Exclusive OR gate	exactly one input occurs	$q_A + q_B - 2q_A q_B$	10.5.4
Static, dynamic fault tree	combinations of states; also their order and timing	AND/OR; PAND and sequence gates	10.5.4

Three ideas carry forward. The top event must be defined before the tree is drawn, and its definition, static or temporal, decides whether a static fault tree is adequate at all. The structure of the tree, which events are single points of failure and which are protected by redundancy, is available before any probability, and in the example of this chapter the single points dominated the result. And the logic of the gates is independent of the probability model: independence, repeated events and time ordering are assumptions of the quantification, each to be justified, and the next chapter's minimal cut sets are the tool for quantifying large trees correctly.

Further reading. Rausand, Barros and Høyland, *System Reliability Theory*, third edition (fault-tree analysis, Boolean reliability models and coherent systems). Kumamoto and Henley, *Probabilistic Risk Assessment and Management for Engineers and Scientists*, second edition

(fault-tree methodology and probabilistic risk assessment). U.S. Nuclear Regulatory Commission, *Fault Tree Handbook*, NUREG-0492 (1981) (the foundational reference for construction, Boolean reduction and quantification). NASA, *Fault Tree Handbook with Aerospace Applications* (detailed methods and engineering examples).

10.8 Exercises

Exercise 10.1. Answer each part and justify the answer in one sentence. (a) Two independent pumps must both fail for cooling flow to be lost. Which gate represents this condition: AND, OR or exclusive OR? (b) A shared electrical supply failure can by itself disable all pumps. Through which gate does it combine with the dual-pump failure event? (c) For $F_{\text{top}} = AB + AC$, why can $\Pr(AB)$ and $\Pr(AC)$ not simply be added? (d) Which property characterizes coherent failure logic: that additional component failures cannot reverse an existing system failure, that all components have identical reliability, or that the events are always independent? (e) Why might a static fault tree be insufficient for an emergency backup system? (Answers: (a) AND, since both pump failures are necessary and the top event is their intersection; (b) OR, since either the common supply failure or the failure of both pumps is sufficient; (c) the two terms overlap through A, and the overlap ABC would be counted twice without inclusion–exclusion; (d) the first, since coherent failure logic is monotone in the component failure states, with component relevance included in the strict definition; (e) because the order and timing of failures and of switching can matter, and standby activation, switching delays and event ordering may require dynamic modeling.)

Exercise 10.2. (a) Reduce $F_{\text{top}} = AB + ABC + BC + BCD + DD$ with the laws of Table 10.6, naming the law used at each step, factor the result, and list its minimal cut sets. (b) With independent basic events of probabilities $q_A = 0.10$, $q_B = 0.20$, $q_C = 0.30$ and $q_D = 0.01$, compute Q_{top} exactly. (c) Compute the value obtained by adding $\Pr(AB)$ and $\Pr(BC)$ for the part without D, and explain the difference from the exact probability of that part. (Answers: (a) $DD = D$ by AND idempotence; $AB + ABC = AB$ and $BCD + D = D$ by absorption; $F_{\text{top}} = AB + BC + D = B(A + C) + D$, with minimal cut sets $\{A, B\}$, $\{B, C\}$, $\{D\}$; (b) $\Pr(B(A+C)) = 0.20 \times 0.37 = 0.074$ and $Q_{\text{top}} = 1 - 0.926 \times 0.99 = 0.08326$; (c) $0.02 + 0.06 = 0.08$, which counts the overlap ABC, of probability 0.006, twice.)

Exercise 10.3. Return to the cooling-water tree of Figure 10.4 with the probabilities of Table 10.4. (a) A second, redundant controller is installed, so that the control function is lost only if both controllers fail; each fails with probability 0.02, independently. Write the new Boolean expression and compute Q_{top}. (b) Return to one controller, and suppose instead that the pumps share load, so that $\Pr(B \mid A) = 0.30$. Compute Q_{top}. (c) Compute Q_{top} with both the redundant controller and the load sharing, and say which single change, the controller or the removal of the load-sharing dependence, matters more. (d) For the transmission corridor of Figure 10.8, the two lines share the weather, so that $\Pr(B \mid A) = 0.30$. Compute Q_{top} and compare it with the independent value. (Answers: (a) $AB + E + C_1C_2 + M$, $Q_{\text{top}} = 1 - 0.9975 \times 0.99 \times 0.9996 \times 0.995 = 0.017806$; (b) $\Pr(AB) = 0.015$, $Q_{\text{top}} = 0.049131$; (c) 0.030114; the redundant controller removes the largest single point, and starting from (b) it lowers Q_{top} by 0.0190, more than removing the dependence, which lowers it by 0.0121; (d) $\Pr(AB) = 0.006$, $Q_{\text{top}} = 1 - 0.994 \times 0.99 = 0.01594$, half as much again as 0.010594.)

Exercise 10.4. (a) Two relays fail independently with probabilities 0.10 and 0.20. Compute the probability of the exclusive-OR event "exactly one relay failed", and show from its truth table that the event is noncoherent. (b) For the chiller plant with the voting gate of Figure 10.7(a), compute Q_{top} when each chiller fails with probability 0.10, and compare with 0.00725 at 0.05. (c) For the inhibit gate of Figure 10.7(b), take $\Pr(A) = 0.037064$ and compare $\Pr(A)\Pr(K \mid A)$

for $\Pr(K \mid A) = 0.30$ with the value obtained by wrongly using an unconditional $\Pr(K) = 0.10$. (d) A protection and a primary unit both have $\lambda = 1 \times 10^{-4}\,\mathrm{h}^{-1}$. Over $H = 1000\,\mathrm{h}$, compute the static AND probability that both fail and the priority AND probability that the protection fails first. (Answers: (a) $0.10 + 0.20 - 2 \times 0.02 = 0.26$; from the state $(1, 0)$, with output 1, the further failure to $(1, 1)$ gives output 0; (b) $3 \times 0.01 \times 0.90 + 0.001 = 0.028$, nearly four times 0.00725 for a doubled chiller probability; (c) 0.0111 against 0.0037, an underestimate by a factor of three; (d) $(1 - e^{-0.1})^2 = 0.009056$, and by symmetry the priority AND is half of it, 0.004528, which equation (10.27) confirms.)

Chapter 11

Minimal Cut Sets and Fault-Tree Quantification

11.1 Cut sets and minimal cut sets

Chapter 10 built fault trees from AND and OR gates and wrote each top event as a Boolean expression of basic events. A cooling-water system with two redundant pumps and a shared electrical supply, for example, loses its required function when both pumps fail or when the supply fails,

$$T = AB + E, \tag{11.1}$$

where A is the failure of pump A, B the failure of pump B, E the failure of the common electrical supply, the sign + denotes OR (union) and juxtaposition denotes AND (intersection). For a tree this small the probability of the top event follows in one line. Real fault trees contain hundreds of basic events, components that appear in several branches, failure combinations that overlap, and many layers of gates. This chapter asks

how a complex fault tree can be reduced to its essential failure combinations, and how the probability of system failure can be calculated without counting any event twice.

The question has five parts: which combinations of failures are sufficient to cause the top event; which of them contain no redundant events; how overlapping combinations are quantified correctly; when an approximate calculation is acceptable; and how software analyzes trees too large for hand calculation. Minimal cut-set analysis answers all five. The chapter first defines cut sets and minimal cut sets and derives them by Boolean reduction; it then quantifies the top event exactly by inclusion–exclusion, states the rare-event approximation and the bounds that bracket the exact value, turns to the success view of minimal path sets, describes the computational methods used for large trees, and ends with the fault tree of the data center cooling plant and the limits of static cut-set analysis.

Jargon: cut set, minimal cut set, order, path set, minimal path set, structure function, coherent fault tree, inclusion–exclusion, rare-event approximation, min-cut upper bound, binary decision diagram, truncation

A *cut set* is a set of basic events whose joint occurrence guarantees the top event, whatever the states of the other basic events. A *minimal cut set* is a cut set from which no basic event can be removed without losing that guarantee; its *order* is the number of basic events it contains. A *path set* is a set of basic events whose joint non-occurrence (the corresponding components all function) guarantees that the top event does not occur, and a *minimal path set* is a path set with no removable element. The *structure function* $\phi(\mathbf{x})$ of a fault tree is the indicator of the top event as a function of the vector $\mathbf{x}$ of basic-event indicators, $x_i = 1$ when basic event i occurs. A fault tree is *coherent* when ϕ is nondecreasing in every x_i (an additional failure never restores the function) and every basic event is relevant to it; trees built from AND, OR and voting gates without negations are coherent. *Inclusion–exclusion* is the exact formula for the probability of a union of overlapping events. The *rare-event approximation* replaces the probability of the top event by the sum of the minimal cut-set probabilities. The *min-cut upper bound* is a sharper upper bound on the top-event probability built from the same cut-set probabilities. A *binary decision diagram* (BDD) is a graph that represents a Boolean function by successive decisions on its variables and supports exact quantification. *Truncation* is the deliberate omission of cut sets above a stated order or below a stated probability.

Key idea

A coherent fault tree is exactly the union of its minimal cut sets. The probability of the top event is the probability of that union, which must be computed with the overlaps among cut sets accounted for; the sum of the cut-set probabilities is an upper bound that is accurate only when the overlaps are negligible.

Throughout the chapter a capital letter denotes both a basic event and its indicator variable, an overline its complement ($\overline{A}$: pump A functions), q_i the probability of basic event i over the stated mission or at the stated instant, and Q_{top} the probability of the top event. Unless stated otherwise the basic events are mutually independent.

11.1.1 Cut sets

A cut set is a set of basic failure events whose joint occurrence guarantees the top event, regardless of the states of the other basic events. In terms of the structure function, a set C of basic events is a cut set when $\phi(\mathbf{x}) = 1$ for every state $\mathbf{x}$ with $x_i = 1$ for all $i \in C$. For a coherent tree it is enough to check the single state in which the events of C occur and all others do not, because further failures cannot restore the function.

For the tree of equation (11.1) there are two such sets that are as small as possible,

$$C_1 = \{A, B\}, \qquad C_2 = \{E\} :$$

the first requires both pumps to fail, the second only the electrical supply. Figure 11.1 draws the tree. The box at the top is the top event, loss of cooling. Below it is an OR gate (the shield shape), whose output occurs if any input occurs. Its left input is the intermediate event "both pumps fail", the output of an AND gate (the flat-based shape), whose output occurs only if all inputs occur; its inputs are the basic events A and B, drawn as circles. The right input of the OR gate is the basic event E. The dashed outlines group the basic events of each of the two cut sets.

A cut set need not contain only the failures that are needed. The set $\{A, B, E\}$ also guarantees the top event: if both pumps and the supply have failed, cooling is lost. That set

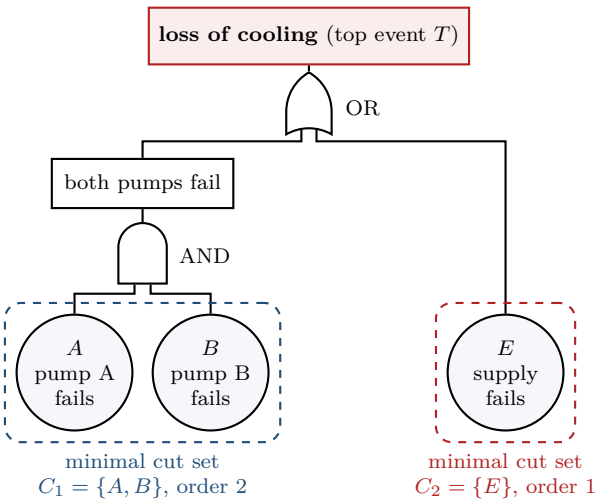


Figure 11.1: Two ways to lose the cooling function, $T = AB + E$. The OR gate below the top event has two inputs: the AND gate of the two pump failures and the supply failure. The dashed outlines are the two minimal cut sets.

contains failures that are not necessary for the guarantee, and the analysis needs a notion that excludes them.

11.1.2 Minimal cut sets and their order

A minimal cut set (MCS) is a cut set from which no basic event can be removed without losing the guarantee that the top event occurs. For $T = AB + E$ the family of minimal cut sets is

$$\mathcal{C}_{\min} = \{\{A, B\},\ \{E\}\}.$$

The set $\{A, B, E\}$ is a cut set but not a minimal cut set, because E alone already causes the system failure: removing A and B from it leaves a cut set. Equivalently, a cut set is minimal when none of its proper subsets is a cut set. Every cut set contains at least one minimal cut set, obtained by removing elements one at a time as long as the remainder is still a cut set.

The order of a minimal cut set is the number of basic events it contains. Table 11.1 lists three orders with their interpretation. A first-order minimal cut set is a single basic event sufficient on its own to cause the system failure, a single point of failure; a second-order set requires a double failure, a third-order set a triple failure. Low-order minimal cut sets often deserve special attention in design reviews, because they show where the redundancy of the system is thinnest.

Table 11.1: Order of a minimal cut set and its interpretation.

Minimal cut set	Order	Interpretation
$\{E\}$	1	Single-point failure
$\{A, B\}$	2	Double failure required
$\{A, B, C\}$	3	Triple failure required

Cut-set order is not a probability and not a risk ranking. With independent basic events, a single-event cut set with probability 10^{-6} is a thousand times less likely than a triple-event cut set whose events each have probability 0.1, since $0.1^3 = 10^{-3}$. A frequent triple-event combination can therefore matter more than an extremely rare single-point failure. Order describes the logical structure of the system; how much each cut set contributes to the probability

of the top event requires the probabilities of its basic events, and the ranking of components by that contribution is the subject of Chapter 12.

Principle 11.1. *Minimality is a logical property, not a probabilistic one. A minimal cut set is a smallest sufficient combination of failures; it is not necessarily probable, and a low order does not by itself make it important.*

11.1.3 Boolean reduction to minimal cut sets

A top event written from a fault tree is usually a sum of products that are cut sets but not all minimal. Consider

$$T = AB + AC + ABC + D, \tag{11.2}$$

in which each product is a candidate combination of basic events that causes the top event. Two identities of Boolean algebra remove the redundant products. Idempotence states that $XX = X$ and $X + X = X$: an event ANDed or ORed with itself is the same event, so a basic event that reaches a product twice through different gates counts once. Absorption states that

$$X + XY = X, \tag{11.3}$$

which follows from $X + XY = X(1 + Y) = X \cdot 1 = X$, since $1 + Y = 1$ for any event Y. In words: if X alone causes the top event, adding the case in which Y also occurs adds nothing. The identities hold for events as sets and require no assumption of probabilistic independence.

With $X = AB$ and $Y = C$, absorption gives $AB + ABC = AB$, and the product ABC is eliminated,

$$T = AB + AC + D = A(B + C) + D,$$

where the second form factors out A. The minimal cut sets are

$$\mathcal{C}_{\min} = \{\{A, B\},\ \{A, C\},\ \{D\}\}.$$

The combination $\{A, B, C\}$ is not minimal, because either $\{A, B\}$ or $\{A, C\}$ is sufficient. The factored form reads directly as an engineering statement: the system fails if A occurs together with either B or C, or if D occurs alone.

What a minimal cut set does not say. A minimal cut set identifies a combination of events sufficient for failure. It does not say that the events in it are statistically independent, and it does not say when they occur or in what order. These are properties of the probability model and of the dynamics of the system, which the logic leaves open and which Section 11.6.6 returns to.

The top event as a union of minimal cut sets. For a conventional static, coherent fault tree with minimal cut sets $C_1, \ldots, C_m$, let C_j also denote the event that all basic events of the j-th set occur, $C_j = \bigcap_{i \in C_j} X_i$, where X_i is basic event i. Then

$$T = \bigcup_{j=1}^{m} C_j, \qquad \phi(\mathbf{x}) = \max_j \prod_{i \in C_j} x_i. \tag{11.4}$$

The derivation takes two steps. If some C_j has fully occurred, the top event occurs, by the definition of a cut set. Conversely, if the top event occurs in a state $\mathbf{x}$, the set of all basic events that have occurred in $\mathbf{x}$ is a cut set, because the tree is coherent, and it contains at least one minimal cut set, which has therefore fully occurred. The representation (11.4) is exact for the complete list of minimal cut sets of the specified coherent tree. A tree with NOT gates, which is not coherent, needs a more general list of minimal combinations that include non-occurrences (prime implicants); the trees of this chapter are coherent.

11.2 Exact quantification of the top event

Once the top event is a union of minimal cut-set events, its probability is the probability of a union. This section shows why that probability is not the sum of the cut-set probabilities, states the inclusion–exclusion formula for any number of events, and works an example with three overlapping cut sets by two exact methods.

11.2.1 Overlapping cut-set events

Return to $T = AB + AC + D$, with the cut-set events

$$C_1 = AB, \qquad C_2 = AC, \qquad C_3 = D,$$

(the cut-set event C_2 is not the basic event C; the subscript keeps them apart). The probability of the top event is $Q_{\text{top}} = \Pr(C_1 \cup C_2 \cup C_3)$. Summing $\Pr(C_1) + \Pr(C_2) + \Pr(C_3)$ is in general wrong, because the cut-set events overlap. The intersection $C_1 \cap C_2 = ABC$ is not empty: both cut sets contain the basic event A, and if A, B and C all occur, both cut sets are satisfied. Adding their probabilities counts that outcome twice. The basic event D is independent of A, B and C, so C_3 overlaps both of the others as well.

Figure 11.2 shows the overlaps as a Venn diagram. The rectangle is the set of all sixteen joint states of the four basic events; each circle is the set of states in which one cut-set event occurs. The number in each region is its probability for the illustrative values $q_A = 0.10$, $q_B = 0.20$, $q_C = 0.30$ and $q_D = 0.01$ of Section 11.2.3; for example, the region inside C_1 only is the event $AB\overline{C}\,\overline{D}$, with probability $0.10 \times 0.20 \times 0.70 \times 0.99 = 0.01386$. The seven regions inside the circles sum to the top-event probability, 0.05356. The sum of the three circle probabilities, $0.02 + 0.03 + 0.01 = 0.06$, counts each of the three regions where exactly two circles overlap twice and the central region three times: the excess is $0.00594 + 0.00014 + 0.00024 + 2 \times 0.00006 = 0.00644$, and $0.05356 + 0.00644 = 0.06$.

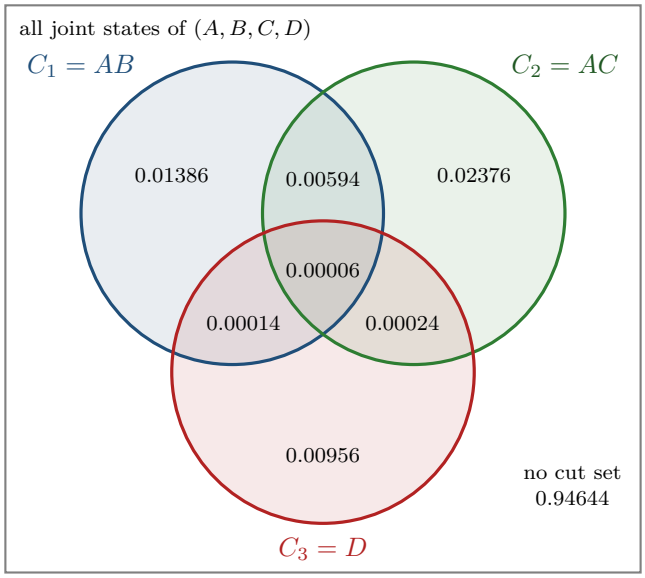


Figure 11.2: Overlap of the three cut-set events of $T = AB + AC + D$ for the illustrative probabilities $q_A = 0.10$, $q_B = 0.20$, $q_C = 0.30$, $q_D = 0.01$, independent. Each number is the probability of its region. The seven regions inside the circles sum to $Q_{\text{top}} = 0.05356$; the circle probabilities 0.02, 0.03 and 0.01 sum to 0.06 because the overlaps are counted more than once.

Common mistake

Computing the top-event probability as the sum of the minimal cut-set probabilities and calling it exact. Cut sets that share a basic event, or that can occur together, overlap; their sum counts the overlapping states more than once. The sum is an upper bound (Section 11.3.1), and it is exact only when the cut-set events are mutually exclusive.

11.2.2 Inclusion–exclusion

The exact probability of a union is given by inclusion–exclusion. For two events,

$$\Pr(C_1 \cup C_2) = \Pr(C_1) + \Pr(C_2) - \Pr(C_1 \cap C_2), \tag{11.5}$$

and for three,

$$\begin{aligned}\Pr(C_1 \cup C_2 \cup C_3) = {} & \Pr(C_1) + \Pr(C_2) + \Pr(C_3) \\ & - \Pr(C_1 \cap C_2) - \Pr(C_1 \cap C_3) - \Pr(C_2 \cap C_3) \\ & + \Pr(C_1 \cap C_2 \cap C_3).\end{aligned} \tag{11.6}$$

For m events the probabilities of the intersections of increasing order are alternately added and subtracted. Write

$$S_k = \sum_{1 \le j_1 < j_2 < \cdots < j_k \le m} \Pr(C_{j_1} \cap C_{j_2} \cap \cdots \cap C_{j_k}) \tag{11.7}$$

for the sum over all intersections of k distinct events; then

$$Q_{\text{top}} = \Pr\Big(\bigcup_{j=1}^{m} C_j\Big) = \sum_{k=1}^{m} (-1)^{k-1} S_k = S_1 - S_2 + S_3 - \cdots + (-1)^{m-1} S_m. \tag{11.8}$$

The formula follows from indicators. The top event does not occur exactly when no cut-set event occurs, so with $\mathbf{1}_{C_j}$ the indicator of C_j,

$$1 - \mathbf{1}_T = \prod_{j=1}^{m} (1 - \mathbf{1}_{C_j}).$$

Expanding the product over all subsets J of $\{1, \ldots, m\}$ gives $\sum_J (-1)^{|J|} \prod_{j \in J} \mathbf{1}_{C_j}$, and the product of indicators is the indicator of the intersection, $\prod_{j \in J} \mathbf{1}_{C_j} = \mathbf{1}_{\cap_{j \in J} C_j}$. Moving the empty subset ($J = \emptyset$, term 1) to the left side, changing signs, and taking expectations, with $\mathbb{E}[\mathbf{1}_E] = \Pr(E)$ for any event E, gives equation (11.8). No independence is used: the formula is an identity of probability.

Intersections of cut-set events. The intersection of several cut-set events is the event that every basic event in the union of their sets occurs. By idempotence a basic event shared by two cut sets appears once, so for independent basic events

$$\Pr(C_{j_1} \cap \cdots \cap C_{j_k}) = \prod_{i \in C_{j_1} \cup \cdots \cup C_{j_k}} q_i. \tag{11.9}$$

The probability of an intersection is therefore not the product of the cut-set probabilities when the sets share events: $\Pr(C_1 \cap C_2) = \Pr(ABC) = q_A q_B q_C$, not $\Pr(C_1)\Pr(C_2) = q_A^2 q_B q_C$.

The cost of the expansion. Equation (11.8) is an exact basis for fault-tree quantification, but S_k has $\binom{m}{k}$ terms, and the whole expansion has $\sum_{k=1}^{m} \binom{m}{k} = 2^m - 1$ intersection terms. Three cut sets need 7 terms, the six minimal cut sets of the cooling plant of Section 11.6 need 63, and twenty cut sets need more than a million. A direct expansion becomes expensive for large fault trees, which is why the approximations and bounds of Section 11.3 and the algorithms of Section 11.5 exist.

11.2.3 Three overlapping minimal cut sets

Consider $T = AB + AC + D$ with mutually independent basic events and the illustrative probabilities

$$q_A = 0.10, \qquad q_B = 0.20, \qquad q_C = 0.30, \qquad q_D = 0.01.$$

Method 1: Boolean factoring. Factor A, $T = A(B + C) + D$. The probability that B or C occurs is

$$\Pr(B + C) = q_B + q_C - q_B q_C = 0.20 + 0.30 - 0.06 = 0.44.$$

Since A is independent of B and C,

$$\Pr(A(B + C)) = q_A \times 0.44 = 0.10 \times 0.44 = 0.044.$$

The event $A(B + C)$ depends only on A, B and C, and D is independent of all three, so the top event is the union of two independent events, and its probability is one minus the probability that neither occurs:

$$Q_{\text{top}} = 1 - (1 - 0.044)(1 - 0.01) = 1 - 0.956 \times 0.99 = 1 - 0.94644 = 0.05356.$$

The top-event probability is 5.356 %.

Method 2: inclusion–exclusion. The cut-set probabilities are

$$\Pr(AB) = 0.10 \times 0.20 = 0.02, \qquad \Pr(AC) = 0.10 \times 0.30 = 0.03, \qquad \Pr(D) = 0.01,$$

so $S_1 = 0.06$. The pairwise intersections, by equation (11.9), are

$$\begin{aligned} \Pr(C_1C_2) &= \Pr(ABC) = 0.10 \times 0.20 \times 0.30 = 0.006, \\ \Pr(C_1C_3) &= \Pr(ABD) = 0.10 \times 0.20 \times 0.01 = 0.0002, \\ \Pr(C_2C_3) &= \Pr(ACD) = 0.10 \times 0.30 \times 0.01 = 0.0003, \end{aligned}$$

so $S_2 = 0.0065$; and the triple intersection is $\Pr(ABCD) = 0.00006 = S_3$. Then

$$Q_{\text{top}} = S_1 - S_2 + S_3 = 0.06 - 0.0065 + 0.00006 = 0.05356,$$

the same value as Method 1, as it must be.

The simple sum $S_1 = 0.06$ overestimates the system failure probability by $0.06 - 0.05356 = 0.00644$, that is, by 0.644 percentage points, or $0.00644/0.05356 = 12.0\,\%$ of the exact value. In this example the overestimate has a closed form: from equation (11.8), $S_1 - Q_{\text{top}} = S_2 - S_3$, and collecting the terms,

$$S_1 - Q_{\text{top}} = q_A q_B q_C + q_A q_D \left(q_B + q_C - q_B q_C\right) = 0.006 + 0.10 \times 0.01 \times 0.44 = 0.00644. \quad (11.10)$$

The first term is the overlap of the two cut sets that share A; the second is the overlap of D with the event $A(B + C)$. Overlapping cut sets cannot in general be treated as mutually exclusive.

> **Principle 11.2.** *The probability of the top event is the probability of the union of the minimal cut-set events. Exact quantification must account for every overlap among them, through inclusion–exclusion, through a factored Boolean expression, or through one of the algorithms of Section 11.5.*

11.3 Rare-event approximation and bounds

The exact union probability is often expensive, and in many applications the cut sets are so improbable that their overlaps are negligible. This section states the rare-event approximation, measures its error on a range of basic-event probabilities, and gives the bounds that bracket the exact value whatever the probabilities are.

11.3.1 The rare-event approximation

In many reliability applications individual failure combinations are unlikely; the probability that two independent high-reliability components both fail during a short mission is very small. If the minimal cut-set events are rare enough, the probability that two or more of them occur together is negligible, the terms $S_2, S_3, \ldots$ of equation (11.8) are small against S_1, and

$$Q_{\text{top}} \approx S_1 = \sum_{j=1}^{m} \Pr(C_j). \tag{11.11}$$

This is the rare-event approximation. When each cut-set probability is computed as the product of independent basic-event probabilities, the approximation is a sum of products, $\sum_j \prod_{i \in C_j} q_i$, and it is also called the sum-of-products approximation.

The sum is always an upper bound,

$$Q_{\text{top}} \leq \sum_{j=1}^{m} \Pr(C_j), \tag{11.12}$$

by the union bound: the indicator of a union is at most the sum of the indicators, $\mathbf{1}_T \leq \sum_j \mathbf{1}_{C_j}$, because the right side counts the number of cut sets that have occurred, which is at least one whenever the top event occurs; taking expectations gives the inequality. The bound holds for any events, independent or not, and whether or not the tree is coherent; coherence is needed only for the union of minimal cut sets to be the top event. Summing all minimal cut-set probabilities therefore gives an upper bound even when the approximation is not accurate, provided that no cut set has been left out.

Rare basic events. Take the same tree, $T = AB + AC + D$, with probabilities ten times smaller,

$$q_A = 0.01, \qquad q_B = 0.02, \qquad q_C = 0.03, \qquad q_D = 0.001,$$

independent. The cut-set probabilities are $\Pr(AB) = 0.0002$, $\Pr(AC) = 0.0003$ and $\Pr(D) = 0.001$, and the rare-event approximation is

$$Q_{\text{top}} \approx 0.0002 + 0.0003 + 0.001 = 0.001500.$$

The exact value, by factoring as in Method 1,

$$\begin{aligned} \Pr(B + C) &= 0.02 + 0.03 - 0.02 \times 0.03 = 0.0494, \\ \Pr\big(A(B + C)\big) &= 0.01 \times 0.0494 = 0.000494, \\ Q_{\text{top}} &= 1 - (1 - 0.000494)(1 - 0.001) = 0.001493506. \end{aligned}$$

The exact probability is 0.14935 % against 0.15000 % for the approximation, a relative error of $(0.0015 - 0.001493506)/0.001493506 = 0.43\,\%$. For this system the approximation is excellent. By equation (11.10), the error is $q_A q_B q_C + q_A q_D (q_B + q_C - q_B q_C)$, a sum of products of three or more basic-event probabilities, while the cut-set probabilities themselves are products of at most two; dividing every probability by ten divides the error by a thousand or more and the cut-set probabilities by at most a hundred.

A small top-event probability does not by itself guarantee a small relative error. The error is set by the overlap of the cut sets, and when the overlap is strong, as when two cut sets share a basic event and their other events are likely, the approximation can be poor even for a small Q_{top}. Section 11.3.2 gives an example.

11.3.2 Accuracy of the approximation

How the error depends on the basic-event probabilities is shown by computing, for the tree $T = AB + AC + D$ with independent basic events, the exact probability, the rare-event approximation and a second upper bound, the min-cut upper bound

$$Q_{\text{MCUB}} = 1 - \prod_{j=1}^{m} (1 - \Pr(C_j)), \tag{11.13}$$

which Section 11.3.3 derives. For this tree,

$$\begin{aligned} Q_{\text{top}} &= q_A(q_B + q_C - q_B q_C)(1 - q_D) + q_D, \\ S_1 &= q_A q_B + q_A q_C + q_D, \\ Q_{\text{MCUB}} &= 1 - (1 - q_A q_B)(1 - q_A q_C)(1 - q_D), \end{aligned}$$

and expanding the last expression and subtracting the first gives its error in closed form, $Q_{\text{MCUB}} - Q_{\text{top}} = q_A q_B q_C (1 - q_A)(1 - q_D)$: the min-cut upper bound removes the overlap with D entirely (the cut set $\{D\}$ shares no event with the others) and keeps only a reduced share of the overlap of the two cut sets that share A.

Table 11.2 lists the three quantities for five sets of basic-event probabilities. Each row gives the four probabilities, the exact value, the rare-event sum S_1 and its relative overestimate, $(S_1 - Q_{\text{top}})/Q_{\text{top}}$, the min-cut upper bound and its relative overestimate, and the lower bound $S_1 - S_2$ of Section 11.3.3. The second and third rows are the two cases worked above; the first divides the probabilities of the second by ten again; the fourth raises them; the fifth is a case of strong overlap, a rare event A shared by two cut sets whose other events are likely. For the fourth row the arithmetic is

$$\begin{aligned} Q_{\text{top}} &= 0.30\,(0.40 + 0.50 - 0.20)(1 - 0.05) + 0.05 = 0.30 \times 0.70 \times 0.95 + 0.05 = 0.2495, \\ S_1 &= 0.12 + 0.15 + 0.05 = 0.32, \quad \text{error } (0.32 - 0.2495)/0.2495 = 28.3\,\%, \\ Q_{\text{MCUB}} &= 1 - 0.88 \times 0.85 \times 0.95 = 0.2894, \quad \text{error } (0.2894 - 0.2495)/0.2495 = 16.0\,\%. \end{aligned}$$

Table 11.2: Exact top-event probability of $T = AB + AC + D$ against the rare-event approximation S_1, the min-cut upper bound Q_{MCUB} and the lower bound $S_1 - S_2$, for independent basic events (illustrative probabilities). The percentages are relative overestimates of the exact value.

q_A	q_B	q_C	q_D	Q_{top}	S_1	error	Q_{MCUB}	error	$S_1 - S_2$
0.001	0.002	0.003	0.0001	0.000104994	0.000105	0.006 %	0.000104999	0.006 %	0.000104994
0.01	0.02	0.03	0.001	0.00149351	0.00150	0.43 %	0.00149944	0.40 %	0.00149350
0.10	0.20	0.30	0.01	0.05356	0.06	12.0 %	0.058906	9.98 %	0.0535
0.30	0.40	0.50	0.05	0.2495	0.32	28.3 %	0.2894	16.0 %	0.2465
0.01	0.50	0.50	0.0001	0.00759925	0.0101	32.9 %	0.010074	32.6 %	0.007599

Three trends appear. First, along the first four rows the error of both upper bounds grows with the probabilities, from 0.006 % to more than 15 %, while the lower bound $S_1 - S_2$ stays within 1.2 % of the exact value. Second, the min-cut upper bound is never worse than the rare-event sum, and in the fourth row it removes nearly half of the error. Third, the fifth row

has a top-event probability of only 0.0076, smaller than that of the original case, yet both upper bounds overestimate it by a third: the two cut sets $\{A, B\}$ and $\{A, C\}$ share the rare event A, and given A the events B and C occur together with probability 0.25, half of the probability 0.5 of each, so half of the probability of each of the two cut sets lies in their overlap. In that case the lower bound $S_1 - S_2 = 0.007599$ is within 0.01 % of the exact value. The rule is to calculate or bound the error of an approximation, rather than to assume that the word "rare" guarantees accuracy.

Figure 11.3 shows the trend continuously. The horizontal axis is a common multiplier s applied to the original probabilities, $(q_A, q_B, q_C, q_D) = s\,(0.10, 0.20, 0.30, 0.01)$, on a logarithmic scale from 0.01 to 3; the vertical axis is the relative overestimate in percent, also logarithmic. The solid curve is the rare-event sum, the dashed curve the min-cut upper bound; the marked points are the cases $s = 0.1$ and $s = 1$ of the table. For small s both errors fall roughly as s^2: the overlap terms are products of three or four basic-event probabilities, of order s^3, while the exact value is dominated by the first-order cut set $\{D\}$, of order s. Both rise steeply as s approaches one, where the basic-event probabilities are no longer small.

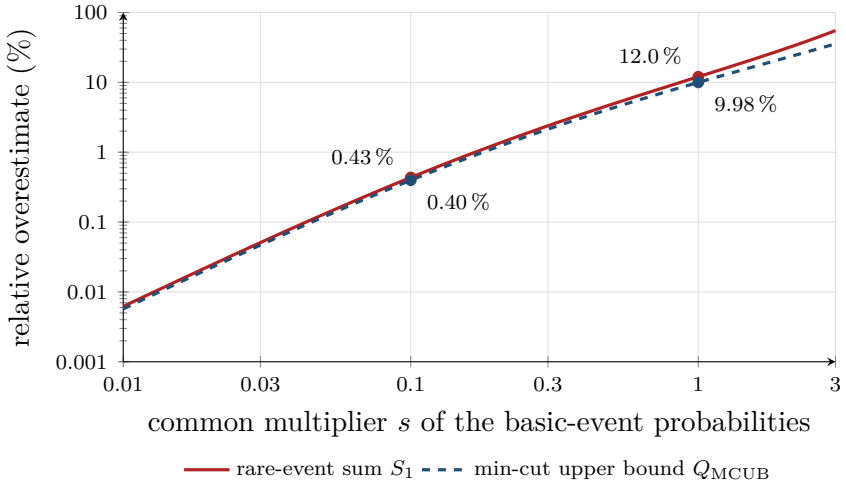


Figure 11.3: Relative overestimate of the exact top-event probability of $T = AB + AC + D$ by the rare-event sum, $(S_1 - Q_{\text{top}})/Q_{\text{top}}$, and by the min-cut upper bound, for independent basic events with $(q_A, q_B, q_C, q_D) = s\,(0.10, 0.20, 0.30, 0.01)$. Exact: $Q_{\text{top}} = q_A(q_B + q_C - q_B q_C)(1 - q_D) + q_D$; overestimates: $q_A q_B q_C + q_A q_D(q_B + q_C - q_B q_C)$ and $q_A q_B q_C(1 - q_A)(1 - q_D)$. Marked: $s = 0.1$ and $s = 1$.

11.3.3 Bounds on the top-event probability

Even when exact quantification is expensive, the top-event probability can be bracketed. For cut-set events $C_1, \ldots, C_m$,

$$\max_j \Pr(C_j) \le Q_{\text{top}} \le \min\Bigl(1, \sum_j \Pr(C_j)\Bigr). \tag{11.14}$$

The lower bound holds because the union contains every individual cut-set event; the upper bound is the union bound of equation (11.12), capped at one because it is a probability.

The second-order lower bound. A sharper lower bound is obtained from the first two terms of inclusion–exclusion,

$$Q_{\text{top}} \ge S_1 - S_2 = \sum_j \Pr(C_j) - \sum_{i<j} \Pr(C_i \cap C_j). \tag{11.15}$$

It is one of the Bonferroni inequalities, a standard result: the partial sums of equation (11.8) alternate around the exact value, $S_1 \ge Q_{\text{top}}$, $S_1 - S_2 \le Q_{\text{top}}$, $S_1 - S_2 + S_3 \ge Q_{\text{top}}$, and so on. When the probabilities are large, $S_1 - S_2$ can fall below the largest single cut-set probability, or below zero, so the lower bound actually used is the larger of the two, $\max\bigl(\max_j \Pr(C_j),\ S_1 - S_2\bigr)$. For the original case of Section 11.2.3, $\max_j \Pr(C_j) = 0.03$ and $S_1 - S_2 = 0.0535$, so the exact value 0.05356 lies in $[0.0535,\ 0.06]$, an interval only 0.0065 wide.

The min-cut upper bound. For a coherent fault tree with independent basic events, a sharper upper bound than the sum is

$$Q_{\text{top}} \le Q_{\text{MCUB}} = 1 - \prod_{j=1}^{m} \bigl(1 - \Pr(C_j)\bigr), \qquad \Pr(C_j) = \prod_{i \in C_j} q_i, \tag{11.16}$$

called the min-cut upper bound (also the Esary–Proschan upper bound). It would be the exact probability if the cut-set events were independent, and it is an upper bound because they are positively correlated. The derivation uses a standard result on monotone events: if the basic events are independent, any two events that are both nonincreasing functions of the basic-event indicators (events that additional failures can only destroy) are positively correlated, and so is any finite collection of them, $\Pr(\bigcap_j G_j) \ge \prod_j \Pr(G_j)$. The events $G_j = \overline{C_j}$, "cut set j has not fully occurred", are of that kind. Hence

$$1 - Q_{\text{top}} = \Pr\Bigl(\bigcap_j \overline{C_j}\Bigr) \ge \prod_j \Pr\bigl(\overline{C_j}\bigr) = \prod_j \bigl(1 - \Pr(C_j)\bigr),$$

which is equation (11.16). The bound is exact when no two minimal cut sets share a basic event, because the cut-set events are then independent. It is never larger than the sum: for numbers x_j in $[0,1]$, $1 - \prod_j (1 - x_j) \le \sum_j x_j$, which follows by induction from $1 - (1-y)(1-x) = y + x - xy \le y + x$. For independent basic events, therefore,

$$\max\bigl(\max_j \Pr(C_j),\ S_1 - S_2\bigr) \le Q_{\text{top}} \le Q_{\text{MCUB}} \le S_1.$$

For the original case, $Q_{\text{MCUB}} = 1 - (1-0.02)(1-0.03)(1-0.01) = 1 - 0.98 \times 0.97 \times 0.99 = 1 - 0.941094 = 0.058906$, and the chain reads $0.0535 \le 0.05356 \le 0.058906 \le 0.06$. Unlike the sum, the min-cut upper bound requires independent basic events; with positively dependent basic events, such as common-cause failures, it can fall below the exact value.

A practical strategy. The bounds give a procedure for large trees:

1. calculate the individual cut-set probabilities;
2. identify the cut sets that overlap importantly, those that share basic events and have large probabilities;
3. calculate their pairwise intersections;
4. determine whether the resulting bounds are narrow enough for the decision at hand;
5. use exact quantification if more precision is necessary.

For high-consequence engineering decisions a bound may be more defensible than an approximation reported with unjustified precision.

11.4 Minimal path sets

The analysis so far describes the system through its failures. The same logic can be read from the success side, through the combinations of functioning components that guarantee the required function, and the two views must agree. This section derives minimal path sets by De Morgan's laws and uses them to compute the system reliability.

11.4.1 Success paths and minimal path sets

A minimal path set is a minimal set of basic component-success events whose joint occurrence guarantees system success: if every component in the set functions, the top event cannot occur, whatever the states of the others, and no component can be removed from the set without losing that guarantee. Consider again the two redundant pumps with a common electrical supply, $T = AB + E$, with A and B the pump failures and E the supply failure. System success is the complement of the top event, $\overline{T}$. De Morgan's laws, $\overline{X + Y} = \overline{X}\,\overline{Y}$ and $\overline{XY} = \overline{X} + \overline{Y}$, give

$$\overline{T} = \overline{AB + E} = \overline{AB}\;\overline{E} = (\overline{A} + \overline{B})\,\overline{E},$$

and distributing,

$$\overline{T} = \overline{A}\,\overline{E} + \overline{B}\,\overline{E}. \tag{11.17}$$

The minimal path sets are

$$P_1 = \{\overline{A}, \overline{E}\}, \qquad P_2 = \{\overline{B}, \overline{E}\} :$$

pump A works and the electrical supply works, or pump B works and the electrical supply works. The general recipe is the same: complement the minimal cut-set expression, apply De Morgan's laws, expand into a sum of products of complemented events, and reduce by absorption.

Figure 11.4 draws the success expression as a reliability block diagram of the kind of Chapter 9. Signal flows from left to right; each block is a component that passes the flow when it functions. The two pumps are in parallel, and the electrical supply is in series with the pair. A path from the input to the output through functioning blocks exists exactly when one of the two minimal path sets has fully occurred, so the diagram is equivalent to equation (11.17). The electrical supply appears in both minimal path sets because either pump needs it to provide the function.

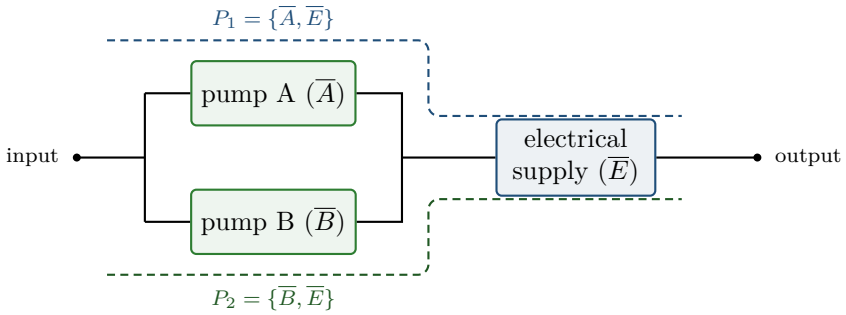


Figure 11.4: Reliability block diagram of the success expression $\overline{T} = \overline{A}\,\overline{E} + \overline{B}\,\overline{E}$: either pump, plus the common supply. The dashed lines trace the two minimal path sets.

Table 11.3 compares the two notions. The minimal cut set describes failure: the included components failing is sufficient, removing any one of the failures removes the guarantee of system failure, and its typical use is vulnerability and risk analysis. The minimal path set describes success: the included components succeeding is sufficient, removing any one removes the guarantee of system success, and its typical use is reliability architecture and the identification of success pathways.

Duality. Every minimal path set shares at least one component with every minimal cut set. If a path set and a cut set had no component in common, the state in which all components of the path set function and all components of the cut set fail would be possible, and it would be both a success and a failure of the system. In the example, E belongs to both path sets and to the cut set $\{E\}$, and each path set contains one of the pumps of the cut set $\{A, B\}$. For the tree $T = AB + AC + D$, the same steps give $\overline{T} = (\overline{A} + \overline{B}\,\overline{C})\,\overline{D}$ and the minimal path sets $\{\overline{A}, \overline{D}\}$ and $\{\overline{B}, \overline{C}, \overline{D}\}$.

Table 11.3: Minimal cut sets and minimal path sets compared.

Property	Minimal cut set	Minimal path set
Describes	Failure	Success
Sufficient condition	Included components fail	Included components succeed
Minimality	Removing any failure removes the guarantee of system failure	Removing any success removes the guarantee of system success
Typical use	Vulnerability and risk analysis	Reliability architecture and success pathways

11.4.2 Exact reliability from minimal path sets

For $\overline{T} = \overline{E}(\overline{A} + \overline{B})$ with independent basic failure events, the supply works with probability $1 - q_E$, at least one pump works with probability $1 - q_A q_B$, and the two events are independent, so

$$R_{\text{system}} = \Pr(\overline{T}) = (1 - q_E)(1 - q_A q_B). \tag{11.18}$$

With the illustrative values $q_A = q_B = 0.05$ and $q_E = 0.01$,

$$R_{\text{system}} = 0.99 \times (1 - 0.0025) = 0.99 \times 0.9975 = 0.987525,$$

and the corresponding failure probability is $Q_{\text{top}} = 1 - 0.987525 = 0.012475$, or $1.2475\,\%$. The same value follows by inclusion–exclusion over the two path-set events, $\Pr(P_1) + \Pr(P_2) - \Pr(P_1 \cap P_2) = 0.99 \times 0.95 + 0.99 \times 0.95 - 0.99 \times 0.95^2 = 1.881 - 0.893475 = 0.987525$, where the intersection contains $\overline{E}$ once, by idempotence.

The same result follows directly from the cut sets,

$$Q_{\text{top}} = \Pr(AB + E) = \Pr(AB) + \Pr(E) - \Pr(AB)\Pr(E) = 0.0025 + 0.01 - 0.000025 = 0.012475.$$

Both calculations agree. Success-path analysis and failure-cut analysis must produce consistent probabilities when they represent the same system under the same assumptions, which makes each an independent check on the other.

> **Principle 11.3.** *The minimal cut sets and the minimal path sets of a system are two descriptions of one structure function. Computed under the same assumptions, the failure probability from the cut sets and one minus the reliability from the path sets must agree; a disagreement shows an error in the logic or in the arithmetic.*

The min-path lower bound. Path sets also give a lower bound on the top-event probability, the counterpart of the min-cut upper bound. With independent basic events, the events "path set j is not fully functioning" are nondecreasing in the failures, hence positively correlated, and the top event is their intersection, so

$$Q_{\text{top}} \geq \prod_j \Big(1 - \prod_{i \in P_j} (1 - q_i)\Big). \tag{11.19}$$

For $T = AB + E$ the bound is $(1 - 0.95 \times 0.99)^2 = 0.0595^2 = 0.00354$, well below the exact 0.012475, because the two path sets share the supply; for $T = AB + AC + D$ with the original values it is $(1 - 0.90 \times 0.99)(1 - 0.80 \times 0.70 \times 0.99) = 0.109 \times 0.4456 = 0.04857$, against the exact 0.05356. With the min-cut upper bound, $0.04857 \leq 0.05356 \leq 0.058906$. The path-set bound is useful when the system is highly unreliable; for reliable systems the cut-set bounds are the sharper ones.

11.5 Computational methods for large fault trees

Hand reduction works for trees of a few events; industrial trees need algorithms. This section explains why the problem grows quickly, and describes four computational approaches, Boolean expansion with absorption, binary decision diagrams, state enumeration and Monte Carlo simulation, before comparing them.

11.5.1 Size of the state space

Consider a fault tree with 100 independent basic events. Each event has two states, failed or not failed, so the system has $2^{100} \approx 1.27 \times 10^{30}$ configurations. Direct enumeration of that state space is impractical: even at 10^9 configurations per second it would take about 4×10^{13} years. With 20 basic events the count is $2^{20} \approx 1.05 \times 10^6$, which is easy; the difficulty grows exponentially with the number of events, not with the size of the system as an engineer perceives it.

A fault tree, however, has a logical structure that can be exploited. The expression $T = AB + AC$ contains three basic events but factors as $T = A(B + C)$, which evaluates the probability without processing the overlapping combinations AB and AC separately and then correcting for their overlap. Every method below exploits structure of this kind, or replaces exhaustive computation by sampling.

11.5.2 Boolean expansion and absorption

One approach constructs the minimal cut sets recursively from the gates of the tree. Each basic event is represented as a family containing one cut set, the singleton $\{X\}$. An AND gate combines the failure requirements of its inputs: every cut set of one input is joined with every cut set of the other. An OR gate collects the alternative sufficient combinations of its inputs. After each operation, any candidate cut set that contains another candidate as a proper subset is removed, which is absorption, equation (11.3), and repeated events inside a candidate are merged, which is idempotence.

Example. Consider $T = (A + B)(C + D)$. Distributing the product,

$$T = AC + AD + BC + BD,$$

and the minimal cut sets are $\{A, C\}$, $\{A, D\}$, $\{B, C\}$ and $\{B, D\}$. Now let an additional branch, the event A alone, be ORed into the top event,

$$T = A + (A + B)(C + D) = A + AC + AD + BC + BD.$$

Absorption gives $A + AC = A$ and $A + AD = A$, so

$$T = A + BC + BD,$$

and the minimal cut sets become $\{A\}$, $\{B, C\}$ and $\{B, D\}$. The cut sets that contained A together with other events were removed because A alone now guarantees the top event.

The basic algorithm. With each cut set represented as a set of basic-event identifiers, and each gate's output represented as a family of such sets, the procedure is:

1. For every basic event X, form the family $\{\{X\}\}$.
2. Process the gates from the bottom of the tree upward, so that the families of a gate's inputs are known when the gate is reached.
3. For an OR gate with input families $\mathcal{F}_1$ and $\mathcal{F}_2$, form the union $\mathcal{F}_1 \cup \mathcal{F}_2$.

4. For an AND gate, form every pairwise union $\{S_1 \cup S_2 : S_1 \in \mathcal{F}_1,\ S_2 \in \mathcal{F}_2\}$; the set union merges a basic event that appears in both.
5. For a gate with more than two inputs, apply step 3 or step 4 to the first two inputs and then to the result and each further input.
6. After every gate, minimize the family: remove duplicates, and remove every set that contains another set of the family as a proper subset.
7. The minimized family at the top gate is the list of minimal cut sets.

Applied to $T = A + (A + B)(C + D)$: the OR gates give $\{\{A\}, \{B\}\}$ and $\{\{C\}, \{D\}\}$; the AND gate gives $\{\{A, C\}, \{A, D\}, \{B, C\}, \{B, D\}\}$; the top OR gate adds $\{A\}$, and minimization removes $\{A, C\}$ and $\{A, D\}$, leaving $\{\{A\}, \{B, C\}, \{B, D\}\}$. Applied to $AB + AC$, the AND gates give $\{\{A, B\}\}$ and $\{\{A, C\}\}$, and the OR gate their union, which is already minimal.

A practical implementation adds a canonical ordering of the events inside each set, efficient duplicate elimination and fast subset tests. The simple algorithm is appropriate for understanding the logic but not necessarily for industrial-scale fault trees: an AND gate whose inputs have n_1 and n_2 cut sets produces up to $n_1 n_2$ candidates, and the intermediate families can grow combinatorially before minimization removes most of them. Large analyses therefore combine the expansion with truncation, Section 11.6.6.

11.5.3 Binary decision diagrams

A binary decision diagram (BDD) represents a Boolean function as a graph of decisions about basic events. Each internal node tests one variable, which is either false (the event does not occur) or true (it occurs); its two branches lead to other decision nodes or to one of two terminal nodes, system success (0) and system failure (1). A basic event that appears in several branches of the fault tree is one variable of the diagram, tested once on any path, so repeated events are represented consistently.

Shannon expansion. For any Boolean function F and any basic-event variable X,

$$F = \overline{X}\, F_{X=0} + X\, F_{X=1}, \tag{11.20}$$

where $F_{X=0}$ is the function with X set to false (the event does not occur) and $F_{X=1}$ the function with X set to true. The identity is checked by cases: if $X = 1$, the right side reduces to $F_{X=1}$, which is F evaluated at $X = 1$; if $X = 0$, it reduces to $F_{X=0}$. The two branches are mutually exclusive, because X cannot equal zero and one at once, so their probabilities add. If X is independent of the other basic events, $F_{X=0}$ and $F_{X=1}$, which do not depend on X, are independent of X, and

$$\Pr(F) = (1 - q_X)\ \Pr(F_{X=0}) + q_X\ \Pr(F_{X=1}). \tag{11.21}$$

This is an exact probability recursion: applied to $F_{X=0}$ and $F_{X=1}$ with the next variable, and so on, it ends at the constants 0 and 1. For dependent basic events, the branch probabilities must instead be evaluated conditionally, $\Pr(F) = \Pr(\overline{X})\ \Pr(F_{X=0} \mid \overline{X}) + \Pr(X)\ \Pr(F_{X=1} \mid X)$.

Worked example. Let $F = AB + AC$ and choose A as the first decision variable. If $A = 0$, $F_{A=0} = 0$; if $A = 1$, $F_{A=1} = B + C$. Therefore

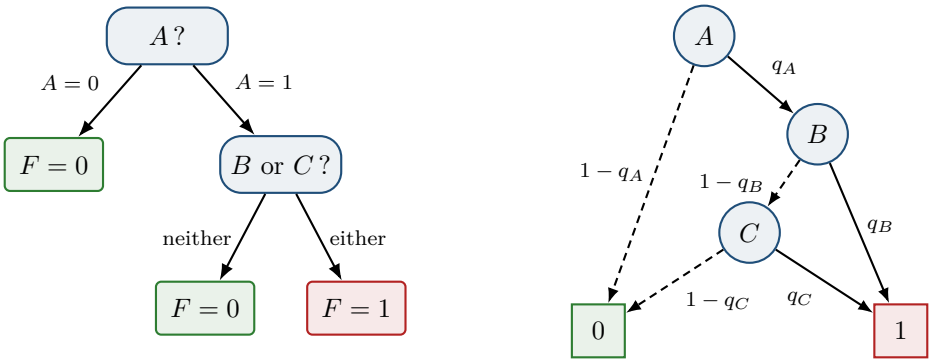
$$\Pr(F) = (1 - q_A) \times 0 + q_A\ \Pr(B + C) = q_A\,(q_B + q_C - q_B q_C).$$

Expanding $B + C$ on B in the same way, $(B + C)_{B=1} = 1$ and $(B + C)_{B=0} = C$, so $\Pr(B + C) = q_B + (1 - q_B)q_C$, and

$$\Pr(F) = q_A q_B + q_A(1 - q_B)q_C.$$

The two products are the probabilities of the two paths of the diagram that end at the failure terminal, AB and $A\overline{B}C$, and they are mutually exclusive: a BDD turns the top event into a sum of disjoint products, whose probabilities may be added. With the values of Section 11.2.3, $\Pr(F) = 0.10 \times 0.20 + 0.10 \times 0.80 \times 0.30 = 0.020 + 0.024 = 0.044$, and adding the variable D at the top of the order gives $Q_{\text{top}} = q_D + (1 - q_D) \times 0.044 = 0.01 + 0.99 \times 0.044 = 0.05356$, the exact value. The recursion keeps the repeated event A as one variable instead of treating the two cut sets AB and AC as independent.

Figure 11.5 shows the decomposition. Panel (a) is a conceptual decision tree: the root asks whether A occurs; the branch $A = 0$ ends in system success, $F = 0$; the branch $A = 1$ leads to a second question, whether B or C occurs, which ends in success if neither occurs and in failure, $F = 1$, if either does. It is not yet a binary decision diagram, because its second node tests a compound event. Panel (b) is the reduced ordered BDD of F for the variable order A, B, C: each circle tests one basic event, solid edges are taken when the event occurs (probability q) and dashed edges when it does not $(1 - q)$, and the two squares are the only terminals. Reading panel (b) from the root, the dashed edge of A goes directly to 0, and the two routes to 1 are AB and $A\overline{B}C$, the disjoint products above.



(a) conceptual decomposition of $F = A(B + C)$

(b) reduced ordered BDD, order A, B, C

Figure 11.5: Decision diagrams of $F = AB + AC = A(B + C)$. (a) Conceptual decomposition on A, whose second node tests the compound event "B or C"; it is not a fully reduced binary decision diagram. (b) Reduced ordered BDD: solid edges, event occurs (probability q); dashed edges, event does not occur $(1 - q)$. The two routes to the terminal 1, AB and $A\overline{B}C$, are disjoint, and $\Pr(F) = q_A q_B + q_A(1 - q_B)q_C$.

A diagram becomes reduced by two rules applied until neither applies: identical subgraphs are merged into one, which is why there is a single 0 and a single 1 terminal, and a node whose two edges lead to the same place is removed, because its variable does not affect the result there. For a fixed variable order the reduced diagram is unique, and the recursion of equation (11.21) visits each node once, so the cost of exact quantification is proportional to the number of nodes. BDDs therefore perform exact quantification efficiently for many fault-tree structures. Their size, however, can depend strongly on the order of the variables, between a size linear in the number of events for a good order and exponential for a bad one for some functions, and some Boolean functions require very large diagrams under every order. Choosing the order is part of the method.

11.5.4 State enumeration

Another method enumerates the system configurations explicitly. Let $\mathbf{X} = (X_1, \ldots, X_n)$ be the vector of basic-event indicators, with $X_i = 1$ when basic event i occurs (the component has

failed). For independent basic events, the probability of a configuration $\mathbf{x}$ is

$$\Pr(\mathbf{X} = \mathbf{x}) = \prod_{i=1}^{n} q_i^{x_i}(1 - q_i)^{1-x_i}, \tag{11.22}$$

since each factor is q_i when $x_i = 1$ and $1 - q_i$ when $x_i = 0$. If $\phi(\mathbf{x})$ is the structure function of the top event, 1 when the system fails in configuration $\mathbf{x}$, then

$$Q_{\text{top}} = \sum_{\mathbf{x} \in \{0,1\}^n} \phi(\mathbf{x}) \prod_{i=1}^{n} q_i^{x_i}(1 - q_i)^{1-x_i}, \tag{11.23}$$

because the configurations are mutually exclusive and exhaust all possibilities. The result is exact under the independence assumption; for dependent events the product must be replaced by the joint probability distribution of the basic events.

Table 11.4 carries out the enumeration for $T = AB + AC + D$ with the original values. Each row is one of the $2^4 = 16$ configurations, ordered as binary numbers $x_A x_B x_C x_D$; the columns are the four indicators, the probability of the configuration and the value of ϕ. The first row, all components functioning, has probability $0.9 \times 0.8 \times 0.7 \times 0.99 = 0.49896$. Eleven configurations have $\phi = 1$; their probabilities sum to 0.05356, the exact value once more, and the five with $\phi = 0$ sum to 0.94644. The seven regions of Figure 11.2 are sums of these rows: the region C_3 only, 0.00956, collects the five rows with $x_D = 1$ in which neither AB nor AC has occurred.

Table 11.4: State enumeration of $T = AB + AC + D$ with $q_A = 0.10$, $q_B = 0.20$, $q_C = 0.30$, $q_D = 0.01$, independent (illustrative). The configurations with $\phi = 1$ sum to $Q_{\text{top}} = 0.05356$.

x_A	x_B	x_C	x_D	Probability	ϕ	x_A	x_B	x_C	x_D	Probability	ϕ
0	0	0	0	0.498960	0	1	0	0	0	0.055440	0
0	0	0	1	0.005040	1	1	0	0	1	0.000560	1
0	0	1	0	0.213840	0	1	0	1	0	0.023760	1
0	0	1	1	0.002160	1	1	0	1	1	0.000240	1
0	1	0	0	0.124740	0	1	1	0	0	0.013860	1
0	1	0	1	0.001260	1	1	1	0	1	0.000140	1
0	1	1	0	0.053460	0	1	1	1	0	0.005940	1
0	1	1	1	0.000540	1	1	1	1	1	0.000060	1

The difficulty is the 2^n state space. State enumeration has an important advantage, however: the function ϕ need not be a Boolean expression. It can be any procedure that decides, for a configuration, whether the required function is met, including a physical simulation. A transmission network outage configuration, for example, can be evaluated with a power-flow model to determine whether the required electrical service remains feasible, and a cooling-plant configuration with a heat balance. That is more general than relying solely on static Boolean relations between components, and it is the bridge between fault-tree logic and physical models of the system.

11.5.5 Monte Carlo simulation

Monte Carlo simulation samples system states according to a probability model instead of enumerating them. For each realization $k = 1, \ldots, N$:

1. sample the states of the basic events, $\mathbf{x}^{(k)}$, from their joint distribution;
2. evaluate the fault-tree Boolean function, or any other model of the required function, $\phi(\mathbf{x}^{(k)})$;
3. record whether the top event occurs;
4. repeat many times.

The estimate of the top-event probability is the fraction of realizations in which it occurs,

$$\hat{Q}_{\text{top}} = \frac{1}{N} \sum_{k=1}^{N} \phi(\mathbf{x}^{(k)}). \tag{11.24}$$

Each $\phi(\mathbf{x}^{(k)})$ is a Bernoulli variable with mean Q_{top} and variance $Q_{\text{top}}(1-Q_{\text{top}})$; for independent realizations the variance of the mean is that variance divided by N, and the standard error of the estimate is estimated by

$$\text{SE}(\hat{Q}_{\text{top}}) \approx \sqrt{\frac{\hat{Q}_{\text{top}}\,(1-\hat{Q}_{\text{top}})}{N}}. \tag{11.25}$$

For $Q_{\text{top}} = 0.05356$ and $N = 10^4$ realizations, the standard error is $\sqrt{0.05356 \times 0.94644/10^4} = 0.00225$, or 4.2 % of the estimate; $N = 10^6$ divides it by ten.

Ordinary Monte Carlo becomes inefficient for very rare events. The relative standard error is $\text{SE}\,/Q_{\text{top}} = \sqrt{(1-Q_{\text{top}})/(NQ_{\text{top}})}$. If the top-event probability is about 10^{-7}, a million realizations produce on average 0.1 occurrences of the top event, and the relative standard error is about 3.2, which is not enough for any estimate; a relative standard error of 10 % needs $N \approx 1/(0.01 \times 10^{-7}) = 10^9$ realizations. Methods such as importance sampling, splitting and conditional Monte Carlo can be far more efficient when the event is rare and the probability model supports them. Monte Carlo keeps the advantage of state enumeration: ϕ can be a physical model, and the basic events can follow any joint distribution, including dependent and time-dependent ones.

11.5.6 Comparison of the methods

Table 11.5 compares the five approaches. For each it gives the major advantage, the main limitation, and whether the result is exact under the stated probability model. Boolean reduction reveals the essential failure combinations, but the cut-set expansion may become very large; inclusion–exclusion gives the exact union probability, at the cost of exponentially many intersection terms; BDDs give efficient exact quantification for many Boolean structures but are sensitive to graph complexity and variable ordering; state enumeration evaluates every configuration directly, of which there are 2^n; Monte Carlo is flexible for complex stochastic and physical models, but rare events make sampling expensive.

Table 11.5: Computational approaches to fault-tree quantification.

Method	Major advantage	Main limitation	Exact?
Boolean reduction	Reveals the essential failure combinations	Cut-set expansion may become very large	logic only
Inclusion–exclusion	Exact union probability	Exponential growth in intersection terms, $2^m - 1$	yes
BDD	Efficient exact quantification for many Boolean structures	Sensitive to graph complexity and variable ordering	yes
State enumeration	Directly evaluates every configuration; admits physical models	2^n possible states	yes
Monte Carlo	Flexible for complex stochastic and physical models	Rare-event sampling can be expensive	estimate

These methods are alternatives and complements, not competing definitions of system reliability. Under the same structure function and the same probability model they must give the same top-event probability, within the sampling error for Monte Carlo; a production analysis often uses Boolean reduction to find and screen the minimal cut sets, a BDD to quantify them exactly, and simulation where the logic is not static.

11.6 Application to the data center cooling plant

The methods of this chapter are now applied to a larger fault tree of the data center cooling plant, the running case of this book. This section constructs the tree, identifies its minimal cut sets, quantifies each subsystem and the top event, reads what the cut sets reveal, and closes with the limits of a static cut-set analysis.

11.6.1 The fault tree of the cooling plant

The cooling plant contains two redundant electrical feeders, three chillers of which two are required, two redundant chilled-water pumps, and one shared cooling controller. The top event is

$$T = \{\text{insufficient cooling capacity at the assessed instant}\}.$$

For simplicity, the plant is taken to fail if any of the following occurs: both feeders are unavailable; at least two chillers are unavailable; both pumps are unavailable; the common controller is unavailable. The basic events are written

$$E_1, E_2: \text{ feeder failures}, \qquad X_1, X_2, X_3: \text{ chiller failures},$$
$$W_1, W_2: \text{ pump failures}, \qquad K: \text{ controller failure},$$

with the letters X and W chosen for the chillers and the pumps because C_j and P_j denote minimal cut and path sets. The top event is

$$T = E_1E_2 + X_1X_2 + X_1X_3 + X_2X_3 + W_1W_2 + K, \tag{11.26}$$

in which all terms are ORed together and the three chiller terms represent the 2-out-of-3 failure condition: at least two of the three chillers unavailable.

Figure 11.6 draws the tree. The top event is the box at the top; below it, a four-input OR gate. Its first input is the intermediate event "both feeders unavailable", the output of an AND gate of E_1 and E_2. Its second input is "two or more chillers unavailable", the output of a voting gate, drawn as an OR gate marked 2/3, whose output occurs when at least two of its three inputs occur. Its third input is the AND gate of the two pump failures, and its fourth input is the basic event K, connected directly.

11.6.2 The minimal cut sets

Expression (11.26) is already a sum of products in which no product contains another, so each product is a minimal cut set; the voting gate expands into the three pairs of chillers. The minimal cut sets are listed in Table 11.6, with their order, the failure combination each represents, and its probability for the basic-event values given in Section 11.6.3. The single controller is a first-order minimal cut set; every other way of losing the plant requires at least two basic events.

Even before any probability is assigned, this structure shows that the shared controller deserves careful attention: it is the only component whose failure alone disables the whole cooling service. The minimal path sets give the success view: the plant functions when one feeder, two chillers, one pump and the controller all function, so there are $2 \times 3 \times 2 \times 1 = 12$ minimal path sets, each of order five, for example $\{\overline{E_1}, \overline{X_1}, \overline{X_2}, \overline{W_1}, \overline{K}\}$. The controller belongs to all twelve, which is the dual statement of its being a first-order cut set.

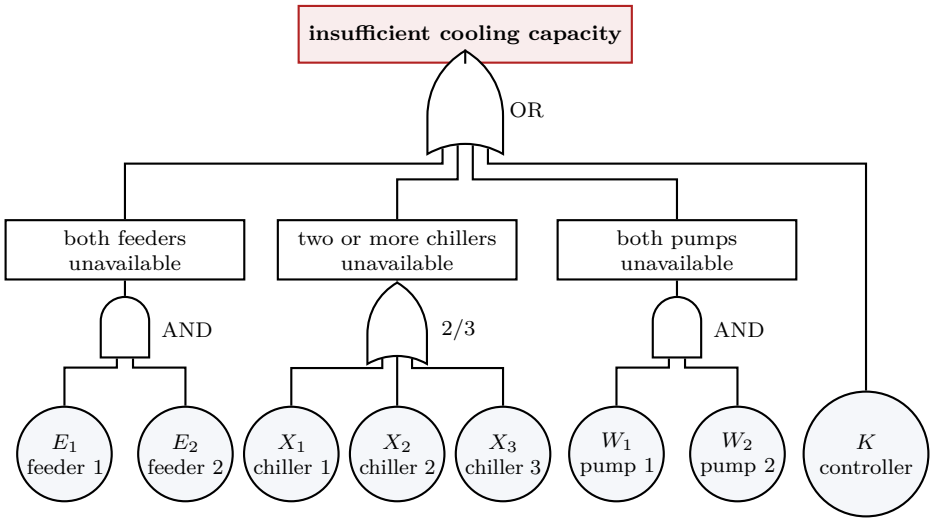


Figure 11.6: Fault tree of the data center cooling plant, equation (11.26). A four-input OR gate combines the AND gate of the two feeders, the 2/3 voting gate of the three chillers (drawn as an OR gate marked 2/3: output when at least two inputs occur), the AND gate of the two pumps, and the controller K, a basic event connected directly.

Table 11.6: Minimal cut sets of the cooling plant, with their probabilities for independent basic events with $q_E = 0.02$ per feeder, $q_X = 0.05$ per chiller, $q_W = 0.03$ per pump and $q_K = 0.01$ (illustrative point unavailabilities).

Cut set	Order	Failure combination	$\Pr(C_j)$
$\{K\}$	1	Common controller unavailable	0.01
$\{E_1, E_2\}$	2	Both electrical feeders unavailable	0.0004
$\{W_1, W_2\}$	2	Both pumps unavailable	0.0009
$\{X_1, X_2\}$	2	Two chillers unavailable	0.0025
$\{X_1, X_3\}$	2	Two chillers unavailable	0.0025
$\{X_2, X_3\}$	2	Two chillers unavailable	0.0025
Sum S_1			0.0188

11.6.3 Quantification of the subsystems

Assume independent basic events with the point unavailabilities $q_E = 0.02$ for each feeder, $q_X = 0.05$ for each chiller, $q_W = 0.03$ for each pump and $q_K = 0.01$ for the shared controller. These are hypothetical probabilities that each component is unavailable at the assessed instant, $1 - A(t)$ for each component in the notation of Chapter 3, not failure rates per hour; the chiller value corresponds to the chiller reliability 0.95 of Chapter 1.

Electrical subsystem. Both feeders must be unavailable,

$$Q_{\text{elec}} = q_E^2 = (0.02)^2 = 0.0004.$$

Pumping subsystem. Both pumps must be unavailable,

$$Q_{\text{pump}} = q_W^2 = (0.03)^2 = 0.0009.$$

Chiller subsystem. At least two of the three chillers must be unavailable. The number of unavailable chillers is binomial with three trials and probability q_X; exactly two are unavailable in $\binom{3}{2} = 3$ ways, each with probability $q_X^2(1 - q_X)$, and all three with probability q_X^3, so

$$Q_{\text{chill}} = \binom{3}{2} q_X^2(1 - q_X) + q_X^3 = 3 \times 0.0025 \times 0.95 + 0.000125 = 0.007125 + 0.000125 = 0.00725. \tag{11.27}$$

The chiller plant reliability is $1 - 0.00725 = 0.99275$, the value $3R^2 - 2R^3$ of Chapter 1 at $R = 0.95$. The sum of the three chiller minimal cut-set probabilities is $3q_X^2 = 3 \times 0.0025 = 0.0075$, slightly more than the exact 0.00725. The difference is the overlap: the state in which all three chillers fail satisfies all three cut sets and is counted three times instead of once, so the sum exceeds the exact value by $2q_X^3 = 0.00025$.

Controller. The controller is a single basic event, $Q_{\text{ctrl}} = q_K = 0.01$.

11.6.4 The top-event probability

The four subsystem failure events involve disjoint sets of basic events, and the basic events are independent, so the four subsystem events are independent. The plant functions only if no subsystem fails, so

$$Q_{\text{top}} = 1 - (1 - Q_{\text{elec}})(1 - Q_{\text{pump}})(1 - Q_{\text{chill}})(1 - Q_{\text{ctrl}}). \tag{11.28}$$

Substituting,

$$\begin{aligned} Q_{\text{top}} &= 1 - 0.9996 \times 0.9991 \times 0.99275 \times 0.99 \\ &= 1 - 0.99870036 \times 0.99275 \times 0.99 \\ &= 1 - 0.99145978 \times 0.99 = 1 - 0.98154518 = 0.0184548. \end{aligned}$$

The exact point unavailability of the cooling plant is 1.8455 %, and the corresponding availability at the assessed instant is $A = 1 - Q_{\text{top}} = 98.1545\,\%$, under the assumptions of independent basic events and the stated subsystem success criteria. A complete enumeration of the $2^8 = 256$ configurations of the eight basic events by equation (11.23) gives the same value.

The approximations and bounds of Section 11.3 bracket it closely. The rare-event sum is $S_1 = 0.0188$, 1.9 % above the exact value; the min-cut upper bound is $1 - (1 - 0.01)(1 -$

$0.0004)(1-0.0009)(1-0.0025)^3 = 0.018683$, 1.2 % above; and the second-order lower bound is $S_1 - S_2 = 0.0188 - 0.000473 = 0.018327$, 0.7 % below. For this plant,

$$0.018327 \le 0.018455 \le 0.018683 \le 0.0188.$$

The approximations are good because every overlap involves either the controller and another cut set, or two chiller pairs, all of small probability.

This number is the probability that the plant is unable to meet its required capacity at an assessed instant. It is not the probability of surviving a 24-hour mission without interruption, which would require models of the timing of failures, of repair, and of the mission success criterion; those are the subjects of later chapters on repairable systems and Markov models.

11.6.5 What the cut sets reveal

Figure 11.7 compares the four subsystems. Each horizontal bar is the unavailability of one subsystem, in units of 10^{-3}, and the number at its end gives the value and, in parentheses, its share of the sum of the four subsystem unavailabilities, 0.01855. The controller has a point unavailability of 1 %, the chiller subsystem 0.725 %; the electrical and pumping subsystems have much smaller unavailabilities, 0.04 % and 0.09 %. The controller and the chillers together account for 93 % of the total.

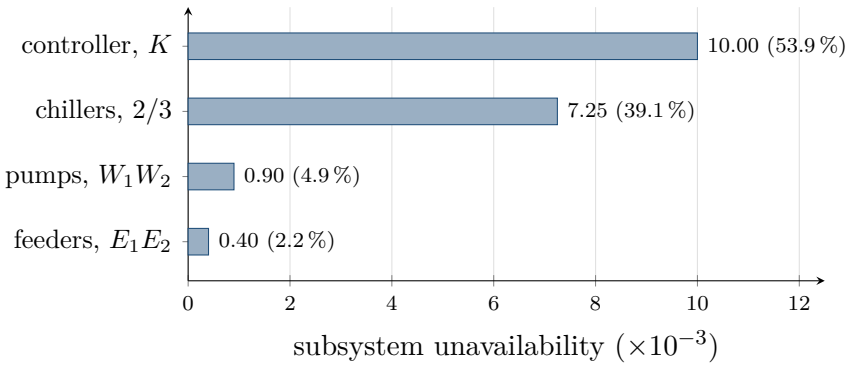


Figure 11.7: Point unavailability of the four subsystems of the cooling plant (illustrative values of Table 11.6): $Q_{\text{elec}} = 0.0004$, $Q_{\text{pump}} = 0.0009$, $Q_{\text{chill}} = 0.00725$ and $Q_{\text{ctrl}} = 0.01$; in parentheses, the share of their sum, 0.01855. The top-event probability is $Q_{\text{top}} = 0.018455$.

This suggests two engineering priorities. The first is to investigate the shared controller, because its failure alone can disable the entire cooling service, and it is the largest single contributor. The second is to investigate the chiller subsystem, because combinations of two chiller outages are a significant contributor to the total unavailability, although each chiller is backed by a spare.

Ranking actual investment decisions requires more information than the cut sets provide. It requires answers to questions such as these: how much a redundant controller would cost; whether two controllers would share software or power dependencies, which would make their failures dependent; how much the chiller reliability could realistically be improved; what the consequences of the different outage scenarios are; and whether simultaneous component failures could arise from a common cause. Minimal cut sets expose vulnerabilities, but they do not by themselves determine the best investment. That requires sensitivity analysis, importance measures and decision optimization.

11.6.6 Limits of static cut-set analysis

Several limitations must be kept in mind when interpreting minimal cut sets and the probabilities computed from them.

Repeated events. If the same basic event, an electrical feeder failure for instance, appears in several branches of a tree, it remains one underlying random event. Assigning separate, independent probabilities to its repeated appearances is generally incorrect. For $T = AB + AC$ with the original values, treating the two appearances of A as independent copies would give $1 - (1 - 0.02)(1 - 0.03) = 0.0494$ instead of the exact 0.044, an overestimate of 12 %. Boolean reduction (by idempotence) and BDDs (by testing each variable once on a path) both keep a repeated event as one event.

Dependent events. Minimal cut sets describe logical sufficiency, not statistical independence. If two chiller failures are correlated, because of a shared cooling-water problem, a common maintenance error or a common power disturbance, then

$$\Pr(X_1 X_2) \neq \Pr(X_1)\ \Pr(X_2)$$

unless independence is justified, and every product formula of this chapter, including equation (11.27), must be replaced by the joint probability. Positive dependence among the events of one cut set raises its probability above the product, often by orders of magnitude for highly reliable redundant components; common-cause failures are the subject of Chapter 13.

Truncation. Large analyses often retain only the cut sets up to a specified order or above a probability threshold. This improves computational efficiency, but the omitted cut sets may contribute to the top-event probability. If cut sets are deliberately omitted, the resulting sum is no longer guaranteed to be an upper bound on the complete top-event probability. Truncating the cooling plant at order one keeps only $\{K\}$ and gives 0.01, well below the exact 0.018455. A truncated analysis should therefore report an estimate of the probability it has discarded.

Dynamic behavior. Static cut sets describe combinations of component states. They do not automatically capture the order in which events occur, activation delays, repair processes or control actions. A cold-standby pump, which starts only when the duty pump fails and can fail on demand, depends on the sequence of events, not only on the combination of states, and requires dynamic reliability modeling; later chapters on repairable systems and Markov models develop it.

The failure logic says which combinations matter and the probability model how likely they are; the next chapter, Chapter 12, uses both to rank the components and the cut sets by their contribution to the top event, through the Birnbaum, Fussell–Vesely, risk achievement worth and risk reduction worth importance measures.

11.7 What to retain

Definition 11.1 (Minimal cut sets and fault-tree quantification)**.** A coherent fault tree is the union of its minimal cut sets, the smallest combinations of basic events sufficient for the top event; its probability is the probability of that union, computed exactly by inclusion–exclusion, factoring, a binary decision diagram or enumeration, and bounded above by the min-cut upper bound and the sum of the cut-set probabilities, which is a good approximation only when the cut sets rarely occur together.

Table 11.7 lists the chapter's terms with their meaning, their expression and units, and the subsection that defines each. All probabilities in it are dimensionless and refer to a stated mission or a stated instant.

Table 11.7: Terms of minimal cut-set analysis and fault-tree quantification.

Term	Meaning	Expression, units	Defined in
Basic-event probability	probability that basic event i occurs	q_i, –	11.1
Cut set	events whose joint occurrence guarantees the top event	$\phi(\mathbf{x}) = 1$ if $x_i = 1$ for all $i \in C$	11.1.1
Minimal cut set	cut set with no removable event	C_j	11.1.2
Order	number of basic events in a minimal cut set	$\lvert C_j \rvert$; 1 = single point of failure	11.1.2
Absorption, idempotence	remove nonminimal and repeated terms	$X + XY = X$; $XX = X$	11.1.3
Coherent fault tree	structure function nondecreasing, all events relevant	$T = \bigcup_j C_j$	11.1.3
Inclusion–exclusion	exact probability of a union	$Q_{\text{top}} = \sum_k (-1)^{k-1} S_k$	11.2.2
Rare-event approximation	sum of cut-set probabilities	$Q_{\text{top}} \approx S_1 = \sum_j \Pr(C_j)$	11.3.1
Union bound	the sum is an upper bound	$Q_{\text{top}} \le S_1$	11.3.1
Second-order lower bound	first two inclusion–exclusion terms	$Q_{\text{top}} \ge S_1 - S_2$	11.3.3
Min-cut upper bound	bound for independent basic events	$1 - \prod_j (1 - \Pr(C_j))$	11.3.3
Minimal path set	smallest set of successes guaranteeing success	P_j	11.4.1
Min-path lower bound	bound from the path sets	$\prod_j [1 - \prod_{i \in P_j} (1 - q_i)]$	11.4.2
Boolean expansion	bottom-up generation of minimal cut sets	AND: pairwise unions; OR: union; minimize	11.5.2
Shannon expansion	decomposition on one variable	$F = \overline{X} F_{X=0} + X F_{X=1}$	11.5.3
Binary decision diagram	graph of decisions; disjoint paths to failure	$\Pr(F) = (1 - q_X) \Pr(F_{X=0}) + q_X \Pr(F_{X=1})$	11.5.3
State enumeration	sum over all configurations	eq. (11.23), 2^n terms	11.5.4
Monte Carlo estimate	fraction of sampled states that fail	$\hat{Q}_{\text{top}}$, SE $\approx \sqrt{\hat{Q}(1 - \hat{Q})/N}$	11.5.5
Voting gate	output when at least k of n inputs occur	2-out-of-3: $3q^2(1 - q) + q^3$	11.6.1
Point unavailability	probability of being unavailable at an instant	$1 - A(t)$	11.6.3
Truncation	omission of cut sets by order or probability	sum no longer an upper bound	11.6.6

Five ideas carry forward. Minimal cut sets identify the essential vulnerabilities, the combinations of failures sufficient to cause a system-level failure. Minimality is logical, not statistical: a minimal cut set is not necessarily probable or important. Overlapping cut sets must be handled correctly, because their probabilities cannot in general be summed to obtain the exact top-event probability. Computational structure matters: Boolean reduction and binary decision diagrams avoid much of the expense of brute-force enumeration. And correct logic is not enough: dependence among components, common causes, time-dependent behavior and physical constraints must also be represented. In short,

system failure logic + probability model = quantified system reliability;

the logic says which combinations matter, the probability model how likely they are, and neither is sufficient by itself.

Further reading. Rausand, Barros and Høyland, *System Reliability Theory*, third edition (minimal cut sets, path sets, Boolean reduction and system reliability). Kumamoto and Henley, *Probabilistic Risk Assessment and Management for Engineers and Scientists*, second edition (fault-tree quantification and risk evaluation). U.S. Nuclear Regulatory Commission, *Fault Tree Handbook* (NUREG-0492), especially the chapters on qualitative and quantitative fault-tree evaluation. NASA, *Fault Tree Handbook with Aerospace Applications* (minimal cut-set generation, rare-event approximations and BDD quantification).

11.8 Exercises

Exercise 11.1. (a) For $T = AB + C$, list the minimal cut sets and their orders. (b) Reduce $T = AB + ABC$ to its minimal expression. (c) For $T = AB + E$, list the minimal path sets. (d) Explain why the sum of the minimal cut-set probabilities can exceed the true top-event probability. (e) Name an approach that represents repeated Boolean variables consistently and supports exact probability calculation. (Answers: (a) $\{A, B\}$ of order 2 and $\{C\}$ of order 1, since A and B jointly cause failure while C causes it alone; (b) AB, because ABC contains AB and is absorbed by the smaller sufficient combination; (c) $\{\overline{A}, \overline{E}\}$ and $\{\overline{B}, \overline{E}\}$, since the supply must work together with at least one pump; (d) the cut sets may overlap, so one system configuration can satisfy several of them and is counted more than once; (e) binary decision diagrams, which represent the Boolean structure recursively and quantify it exactly under a specified probability model.)

Exercise 11.2. For $T = AB + C$ with independent basic events, $q_A = q_B = 0.10$ and $q_C = 0.05$ (illustrative): (a) compute the exact top-event probability; (b) compute the rare-event approximation and its relative error; (c) compute the min-cut upper bound and explain why it equals the exact value here; (d) compute R_{system} from the minimal path sets and check it against (a). (Answers: (a) $1 - (1 - 0.01)(1 - 0.05) = 0.0595$; (b) 0.06, 0.84 % high; (c) 0.0595, because the two cut sets share no basic event and are independent; (d) $(1 - q_C)(1 - q_A q_B) = 0.9405 = 1 - 0.0595$.)

Exercise 11.3. A fault tree has the top event $T = (A + B)(A + C) + BD$, with independent basic events, $q_A = 0.01$, $q_B = 0.10$, $q_C = 0.20$ and $q_D = 0.05$ (illustrative). (a) Expand and reduce the expression by absorption and list the minimal cut sets with their orders. (b) Compute the exact top-event probability by Shannon expansion, first on A and then on B, and draw the reduced ordered BDD for the order A, B, C, D. (c) Check the result by inclusion–exclusion, giving S_1, S_2 and S_3. (d) Compute the rare-event approximation, the min-cut upper bound and the second-order lower bound, and order them with the exact value. (e) Derive the minimal path sets by De Morgan's laws and evaluate the min-path lower bound. (Answers: (a) $T = A + BC + BD$, minimal cut sets $\{A\}$, $\{B, C\}$, $\{B, D\}$; (b) $Q_{\text{top}} = 0.01 + 0.99 \times 0.10 \times (0.20 + 0.05 - 0.01) = 0.03376$; (c) $S_1 = 0.035$, $S_2 = 0.00125$, $S_3 = 0.00001$; (d) $0.03375 \le 0.03376 \le 0.034651 \le 0.035$; (e) $\{\overline{A}, \overline{B}\}$ and $\{\overline{A}, \overline{C}, \overline{D}\}$, bound 0.02699.)

Exercise 11.4. Return to the cooling plant of Section 11.6, with the values of Table 11.6. (a) The single controller is replaced by two redundant controllers, either of which suffices, each unavailable with probability 0.01, independently. List the new minimal cut sets and compute Q_{top}. (b) Instead, the chillers are improved to $q_X = 0.025$ and the single controller is kept. Compute Q_{chill} and Q_{top}, and compare the two improvements. (c) In (a), the two controllers run the same software, and a software fault disables both with probability 0.005, independently of

their other failures. Compute the controller-subsystem unavailability and Q_{top}, and state what the new first-order cut set is. (d) How many Monte Carlo realizations are needed to estimate the original $Q_{\text{top}} = 0.018455$ with a relative standard error of $10\,\%$? (Answers: (a) $\{K_1, K_2\}$ replaces $\{K\}$, $Q_{\text{top}} = 0.008639$; (b) $Q_{\text{chill}} = 0.001844$, $Q_{\text{top}} = 0.013110$, so the redundant controller removes more unavailability; (c) $0.005 + 0.0001 - 0.0000005 = 0.0050995$, $Q_{\text{top}} = 0.013596$, with the software fault as a first-order cut set; (d) $N = (1 - Q_{\text{top}})/(0.01\,Q_{\text{top}}) \approx 5300$.)

Chapter 12

Reliability Importance Measures

12.1 Component importance in a quantified system model

Chapters 9, 10 and 11 built the tools that turn a system's logic and its component probabilities into one number: the probability that the system fails. For the data center cooling plant of this book that number is the probability that the cooling function is unavailable at a randomly selected instant, about 1.85 %. The number states how often the required function is lost; it does not say what to do about it. An engineer who can improve the electrical feeders, replace the pumps, add a chiller or install a redundant controller needs to know which of these changes the system failure probability most, and which components are essential to keeping it at its present value. The question of this chapter is

how the contribution of each basic event to the probability of system failure is measured, which engineering question each measure answers, and why a ranking of components is not by itself a ranking of improvements.

Probabilistic risk assessment answers it with *importance measures*. Table 12.1 lists the four that the chapter develops, with the question each one answers. Practice distinguishes two uses among them: measures that identify opportunities to reduce the present risk, and measures that show which features are essential to maintaining the present risk level. The first use is served by the Fussell–Vesely importance and the risk reduction worth, the second by the risk achievement worth, and the Birnbaum importance underlies both.

Table 12.1: The four importance measures of this chapter and the question each answers about basic event i.

Measure	Symbol	Question
Birnbaum importance	I_i^{B}	How sensitive is the system failure probability to the probability of this basic event?
Fussell–Vesely importance	I_i^{FV}	What fraction of the present system failure probability would be eliminated if this basic event were made impossible?
Risk achievement worth	RAW_i	By what factor would the system failure probability grow if this basic event were certain?
Risk reduction worth	RRW_i	By what factor would the system failure probability fall if this basic event were made impossible?

Jargon: basic event, top event, coherent system, importance measure, critical component, Birnbaum, Fussell–Vesely, RAW, RRW

A *basic event* is an event at the bottom of a fault tree, usually the failure or unavailability of one component, to which a probability q_i is assigned. The *top event* is the undesired system event the tree describes; its probability is written Q_{sys} in this chapter. A system (or fault tree) is *coherent* when the occurrence of an additional basic event can never prevent the top event: failing a component never repairs the system. An *importance measure* is a number that ranks basic events by their contribution to the probability of the top event. A component is *critical* (pivotal) in a state of the other components when its own state decides whether the top event occurs. The *Birnbaum importance* I_i^{B} is the partial derivative of Q_{sys} with respect to q_i; the *Fussell–Vesely importance* I_i^{FV} is the fraction of Q_{sys} removed when event i is made impossible; the *risk achievement worth* RAW_i (RAW) is the ratio of the top-event probability with event i certain to Q_{sys}; the *risk reduction worth* RRW_i (RRW) is the ratio of Q_{sys} to the top-event probability with event i impossible. All four are dimensionless.

Key idea

Every importance measure compares the system failure probability with one basic-event probability changed, to the system failure probability of the present model. The four measures differ only in which change they make and in how they report the comparison, and none of them is, by itself, the benefit of an engineering project.

This section restates the running cooling-plant model of Chapter 11 and introduces the three probabilities from which all four measures are computed.

12.1.1 The running example: the data center cooling plant

The cooling plant of this book has four groups of equipment: two electrical feeders, of which one suffices; three chillers, of which two are required; two chilled-water pumps, of which one suffices; and one shared controller, whose failure alone disables the cooling function. The top event is T, "the required cooling function is unavailable at the assessed instant". Writing E_1, E_2 for the unavailability of feeders 1 and 2, C_1, C_2, C_3 for the unavailability of the three chillers, P_1, P_2 for the two pumps and K for the controller, the minimal cut sets derived in Chapter 11 give the top event as the Boolean sum

$$T = E_1E_2 + P_1P_2 + C_1C_2 + C_1C_3 + C_2C_3 + K, \tag{12.1}$$

in which juxtaposition is AND and $+$ is OR. The six minimal cut sets are $\{K\}$, of order one, and $\{E_1, E_2\}$, $\{P_1, P_2\}$, $\{C_1, C_2\}$, $\{C_1, C_3\}$ and $\{C_2, C_3\}$, of order two.

Figure 12.1 draws the fault tree. The top event is the red box at the top. Below it an OR gate, the shield-shaped symbol whose output occurs when any input occurs, collects four branches. The first and third branches are AND gates, the flat-based symbol whose output occurs only when all inputs occur: both feeders and both pumps must be unavailable. The second branch is a voting gate marked 2/3, which occurs when at least two of its three inputs occur. The fourth branch is the controller, a basic event that feeds the OR gate directly. Basic events are the circles at the bottom.

The basic-event probabilities are *snapshot unavailabilities*: the probability that a component is unavailable at a randomly selected instant, or equivalently its long-run fraction of time unavailable. They are not failure rates per hour. Table 12.2 lists them; they are the illustrative values of Chapter 11, kept unchanged. The basic events are assumed mutually independent throughout the numerical work of this chapter; Section 12.6.4 returns to that assumption.

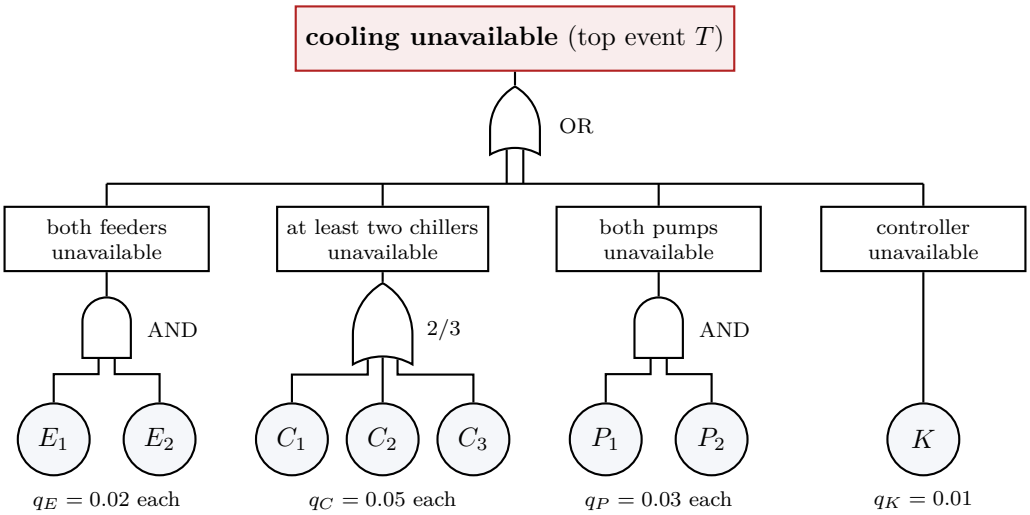


Figure 12.1: Fault tree of the data center cooling plant, equation (12.1). Any of the four branches causes the top event: both feeders, at least two of three chillers (voting gate 2/3), both pumps, or the shared controller. The basic events (circles) are assumed independent, with the snapshot unavailabilities of Table 12.2 (illustrative).

Table 12.2: Basic events of the cooling plant and their snapshot unavailabilities (illustrative, as in Chapter 11).

Basic event	Symbols	Number of events	Probability
Electrical feeder unavailable	E_1, E_2; q_E	2	0.02 each
Pump unavailable	P_1, P_2; q_P	2	0.03 each
Chiller unavailable	C_1, C_2, C_3; q_C	3	0.05 each
Controller unavailable	K; q_K	1	0.01

With independent basic events, each branch of Figure 12.1 has its own unavailability. The feeder and pump branches are AND gates of two events,

$$Q_E = q_{E_1} q_{E_2} = q_E^2 = 0.0004, \qquad Q_P = q_{P_1} q_{P_2} = q_P^2 = 0.0009. \tag{12.2}$$

The chiller branch occurs when exactly two chillers are unavailable, in $\binom{3}{2} = 3$ ways each of probability $q_C^2(1 - q_C)$, or when all three are,

$$Q_C = 3q_C^2(1 - q_C) + q_C^3 = 3q_C^2 - 2q_C^3 = 3(0.0025) - 2(0.000125) = 0.00725. \tag{12.3}$$

The four branches involve disjoint sets of independent basic events, so they are independent, and the top event fails to occur only when none of them occurs:

$$Q_{\text{sys}} = 1 - (1 - Q_E)(1 - Q_P)(1 - Q_C)(1 - q_K). \tag{12.4}$$

Substituting,

$$Q_{\text{sys}} = 1 - (0.9996)(0.9991)(0.99275)(0.99) = 1 - 0.9815452 = 0.0184548,$$

a snapshot unavailability of the cooling function of 1.8455 %. This is the exact top-event probability under the stated model, including the overlap among cut sets; the sum of the six cut-set probabilities, $0.0004 + 0.0009 + 3(0.0025) + 0.01 = 0.0188$, is its rare-event upper bound. Every importance value in this chapter is computed relative to this baseline.

12.1.2 System failure probability as a function of the basic-event probabilities

The four measures are all built from one function. Collect the basic-event probabilities in the vector $\mathbf{q} = (q_1, q_2, \ldots, q_n)$, and write

$$Q(\mathbf{q}) = \Pr(T) \tag{12.5}$$

for the top-event probability that the model assigns to them. For the cooling plant, $n = 8$ and $Q(\mathbf{q})$ is equation (12.4) with each subsystem term written in the individual q_i. The probability Q is whatever probability of system failure the model computes: a snapshot unavailability, as here, or an unreliability $1 - R(H)$ over a mission, when the basic-event probabilities are failure probabilities over that mission. The measures of this chapter apply unchanged to either.

For each basic event i, three evaluations of Q are used:

$$Q_{\text{sys}} = Q(\mathbf{q}), \qquad \text{the baseline system failure probability;} \tag{12.6}$$

$$Q(0_i) = Q(\mathbf{q} \mid q_i = 0), \qquad \text{the same with event } i \text{ made impossible;} \tag{12.7}$$

$$Q(1_i) = Q(\mathbf{q} \mid q_i = 1), \qquad \text{the same with event } i \text{ certain.} \tag{12.8}$$

The notation $Q(0_i)$ and $Q(1_i)$ is short for Q evaluated at the vector $\mathbf{q}$ with its ith entry replaced by 0 or 1 and all other entries kept. For the controller, for example, $Q(0_K) = 1 - (1 - Q_E)(1 - Q_P)(1 - Q_C)$ and $Q(1_K) = 1$.

These are *model interventions*: one basic-event probability is set to an extreme value while the rest of the model, its logic and every other probability, is held fixed. They are counterfactual calculations on the model, not predictions of what happens when a component is physically removed or replaced. A real redesign can change other probabilities, add failure modes or introduce dependencies, and then the whole model must be updated (Section 12.6). Under this reading, the three quantities of equations (12.6)–(12.8) are sufficient to define all four importance measures.

Remark 12.1. In this chapter C_1, C_2 and C_3 denote the chiller basic events of the running example. Where a general formula needs the minimal cut sets of an arbitrary fault tree, they are written C_j, $j = 1, \ldots, m$, as in Chapter 11; the context always makes clear which is meant, and the cut sets of the cooling plant are always written out as sets of events.

12.2 Birnbaum importance

The Birnbaum importance measures how sensitive the system failure probability is to the probability of one basic event. This section derives it as a partial derivative and as a difference of two conditional probabilities, works it out for one feeder and the controller, gives its physical meaning, and computes it for every component of the cooling plant.

12.2.1 Definition and derivation

For independent basic events in a static Boolean model, the *Birnbaum importance* of basic event i is

$$I_i^{\mathrm{B}} = \frac{\partial Q}{\partial q_i}, \tag{12.9}$$

the rate of change of the system failure probability with the probability of that event, all other probabilities fixed.

The derivative has a closed form. Let X_i be the indicator of basic event i, $X_i = 1$ when it occurs and $X_i = 0$ when it does not, so that $\Pr(X_i = 1) = q_i$. By the law of total probability, conditioning on X_i,

$$Q = \Pr(T \mid X_i = 0)\,(1 - q_i) + \Pr(T \mid X_i = 1)\,q_i. \tag{12.10}$$

Because the basic events are independent, conditioning on X_i does not change the probabilities of the other events. The conditional probability $\Pr(T \mid X_i = 1)$ is therefore the top-event probability computed with $q_i = 1$ and every other q_j unchanged, which is $Q(1_i)$; likewise $\Pr(T \mid X_i = 0) = Q(0_i)$. Equation (12.10) becomes

$$Q = (1 - q_i)\,Q(0_i) + q_i\,Q(1_i) = Q(0_i) + q_i\big[Q(1_i) - Q(0_i)\big], \tag{12.11}$$

which is the *pivotal decomposition* of Q on event i. Neither $Q(0_i)$ nor $Q(1_i)$ depends on q_i, so Q is a linear function of q_i, and differentiating equation (12.11) gives

$$I_i^{\mathrm{B}} = Q(1_i) - Q(0_i). \tag{12.12}$$

The Birnbaum importance is the change in the system failure probability between forcing event i to be absent and forcing it to occur. This second form is the more useful one: it needs two evaluations of the model and no differentiation, and it holds for a component of any system whose basic events are independent.

Two properties follow. First, in a coherent system the occurrence of event i can only add to the states in which the top event occurs, so $Q(1_i) \geq Q(0_i)$; both are probabilities, so

$$0 \leq I_i^{\mathrm{B}} \leq 1. \tag{12.13}$$

Second, I_i^{B} is dimensionless, since Q and q_i are both probabilities. Because Q is linear in q_i, the derivative is exact for any change of q_i alone: changing q_i by Δq_i changes Q by exactly $I_i^{\mathrm{B}}\,\Delta q_i$. It is only when several probabilities change together that the first-order estimate $\sum_i I_i^{\mathrm{B}}\,\Delta q_i$ becomes an approximation (Section 12.5.1). Finally, I_i^{B} does not contain q_i itself: it measures how much the system depends on event i, not how often event i occurs.

12.2.2 One feeder and the controller at subsystem level

Consider the electrical subsystem alone. It fails when both feeders fail, $Q_E = q_{E_1} q_{E_2}$, and its Birnbaum importance with respect to feeder 1 is

$$I_{E_1}^{\mathrm{B}} = \frac{\partial Q_E}{\partial q_{E_1}} = q_{E_2} = 0.02.$$

The same value follows from equation (12.12): with feeder 1 certainly failed the subsystem fails if feeder 2 fails, $Q_E(1_{E_1}) = q_{E_2} = 0.02$, and with feeder 1 never failing it cannot fail, $Q_E(0_{E_1}) = 0$. An increase of feeder 1's unavailability by one percentage point therefore increases the electrical subsystem's unavailability by 0.02 percentage points, exactly, since Q_E is linear in q_{E_1}. The second feeder shields the subsystem from almost all of feeder 1's failures: only the 2 % of the time when feeder 2 is also down does a failure of feeder 1 matter.

The controller is the opposite case. A subsystem consisting of the one indispensable controller fails exactly when the controller does, $Q_K = q_K$, and

$$I_K^{\mathrm{B}} = \frac{\partial Q_K}{\partial q_K} = 1,$$

the largest value allowed by inequality (12.13). Every change in the controller's unavailability passes unattenuated to the subsystem, because a controller failure directly causes loss of the required function. The two examples differ by a factor of fifty in Birnbaum importance although the controller's own unavailability, 0.01, is half the feeder's: the difference comes from the architecture, not from the failure probabilities of the components themselves.

12.2.3 Critical states and the physical meaning

The difference of equation (12.12) has a direct probabilistic reading. Call basic event i *critical* in a given state of the other basic events when the top event occurs if event i occurs and does not occur if it does not: changing component i's state alone changes whether the system functions. In a coherent system, every state of the other events in which T occurs with $X_i = 0$ is also one in which T occurs with $X_i = 1$. The difference $Q(1_i) - Q(0_i)$ is therefore the probability of the states of the other events in which T occurs with $X_i = 1$ but not with $X_i = 0$, which is

$$I_i^{\mathrm{B}} = \Pr(\text{basic event } i \text{ is critical}). \tag{12.14}$$

The Birnbaum importance is high when the component is frequently pivotal to system success, and low when the rest of the system usually either fails without it or survives without it.

For chiller 1 in the full plant, the event is critical when exactly one of chillers 2 and 3 is unavailable, so that the voting gate turns on chiller 1, and when the feeder, pump and controller branches all function, so that no other branch has already caused the top event. With independent events,

$$I_{C_1}^{\mathrm{B}} = \underbrace{2\,q_C(1-q_C)}_{\text{one other chiller down}} \quad (1-Q_E)(1-Q_P)(1-q_K) = 0.095 \times 0.988713 = 0.093928.$$

For feeder 1 in the full plant, the event is critical when feeder 2 is down and the other three branches function, $I_{E_1}^{\mathrm{B}} = q_E(1-Q_P)(1-Q_C)(1-q_K) = 0.02 \times 0.981938 = 0.019639$. The criticality reading explains why importance depends on the architecture: it is the probability of the configurations of the other components in which this component's state is decisive, and that probability is set by the redundancy around the component and by the reliability of everything else.

12.2.4 Birnbaum importance of every component of the plant

For the controller, equation (12.4) has the form

$$Q_{\mathrm{sys}} = 1 - (1 - q_K)\, A_{-K}, \qquad A_{-K} = (1-Q_E)(1-Q_P)(1-Q_C),$$

where A_{-K} is the probability that the feeder, chiller and pump branches all function, the snapshot availability of the plant without its controller. Differentiating,

$$I_K^{\mathrm{B}} = \frac{\partial Q_{\mathrm{sys}}}{\partial q_K} = A_{-K} = (0.9996)(0.9991)(0.99275) = 0.991460. \tag{12.15}$$

The controller is critical whenever the rest of the plant functions, which is 99.15 % of the time.

For pump 1, $Q_P = q_{P_1} q_{P_2}$ enters equation (12.4) through the factor $1 - q_{P_1} q_{P_2}$, whose derivative with respect to q_{P_1} is $-q_{P_2}$, so

$$I^{\mathrm{B}}_{P_1} = q_{P_2}\,(1 - Q_E)(1 - Q_C)(1 - q_K) = 0.03\,(0.9996)(0.99275)(0.99) = 0.029473. \qquad (12.16)$$

The pump's Birnbaum importance is thirty-four times smaller than the controller's because of its redundant partner: pump 1 is critical only when pump 2 is down. This does not mean that pump reliability is unimportant; it means that a change in one pump's unavailability has comparatively little effect on the present system failure probability. By the same argument the feeder has $I^{\mathrm{B}}_{E_1} = 0.019639$, slightly below the subsystem value 0.02 of Section 12.2.2 because the feeder is critical only when the other three branches also function, and the chiller has $I^{\mathrm{B}}_{C_1} = 0.093928$ from Section 12.2.3. The ranking by Birnbaum importance is controller, chiller, pump, feeder.

Figure 12.2 draws the pivotal decomposition of equation (12.11) for chiller 1. The horizontal axis is the probability q_{C_1} of the chiller 1 basic event, from 0 to 1, with the other seven probabilities held at their values in Table 12.2; the vertical axis is the system failure probability, in percent. The solid line is Q as a function of q_{C_1}. It is a straight line, as the derivation requires, from $Q(0_{C_1}) = 1.3758\,\%$ at the left end to $Q(1_{C_1}) = 10.7686\,\%$ at the right end, and its slope is the Birnbaum importance 0.093928. The dot marks the present model, $q_{C_1} = 0.05$ and $Q_{\mathrm{sys}} = 1.8455\,\%$, and the dashed horizontal line is that baseline. The vertical gap between the baseline and the left end is what would be removed by making the event impossible; the gap between the right end and the baseline is what would be added by making it certain. These two gaps are the subject of the next section.

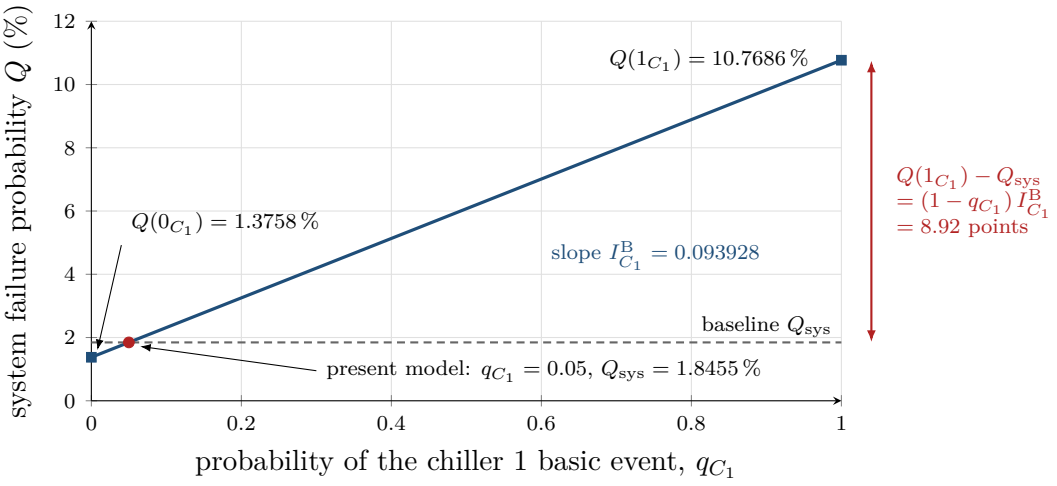


Figure 12.2: Pivotal decomposition of the cooling-plant failure probability on the chiller 1 basic event, $Q = Q(0_{C_1}) + q_{C_1}\,[Q(1_{C_1}) - Q(0_{C_1})]$ (equation (12.11)), with $Q(0_{C_1}) = 0.0137584$, $Q(1_{C_1}) = 0.1076862$ and the other basic-event probabilities of Table 12.2; illustrative values. The line is straight because Q is linear in each single q_i; its slope is the Birnbaum importance. The gap between the baseline and the left end, $Q_{\mathrm{sys}} - Q(0_{C_1}) = q_{C_1}\, I^{\mathrm{B}}_{C_1} = 0.47$ percentage points, is the reduction available by making the event impossible.

12.3 Fussell–Vesely importance, risk achievement worth and risk reduction worth

The Birnbaum importance measures sensitivity; the other three measures report what happens to the system failure probability at the two extremes of one basic event, relative to the baseline.

This section defines the Fussell–Vesely importance and relates it to the Birnbaum importance and to the minimal cut sets, defines the risk achievement worth and interprets it for maintenance, and defines the risk reduction worth and relates it to the Fussell–Vesely importance.

12.3.1 Fussell–Vesely importance

The Fussell–Vesely importance asks how much of the existing system failure probability could be eliminated by making a particular basic event impossible. The absolute reduction achievable by perfect prevention of event i is

$$\Delta Q_i^- = Q_{\text{sys}} - Q(0_i), \tag{12.17}$$

and the *Fussell–Vesely importance* normalizes it by the baseline,

$$I_i^{\text{FV}} = \frac{Q_{\text{sys}} - Q(0_i)}{Q_{\text{sys}}} = 1 - \frac{Q(0_i)}{Q_{\text{sys}}}, \qquad Q_{\text{sys}} > 0. \tag{12.18}$$

It is the fraction of the baseline top-event probability that would be eliminated if the basic event were made impossible, a number between 0 and 1, often quoted as a percentage.

For the controller, the baseline is $Q_{\text{sys}} = 0.0184548$. Setting $q_K = 0$ in equation (12.4) removes the last factor,

$$Q(0_K) = 1 - (0.9996)(0.9991)(0.99275) = 1 - 0.991460 = 0.0085402,$$

so the maximum reduction is $\Delta Q_K^- = 0.0184548 - 0.0085402 = 0.0099146$, and

$$I_K^{\text{FV}} = \frac{0.0099146}{0.0184548} = 0.5372.$$

Making controller failure impossible would remove 53.72 % of the baseline snapshot unavailability of the cooling function. That is a substantial opportunity. The value is a counterfactual result under the stated model: it does not mean that the controller caused 53.72 % of the plant's historical outages, and it does not mean that a redesigned controller would deliver that reduction, since no real controller has zero unavailability and a redesign may change other parts of the model.

12.3.2 Relation to the Birnbaum importance and to the cut sets

The pivotal decomposition (12.11) evaluated at the baseline reads $Q_{\text{sys}} = Q(0_i) + q_i\, I_i^{\text{B}}$, so

$$Q_{\text{sys}} - Q(0_i) = q_i\, I_i^{\text{B}}. \tag{12.19}$$

Substituting into the definition (12.18),

$$I_i^{\text{FV}} = \frac{q_i\, I_i^{\text{B}}}{Q_{\text{sys}}}. \tag{12.20}$$

For the controller, $0.01 \times 0.991460 / 0.0184548 = 0.5372$, as before. The Fussell–Vesely importance therefore depends on three things: the probability q_i of the event itself, the sensitivity I_i^{B} of the system to it, and the baseline Q_{sys}. It is the Birnbaum importance weighted by how often the event occurs.

A component can thus have a high Birnbaum importance and a low Fussell–Vesely importance, when its own failure probability is very small. As an illustrative case, suppose a ninth component, indispensable like the controller, were added in series to the cooling plant with $q_i = 10^{-7}$. Its Birnbaum importance would be the probability that the rest of the plant functions, $1 -$

$0.0184548 = 0.98155$, close to one, while its Fussell–Vesely importance would be $10^{-7} \times 0.98155/0.0184549 = 5.3 \times 10^{-6}$. The system is fully vulnerable to that component, and the component contributes almost nothing to the present failure probability. The distinction between vulnerability and present contribution is the central distinction of importance analysis.

Principle 12.1. *The Birnbaum importance measures how strongly the system depends on a basic event; the Fussell–Vesely importance measures how much that event contributes to the present failure probability. The second is the first weighted by the event's own probability,* $I_i^{\mathrm{FV}} = q_i I_i^{\mathrm{B}}/Q_{\mathrm{sys}}$.

In fault-tree practice the Fussell–Vesely importance is also computed from the minimal cut sets, as the share of the top-event probability carried by the cut sets that contain event i. Three versions are in use, and they agree only when cut-set overlaps are negligible. Writing U_i for the event that at least one minimal cut set containing i occurs, and U_i' for the event that at least one of the other minimal cut sets occurs, the top event is $T = U_i \cup U_i'$, and with independent basic events $Q(0_i) = \Pr(U_i')$. Then

$$Q_{\mathrm{sys}} - Q(0_i) = \Pr(U_i \cup U_i') - \Pr(U_i') = \Pr(U_i \cap \overline{U_i'}),$$

the probability that a cut set containing i occurs and no other cut set does. The counterfactual definition (12.18) is therefore $\Pr(U_i \cap \overline{U_i'})/Q_{\mathrm{sys}}$. The cut-set definition used in many textbooks is $\Pr(U_i)/Q_{\mathrm{sys}}$, the probability that at least one cut set containing i occurs given that the top event occurs, which is never smaller. The rare-event version replaces both probabilities by sums of cut-set probabilities, $\sum_{j:i\in C_j} \Pr(C_j)/\sum_j \Pr(C_j)$. Table 12.3 compares the three for the cooling plant. They agree to three decimals for the feeder and the pump, whose cut sets overlap little with the others, and differ in the second decimal for the chiller, whose two cut sets $\{C_1, C_2\}$ and $\{C_1, C_3\}$ overlap with each other and with $\{C_2, C_3\}$. Naively summing the probabilities of overlapping cut sets is not, in general, an exact Fussell–Vesely calculation; this chapter uses the counterfactual definition (12.18) throughout.

Table 12.3: Three versions of the Fussell–Vesely importance for the cooling plant: the counterfactual definition (12.18), the cut-set definition $\Pr(U_i)/Q_{\mathrm{sys}}$, and the rare-event cut-set sum. Independent basic events, Table 12.2.

Basic event	Cut sets containing it	Counterfactual	$\Pr(U_i)/Q_{\mathrm{sys}}$	Rare-event sum
Feeder E_1	$\{E_1, E_2\}$	0.02128	0.02167	0.02128
Pump P_1	$\{P_1, P_2\}$	0.04791	0.04877	0.04787
Chiller C_1	$\{C_1, C_2\}$, $\{C_1, C_3\}$	0.25448	0.26416	0.26596
Controller K	$\{K\}$	0.53724	0.54186	0.53191

12.3.3 Risk achievement worth

The risk achievement worth asks how much greater the system failure probability would become if a particular component were unavailable with certainty. It is the ratio

$$\mathrm{RAW}_i = \frac{Q(1_i)}{Q_{\mathrm{sys}}}, \tag{12.21}$$

a multiplicative factor. A value $\mathrm{RAW}_i = 5$ means that forcing the basic event to occur raises the modeled system failure probability to five times its baseline. In a coherent system $Q(1_i) \geq Q_{\mathrm{sys}}$, so $\mathrm{RAW}_i \geq 1$; and since $Q(1_i) \leq 1$, $\mathrm{RAW}_i \leq 1/Q_{\mathrm{sys}}$, the value reached by a component whose

failure alone causes the top event. RAW identifies the components whose loss would leave the system in a highly vulnerable condition, which is the question asked in outage planning, maintenance scheduling and configuration risk management.

For chiller 1, the baseline unavailability is $q_{C_1} = 0.05$. Forcing $q_{C_1} = 1$ leaves two chillers, both of which must now operate, so the chiller branch fails if either remaining chiller fails:

$$Q_C(1_{C_1}) = 1 - (1 - q_{C_2})(1 - q_{C_3}) = 1 - (0.95)^2 = 0.0975,$$

thirteen times the baseline $Q_C = 0.00725$. The system failure probability becomes

$$Q(1_{C_1}) = 1 - (0.9996)(0.9991)(1 - 0.0975)(0.99) = 1 - 0.892314 = 0.107686,$$

and

$$\mathrm{RAW}_{C_1} = \frac{0.107686}{0.0184548} = 5.84.$$

Chiller 1 has only a 5 % baseline unavailability, and the plant tolerates its loss without losing the cooling function. With chiller 1 out of service, however, the plant is close to six times as likely to be without cooling, because a single further chiller failure now suffices. RAW captures this conditional vulnerability, which the chiller's own unavailability does not show. For the controller the ratio reaches its upper bound, $\mathrm{RAW}_K = 1/0.0184548 = 54.19$, because its failure alone causes the top event.

12.3.4 Risk achievement worth and maintenance configurations

Suppose the operator considers taking chiller 1 offline for preventive maintenance. The baseline system unavailability is $Q_{\mathrm{sys}} = 1.8455\,\%$; in the configuration with chiller 1 unavailable it is $Q(1_{C_1}) = 10.7686\,\%$. If the maintenance is carried out at a randomly selected instant, with the other components in their usual states, the plant is substantially more vulnerable while the chiller is out.

Figure 12.3 shows this as a configuration risk profile over three days. The horizontal axis is time in hours; the vertical axis is the snapshot unavailability of the cooling function under the configuration in force, in percent. The plant runs at the baseline 1.8455 % for the first day, at 10.7686 % during a 24 h window in which chiller 1 is out of service, and at the baseline again afterwards. The step height is the RAW ratio in absolute form, $Q(1_{C_1}) - Q_{\mathrm{sys}} = (1 - q_{C_1})\, I^{\mathrm{B}}_{C_1} = 8.92$ percentage points. If the other components remain in their stationary regime during the window, the expected time without cooling during the window is the area under the profile: $0.107686 \times 24\,\mathrm{h} = 2.58\,\mathrm{h}$, against $0.018455 \times 24\,\mathrm{h} = 0.44\,\mathrm{h}$ for the same 24 h at the baseline, an increase of 2.14 h, the shaded area.

These are snapshot probabilities, and the figure is only as good as the assumption that the other components keep their stationary unavailabilities throughout the window. A real maintenance window has a start chosen by the operator, who will typically confirm that chillers 2 and 3 are running before releasing chiller 1, so at the start of the window the conditional unavailability is lower than 10.77 % and rises toward it as failures accumulate. Evaluating an actual maintenance period requires a time-dependent model of its duration and of the failures and repairs of the other components during it, the subject of a later chapter on repairable systems and Markov models. RAW identifies the significance of the component's unavailability; it does not, by itself, give the probability of losing cooling at some time during a multihour window.

12.3.5 Risk reduction worth

The risk reduction worth asks by what factor the present system failure probability could be reduced if a particular basic event became impossible. It is defined as

$$\mathrm{RRW}_i = \frac{Q_{\mathrm{sys}}}{Q(0_i)}, \qquad Q(0_i) > 0, \tag{12.22}$$

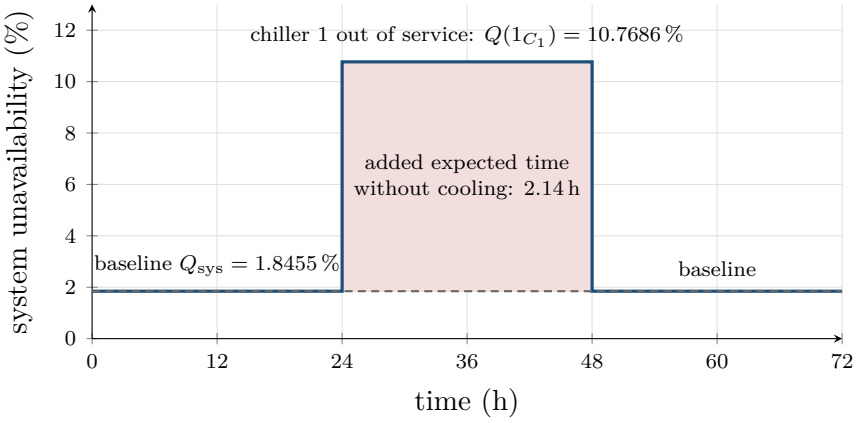


Figure 12.3: Configuration risk profile of the cooling plant over three days with chiller 1 out of service from hour 24 to hour 48 (illustrative). Outside the window the snapshot unavailability is the baseline 0.0184548; inside it is $Q(1_{C_1}) = 0.107686$, a factor $\text{RAW}_{C_1} = 5.84$ higher. The shaded area, $(0.107686 - 0.018455) \times 24\,\text{h} = 2.14\,\text{h}$, is the added expected time without cooling if the other components stay in their stationary regime.

so that $\text{RRW}_i = 2$ means that making the basic event impossible would halve the system failure probability. The two ratio measures look in opposite directions: RAW examines the consequence of certain failure, RRW the maximum benefit of eliminating failure.

For the controller, with $Q_{\text{sys}} = 0.0184548$ and $Q(0_K) = 0.0085402$,

$$\text{RRW}_K = \frac{0.0184548}{0.0085402} = 2.161.$$

If controller failure could be eliminated completely, the system failure probability would fall by a factor of about 2.16. Equivalently, it would fall by 53.72 %, the Fussell–Vesely value of Section 12.3.1.

The equivalence is exact. From equation (12.18), $Q(0_i)/Q_{\text{sys}} = 1 - I_i^{\text{FV}}$, and inverting,

$$\text{RRW}_i = \frac{1}{1 - I_i^{\text{FV}}}. \tag{12.23}$$

For the controller, $1/(1 - 0.5372) = 2.161$. The function $x \mapsto 1/(1-x)$ is increasing on $[0, 1)$, so for the same baseline and the same top event RRW and Fussell–Vesely contain the same ranking information: a component with the higher Fussell–Vesely importance also has the higher RRW. The two measures are two expressions of one counterfactual, a fraction removed and a factor of reduction; their interpretations differ, but their rankings do not.

12.4 Comparison of the four measures

With the four measures defined, they can be computed for every component of the plant and set side by side. This section tabulates the two counterfactual evaluations for each component type, collects the four measures in one table and two charts, and derives the relations that express all of them through the Birnbaum importance.

12.4.1 The counterfactual calculations

The two feeders are identical and play symmetric roles, so they have equal importance; so do the two pumps and the three chillers. One representative of each type is enough. For each, two

counterfactual evaluations of equation (12.4) are needed, $Q(0_i)$ and $Q(1_i)$, from which

$$I_i^{\mathrm{B}} = Q(1_i) - Q(0_i), \quad I_i^{\mathrm{FV}} = \frac{Q_{\mathrm{sys}} - Q(0_i)}{Q_{\mathrm{sys}}}, \quad \mathrm{RAW}_i = \frac{Q(1_i)}{Q_{\mathrm{sys}}}, \quad \mathrm{RRW}_i = \frac{Q_{\mathrm{sys}}}{Q(0_i)}. \tag{12.24}$$

Table 12.4 lists the counterfactual values. For feeder 1 the arithmetic is: with $q_{E_1} = 0$ the feeder branch cannot fail, $Q_E = 0$, and $Q(0_{E_1}) = 1 - (0.9991)(0.99275)(0.99) = 1 - 0.981938 = 0.0180620$; with $q_{E_1} = 1$ the branch fails when feeder 2 does, $Q_E = 0.02$, and $Q(1_{E_1}) = 1 - (0.98)(0.981938) = 0.0377008$. The other rows follow in the same way, by setting the branch unavailability to its value with the chosen event absent or certain: $Q_P = 0$ or 0.03 for a pump, $Q_C = 0.0025$ or 0.0975 for a chiller, $q_K = 0$ or 1 for the controller.

Table 12.4: Counterfactual evaluations of the cooling-plant failure probability for one representative basic event of each type, with the branch unavailability each implies. Baseline $Q_{\mathrm{sys}} = 0.0184548$.

Basic event	q_i	Branch with $q_i = 0$	$Q(0_i)$	Branch with $q_i = 1$	$Q(1_i)$
One feeder	0.02	$Q_E = 0$	0.0180620	$Q_E = 0.02$	0.0377008
One pump	0.03	$Q_P = 0$	0.0175706	$Q_P = 0.03$	0.0470435
One chiller	0.05	$Q_C = 0.0025$	0.0137584	$Q_C = 0.0975$	0.1076862
Controller	0.01	$q_K = 0$	0.0085402	$q_K = 1$	1.0000000

The controller has $Q(1_K) = 1$ because its failure alone guarantees the top event. For one chiller, $Q(1_{C_1})$ is far below one because the other two chillers may still function. The chiller's $Q(0_{C_1})$ is the lowest of the three redundant components because removing one chiller's failures turns the 2-out-of-3 chiller branch into one that fails only when both of the other chillers fail, $Q_C = 0.05^2 = 0.0025$, a factor of three below the baseline 0.00725.

12.4.2 The complete results

Table 12.5 applies equations (12.24) to Table 12.4. All values come from the exact Boolean model, including the overlap among cut sets. The controller ranks first on every measure, followed by a chiller, a pump and a feeder.

Table 12.5: The four importance measures for one basic event of each type in the cooling plant, exact model with independent basic events, $Q_{\mathrm{sys}} = 0.0184548$ (illustrative values).

Basic event	q_i	I_i^{B}	I_i^{FV}	RAW_i	RRW_i
One feeder	0.02	0.01964	0.0213 (2.13 %)	2.04	1.022
One pump	0.03	0.02947	0.0479 (4.79 %)	2.55	1.050
One chiller	0.05	0.09393	0.2545 (25.45 %)	5.84	1.341
Controller	0.01	0.99146	0.5372 (53.72 %)	54.19	2.161

Three readings follow from the table. First, the component with the highest unavailability, the chiller at 0.05, is not the most important on any measure; the controller, with the lowest unavailability, is first on all four, because it is the only first-order cut set. Second, the measures span very different ranges: the Birnbaum importance of the controller is fifty times the feeder's,

while its RRW is only about twice the feeder's, because RRW compresses large reductions into modest factors. Third, the Fussell–Vesely values do not add to one, and are not meant to: $0.0213 \times 2 + 0.0479 \times 2 + 0.2545 \times 3 + 0.5372 = 1.44$, since a top event can involve several basic events at once.

12.4.3 Rankings in charts

Figures 12.4 and 12.5 display two of the columns of Table 12.5. In both, the horizontal axis lists one basic event of each type, in decreasing order of importance, and the vertical axis is the measure, starting at zero; each bar carries its value. Figure 12.4 is the Fussell–Vesely importance in percent, the contribution of each event to the present unavailability. Figure 12.5 is the risk achievement worth, the vulnerability if the component is unavailable, with a dashed line at RAW = 1, the value of a component whose certain failure would change nothing.

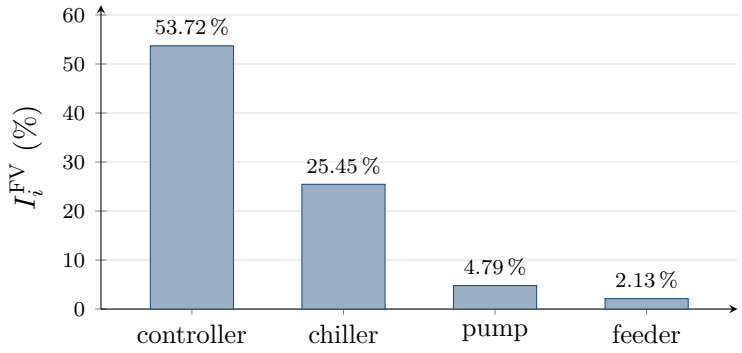


Figure 12.4: Fussell–Vesely importance of one basic event of each type in the cooling plant, $I_i^{\mathrm{FV}} = [Q_{\mathrm{sys}} - Q(0_i)]/Q_{\mathrm{sys}}$ (Table 12.5): the fraction of the present snapshot unavailability removed if the event were made impossible. Illustrative values.

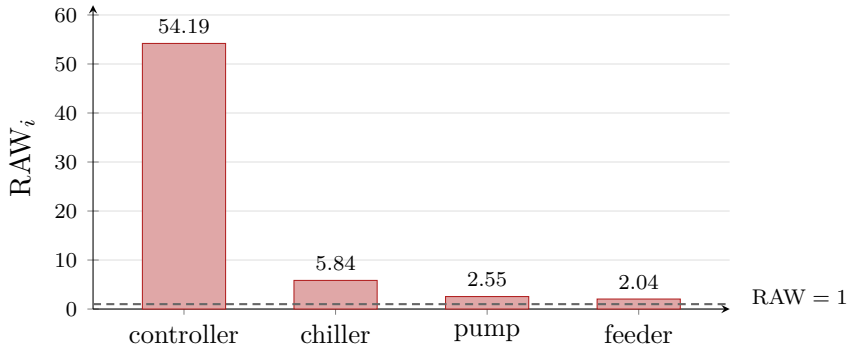


Figure 12.5: Risk achievement worth of one basic event of each type in the cooling plant, $\mathrm{RAW}_i = Q(1_i)/Q_{\mathrm{sys}}$ (Table 12.5): the factor by which the snapshot unavailability grows if the component is certainly unavailable. The dashed line marks RAW = 1, no change. Illustrative values.

Both charts place the controller first, followed by the chiller, the pump and the feeder. The measures do not always produce the same ranking, however, and this example agrees only because the controller is indispensable and its unavailability, 0.01, though the lowest in the plant, is large compared with the probabilities of the second-order cut sets, 0.0004 to 0.0025. A component with a very small baseline unavailability can have a high RAW, because its failure would be catastrophic for the system, and a low Fussell–Vesely importance, because it rarely fails in the baseline model; the illustrative component with $q_i = 10^{-7}$ of Section 12.3.2 has

RAW = 54.19, equal to the controller's, and $I^{\mathrm{FV}} = 5.3 \times 10^{-6}$, far below the feeder's. The rankings also move with the parameters: if the chiller unavailability doubled to 0.10, the chiller's Fussell–Vesely importance would rise to 0.457 and overtake the controller's, which would fall to 0.249, while the controller would keep the higher RAW (Exercise 12.3). The distinction matters most when highly reliable but indispensable components are compared with more failure-prone redundant ones.

12.4.4 Relations among the four measures

All four measures can be written through the Birnbaum importance, the event probability q_i and the baseline Q_{sys}. The pivotal decomposition (12.11) gives two differences. The first is equation (12.19), $Q_{\mathrm{sys}} - Q(0_i) = q_i\, I_i^{\mathrm{B}}$. For the second, subtract the decomposition from $Q(1_i)$:

$$Q(1_i) - Q_{\mathrm{sys}} = Q(1_i) - (1-q_i)\,Q(0_i) - q_i\,Q(1_i) = (1-q_i)\big[Q(1_i) - Q(0_i)\big] = (1-q_i)\,I_i^{\mathrm{B}}. \quad (12.25)$$

The interval from $Q(0_i)$ to $Q(1_i)$, of length I_i^{B}, is split by the baseline in the proportion $q_i : (1-q_i)$, as Figure 12.2 shows. Dividing the two differences by Q_{sys},

$$I_i^{\mathrm{FV}} = \frac{q_i\, I_i^{\mathrm{B}}}{Q_{\mathrm{sys}}}, \quad (12.26)$$

$$\mathrm{RAW}_i = 1 + \frac{(1-q_i)\, I_i^{\mathrm{B}}}{Q_{\mathrm{sys}}}, \quad (12.27)$$

$$\mathrm{RRW}_i = \frac{Q_{\mathrm{sys}}}{Q_{\mathrm{sys}} - q_i\, I_i^{\mathrm{B}}}. \quad (12.28)$$

For the chiller, equation (12.27) gives $1 + 0.95 \times 0.093928/0.0184548 = 1 + 4.835 = 5.835$, the value of Table 12.5. Eliminating I_i^{B} between equations (12.26) and (12.27) relates the two measures that rank differently:

$$\mathrm{RAW}_i - 1 = \frac{1-q_i}{q_i}\, I_i^{\mathrm{FV}}. \quad (12.29)$$

For a fixed Fussell–Vesely importance, RAW grows as the odds $q_i/(1-q_i)$ of the event fall: a very reliable event has a large RAW for every unit of present contribution. For the controller, $1 + (0.99/0.01) \times 0.5372 = 54.19$. This is the mathematical form of the observation that rare but indispensable components rank high on RAW and low on Fussell–Vesely.

Read through these formulas, the four measures are four perspectives on the same probability model. The Birnbaum importance is the sensitivity, the probability that the event is critical. The Fussell–Vesely importance weights that sensitivity by how often the event occurs. RAW weights it by how often the event does not occur and reports the result as a factor of increase, the consequence of forcing failure. RRW reports the Fussell–Vesely reduction as a factor. Their ranges are $0 \le I_i^{\mathrm{B}} \le 1$, $0 \le I_i^{\mathrm{FV}} \le 1$, $1 \le \mathrm{RAW}_i \le 1/Q_{\mathrm{sys}}$ and $\mathrm{RRW}_i \ge 1$ in a coherent system.

12.5 Importance measures and engineering decisions

An importance ranking identifies where the system has leverage; it does not by itself rank engineering projects. This section evaluates four improvement projects directly, states the general decision formula, turns the improvement calculation into a worked table, and compares candidate improvements by unreliability reduction per unit cost.

12.5.1 A highly important component is not automatically the best investment

Suppose money can be spent on one of four improvements, each of which halves the unavailability of one type of equipment: the controller from 1 % to 0.5 %; every chiller from 5 % to 2.5 %; every pump from 3 % to 1.5 %; or every feeder from 2 % to 1 %. The benefit of each is found by recomputing the system failure probability (12.4) for the proposed design. For the chiller project, $Q_C = 3(0.025)^2 - 2(0.025)^3 = 0.0018438$ and $Q_{\text{sys}} = 1 - (0.9996)(0.9991)(1 - 0.0018438)(0.99) = 0.0131096$, a reduction of $0.0184548 - 0.0131096 = 0.0053452$, or 0.5345 percentage points.

Table 12.6 gives the four results. Its columns are the new system unavailability, the absolute reduction in percentage points, the reduction as a fraction of the baseline, and the first-order estimate $\sum_i I_i^{\text{B}} \, \Delta q_i$ that the Birnbaum importances of Table 12.5 would predict.

Table 12.6: Four improvement projects, each halving the unavailability of every component of one type, evaluated by recomputing equation (12.4). Baseline 1.8455 %; illustrative values.

Improvement	New Q_{sys} (%)	Reduction (points)	Relative reduction	First-order estimate (points)
Controller, 1 % to 0.5 %	1.3498	0.4957	26.9 %	0.4957
Three chillers, 5 % to 2.5 % each	1.3110	0.5345	29.0 %	0.7045
Two pumps, 3 % to 1.5 % each	1.7792	0.0663	3.6 %	0.0884
Two feeders, 2 % to 1 % each	1.8160	0.0295	1.6 %	0.0393

The chiller project produces the largest reduction, although the controller has the highest Fussell–Vesely importance and RRW of any single component. The reason is that the chiller project improves three basic events at once, while each importance measure in Table 12.5 describes one basic event. Individual component importance does not rank multi-component projects; the proposed intervention itself must be evaluated.

The last column shows a second reason. For the controller, the first-order estimate $I_K^{\text{B}} \times 0.005 = 0.4957$ points is exact, because Q is linear in the single probability q_K (Section 12.2.1). For the chillers, the estimate $3 \times 0.093928 \times 0.025 = 0.7045$ points overstates the true 0.5345 by a third: when all three chiller probabilities fall together, each chiller's Birnbaum importance falls with them, since each is critical only when another chiller is down. The Birnbaum importance is a local, one-at-a-time sensitivity, and a large change of several probabilities must be evaluated on the model.

12.5.2 The general decision formula

Let a denote an engineering action, such as installing a redundant controller or replacing several chillers. The action changes the model parameters, $\mathbf{q} \to \mathbf{q}(a)$, and possibly the logic. Its effect on the system failure probability is the absolute reduction

$$\Delta Q(a) = Q(\mathbf{q}) - Q(\mathbf{q}(a)), \tag{12.30}$$

computed with the complete model for the modified system. A risk-based decision also needs consequences and costs. A simplified net benefit is

$$B_{\text{net}}(a) = \mathbb{E}[C \mid \text{baseline}] - \mathbb{E}[C \mid a] - C_{\text{impl}}(a), \tag{12.31}$$

where $\mathbb{E}[C \mid \cdot]$ is the expected consequence of system failures over the planning horizon, in the system as it is and with the action taken, and $C_{\text{impl}}(a)$ is the cost of implementing the action. All three terms must refer to the same planning horizon, and the expected consequences must reflect the relevant outage scenarios and be expressed in the same monetary unit. As an illustrative case, an action that lowers the expected production losses from cooling outages by R\$ 500 000 per year and costs R\$ 1 000 000 up front pays for itself in two years if nothing else changes, but the decision must also account for the equipment lifetime, discounting, maintenance costs and any further benefits or new failure modes.

Importance measures are therefore screening and prioritization tools. They show where the model has leverage and which components the present risk level depends on; they are not an optimization of engineering actions, which requires equations (12.30) and (12.31).

Principle 12.2. *Importance identifies leverage; decisions require consequences, feasibility and cost. Every candidate action is evaluated by recomputing the system failure probability for the modified model, not read off an importance ranking.*

12.5.3 A worked table of component improvements

The improvement experiment generalizes Table 12.6: choose one type of equipment, reduce the unavailability of every component of that type by a fraction f, and recompute the system failure probability. The model is equation (12.4) with q_E, q_P, q_C or q_K replaced by $(1 - f)$ times its baseline value, all components of a type modified together, the logic unchanged, independence kept, and no new dependence or failure mode introduced by the improvement. Table 12.7 evaluates it for $f = 25\,\%$, $50\,\%$, $75\,\%$ and $100\,\%$. Its three blocks give the new system unavailability, the absolute reduction and the reduction relative to the baseline $1.8455\,\%$.

Table 12.7: System unavailability after reducing the unavailability of every component of one type by a fraction f, independent basic events, unchanged logic (illustrative values from Table 12.2).

Equipment improved	$f = 25\,\%$	$f = 50\,\%$	$f = 75\,\%$	$f = 100\,\%$
New system unavailability Q_{sys} (%)				
Shared controller	1.5976	1.3498	1.1019	0.8540
All three chillers	1.5354	1.3110	1.1746	1.1287
Both pumps	1.8068	1.7792	1.7626	1.7571
Both feeders	1.8283	1.8160	1.8087	1.8062
Absolute reduction (percentage points)				
Shared controller	0.2479	0.4957	0.7436	0.9915
All three chillers	0.3101	0.5345	0.6709	0.7168
Both pumps	0.0387	0.0663	0.0829	0.0884
Both feeders	0.0172	0.0295	0.0368	0.0393
Relative reduction (% of baseline)				
Shared controller	13.43	26.86	40.29	53.72
All three chillers	16.80	28.96	36.35	38.84
Both pumps	2.10	3.59	4.49	4.79
Both feeders	0.93	1.60	2.00	2.13

One row of the arithmetic, the chillers at $f = 75\,\%$: the chiller unavailability becomes $0.25 \times 0.05 = 0.0125$, so $Q_C = 3(0.0125)^2 - 2(0.0125)^3 = 0.00046875 - 0.0000039 = 0.0004648$,

and $Q_{\text{sys}} = 1 - (0.9996)(0.9991)(0.9995352)(0.99) = 0.0117462$, a reduction of 0.6709 percentage points, 36.35 % of the baseline.

Three trends appear. First, the controller row is linear in f: each 25 % step removes the same 0.2479 points, because Q is linear in q_K. The chiller row shows diminishing returns: the first quarter removes 0.3101 points, the last only 0.0459, because Q_C is quadratic in q_C near zero. The two rows therefore cross: the chiller project is better at 25 % and 50 %, the controller project at 75 % and 100 %. Which component is "the best to improve" depends on how large an improvement is feasible. Second, at $f = 100\,\%$ the controller's relative reduction is its Fussell–Vesely importance, 53.72 %, while the chillers' is 38.84 %, larger than one chiller's 25.45 % because three events are removed together; and the pump and feeder rows equal one pump's and one feeder's Fussell–Vesely values, 4.79 % and 2.13 %, because removing one pump's failures already removes the pump branch. Third, no improvement of one equipment type can remove the failure pathways that run through the others: even a perfect controller leaves 0.854 %, the contribution of the feeders, chillers and pumps.

12.5.4 Unreliability reduction per unit cost

The decision formula (12.31) needs costs. Table 12.8 attaches illustrative implementation costs to six candidate actions, the four improvements of Table 12.6 and two structural changes: a redundant controller, a second controller in parallel with the first so that both must fail, which with independent controllers lowers the controller branch to $q_K^2 = 0.0001$; and a fourth chiller, which turns the chiller branch into 2-out-of-4 required, failing when at least three of four chillers are unavailable, $Q_C = 4q_C^3(1 - q_C) + q_C^4 = 0.00048125$. For each action the table gives the new system unavailability, the absolute reduction $\Delta Q(a)$, the reduction per R$ 100 000 of cost, and the expected hours of cooling unavailability avoided per year, 8760 h $\times\, \Delta Q(a)$, which uses the snapshot unavailability as the long-run fraction of time unavailable. The costs are invented for the illustration and labeled as such.

Table 12.8: Candidate improvements of the cooling plant ranked by unreliability reduction and by reduction per unit cost. Model of equation (12.4) with independent basic events; costs illustrative.

Action a	New Q_{sys} (%)	$\Delta Q(a)$ (points)	Cost (R$)	Points per R$ 100 000	Hours avoided per year
Controller, 1 % to 0.5 %	1.3498	0.4957	120 000	0.413	43.4
Redundant controller	0.8639	0.9815	250 000	0.393	86.0
Three chillers, 5 % to 2.5 %	1.3110	0.5345	450 000	0.119	46.8
Fourth chiller	1.1762	0.6692	600 000	0.112	58.6
Two pumps, 3 % to 1.5 %	1.7792	0.0663	40 000	0.166	5.8
Two feeders, 2 % to 1 %	1.8160	0.0295	200 000	0.015	2.6

For the redundant controller the arithmetic is $Q_{\text{sys}} = 1 - 0.991460 \times (1 - 0.0001) = 0.0086394$, $\Delta Q = 0.0184548 - 0.0086394 = 0.0098154$, or 0.9815 points, which per R$ 100 000 of a R$ 250 000 cost is $0.9815/2.5 = 0.393$ points, and $8760 \times 0.0098154 = 86.0$ h of unavailability avoided per year.

Figure 12.6 draws the two rankings. Each panel lists the six actions on the vertical axis, in the order of Table 12.8; the upper panel's horizontal axis is the absolute reduction $\Delta Q(a)$ in percentage points, the lower panel's the reduction per R$ 100 000. By absolute reduction the order is redundant controller, fourth chiller, chiller upgrade, controller upgrade, pumps, feeders. By reduction per unit cost it is controller upgrade, redundant controller, pumps, chiller upgrade, fourth chiller, feeders. The pump upgrade, a minor contributor in absolute terms, is

more cost-effective than either chiller action; the chiller actions, which lead Table 12.6, fall to the middle once their cost is counted. With a budget, the decision is a selection among actions, possibly several, by their benefit against their cost, which none of the four importance measures expresses.

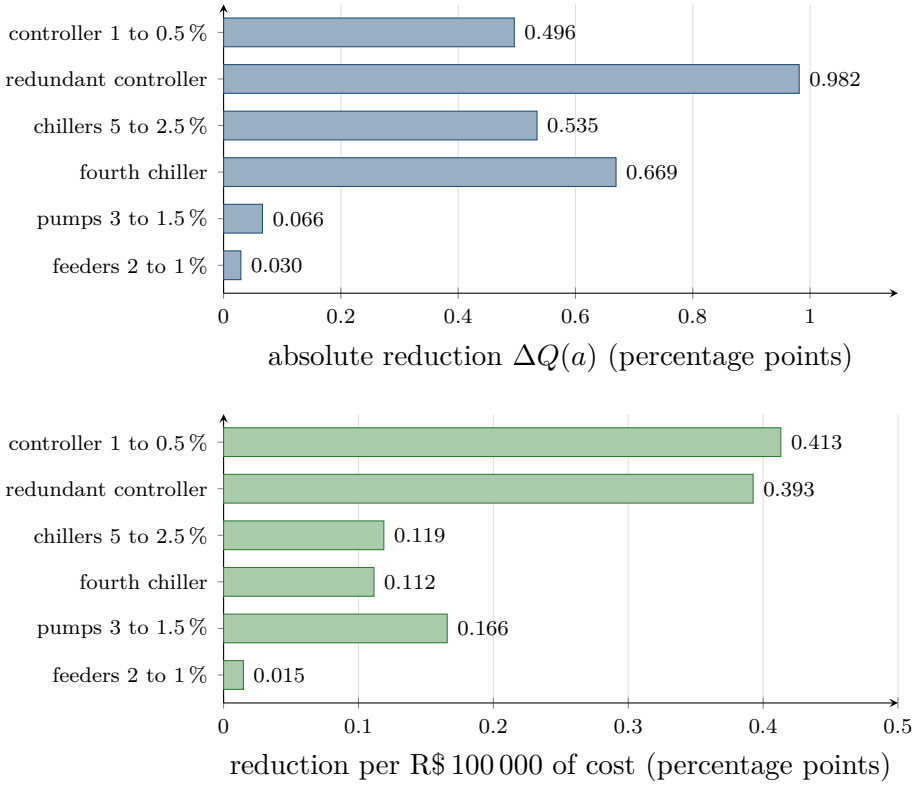


Figure 12.6: Six candidate improvements of the cooling plant (Table 12.8). Upper panel: absolute reduction of the snapshot unavailability, $\Delta Q(a) = Q(\mathbf{q}) - Q(\mathbf{q}(a))$. Lower panel: the same reduction per R$ 100 000 of illustrative implementation cost. The two rankings differ.

The redundant controller deserves a caution that anticipates the next section. Its benefit in Table 12.8 assumes that the two controllers fail independently. Two controllers of the same design, running the same software on the same power supply, are a textbook candidate for common-cause failure, and even a small common-cause fraction would remove much of the computed benefit.

12.6 Limitations of importance measures

Importance measures are conditional on the model that produces them: on its top event, its parameters and its independence assumptions. This section states four limits on their interpretation: the dependence on the chosen top event, the distinction between structural, reliability and uncertainty importance, the distinction between importance and the value of information, and the effect of dependent and common-cause failures.

12.6.1 Dependence on the top event

A component's importance is not an intrinsic property of the component. It is a property of the component's role in a model of one defined system outcome. Consider two top events for the same data center:

$$T_1 = \{\text{loss of cooling capacity}\}, \qquad T_2 = \{\text{thermal damage to IT equipment}\}.$$

The cooling controller is highly important for T_1, as this chapter has computed. The probability of T_2, however, also depends on the thermal inertia of the room, the protective shutdown of the IT equipment, the IT load, any backup cooling and the duration of the outage; a short controller outage that the room's thermal mass rides through causes T_1 but not T_2. The controller's importance for T_2 can therefore differ substantially from its importance for T_1. In the same way, a transmission-line outage can be important for the loss of transfer capacity across an interface and much less important for actual load shedding, when redispatch and network redundancy are available. The required function, set out in Chapter 2, determines the top event, and the top event determines the importance.

Principle 12.3. *Importance is always relative to a defined system outcome. A ranking computed for one top event does not transfer to another.*

12.6.2 Structural, reliability and uncertainty importance

Three related notions of importance are distinguished in the literature, and the Birnbaum importance sits between the first two. *Structural importance* depends on the system logic alone. For a coherent system of n basic events, the structural importance of event i, written I_i^{ϕ}, is the fraction of the 2^{n-1} states of the other events in which event i is critical; it equals the Birnbaum importance evaluated with every basic-event probability set to $1/2$. For the cooling plant, $n = 8$ and the other events have $2^7 = 128$ states. The controller is critical in 36 of them, a chiller in 18, a pump or a feeder in 12, so $I_K^{\phi} = 0.281$, $I_{C_1}^{\phi} = 0.141$ and $I_{P_1}^{\phi} = I_{E_1}^{\phi} = 0.094$; the structure alone places the controller first and does not separate pumps from feeders. The Birnbaum importance of Section 12.2 is a *reliability importance*: it depends on the structure and on the probabilities of the other events, which is why the pump (0.0295) ranks above the feeder (0.0196) once their partners' unavailabilities, 0.03 against 0.02, are known. The Birnbaum importance is therefore not purely a structural quantity.

Uncertainty importance asks a third question. The basic-event probabilities are themselves estimates, with epistemic uncertainty from finite operating records. A basic event may have a modest Birnbaum importance and a very uncertain probability, while another has a high Birnbaum importance and a precisely known probability. Which uncertainty to reduce first depends on how the uncertainty in each parameter propagates to the uncertainty in Q. For small parameter variations, a first-order expansion of Q about the estimates gives

$$\operatorname{Var}(Q) \approx \nabla Q^{\mathsf{T}}\, \Sigma_{\mathbf{q}}\, \nabla Q, \tag{12.32}$$

where ∇Q is the vector of sensitivities $\partial Q/\partial q_i$ and $\Sigma_{\mathbf{q}}$ is the covariance matrix of the uncertain basic-event probabilities. When the parameter uncertainties are independent, $\Sigma_{\mathbf{q}}$ is diagonal and

$$\operatorname{Var}(Q) \approx \sum_i \left(I_i^{\mathrm{B}}\right)^2 \operatorname{Var}(q_i). \tag{12.33}$$

When several identical components share one estimated parameter, as the three chillers would if their common unavailability q_C were estimated from pooled records, the sensitivity to that shared parameter is the sum of their Birnbaum importances, $3 \times 0.093928 = 0.2818$ for q_C, because one error moves all three probabilities together. As an illustration with invented standard deviations of 0.002 for q_K, 0.02 for q_C and 0.01 for q_P and q_E, each shared by its equipment type, equation (12.33) gives contributions $(0.99146 \times 0.002)^2 = 3.9 \times 10^{-6}$ from the controller, $(0.2818 \times 0.02)^2 = 3.18 \times 10^{-5}$ from the chillers, 3.5×10^{-7} from the pumps and 1.5×10^{-7} from the feeders, a total variance of 3.62×10^{-5} (standard deviation 0.0060). The chillers carry 88 % of the uncertainty in Q and the controller 11 %, the reverse of their reliability importance. Reliability importance asks which components most influence system

failure; uncertainty importance asks which uncertain inputs most influence the uncertainty of the prediction; the two need not rank alike. The first-order approximation suits small uncertainties and a smooth model; for broad or strongly skewed uncertainty distributions, direct propagation by sampling is preferable, and the Bayesian treatment of parameter uncertainty in Chapter 8 provides the distributions to propagate.

12.6.3 Importance and the value of information

An engineer may be able either to improve the cooling controller or to install a sensor that estimates its condition, and with it its unavailability, more accurately. The two actions are of different kinds. Improving the controller changes the physical system and may reduce the risk. Improving the estimate changes the evidence: it reduces the uncertainty about the risk without, by itself, changing any failure probability. A measurement is valuable when it can lead to a better subsequent decision, for instance by revealing that a planned replacement is unnecessary, or that an outage is imminent and the load should be shifted.

The value of information is defined on a decision model: it is the expected value of the best decision taken after the observation minus the expected value of the best decision taken without it, before the observation's outcome is known. It depends on how likely the observation is to change the optimal action and on what that change is worth. A component with a high Birnbaum importance does not automatically have a high value of information: if its probability is already well known, or if no plausible measurement result would change the decision, measuring it is worth little. Importance measures are computed from the model; the value of information requires a model of decisions and their consequences on top of it, the subject of a later chapter on observability and the value of information.

12.6.4 Dependence and common-cause failures

The numerical work of this chapter assumes independent basic events. If the three chillers share an electrical bus or a cooling-water source, their unavailabilities can be strongly dependent, and

$$\Pr(C_1 \cap C_2) \neq \Pr(C_1)\,\Pr(C_2).$$

An analysis that ignores such common-cause failures overstates the benefit of redundancy, as the redundant controller of Section 12.5.4 illustrates, and misrepresents the importance rankings: the chiller cut sets would carry more probability than the independent model assigns, and the importance of the shared bus or water source, absent from the model, would be invisible.

Dependence also affects the definitions themselves. The derivative $I_i^{\mathrm{B}} = \partial Q/\partial q_i$ of equation (12.9) presumes a clear statement of what changes when q_i changes and what stays fixed; with dependent events, setting q_i to 0 or 1 does not determine the conditional probabilities of the other events, and the identity $\Pr(T \mid X_i = 1) = Q(1_i)$ used in Section 12.2.1 fails. When a physical modification changes a shared cause or the joint distribution of failure events, the correct procedure is to update the complete probabilistic model and recompute the top-event probability, as equation (12.30) prescribes. The next chapter, Chapter 13, takes up dependent and common-cause failures and the conditional probabilities needed to model them.

12.7 What to retain

Definition 12.1 (Reliability importance measures)**.** For a system failure probability $Q(\mathbf{q})$ with baseline Q_{sys} and the counterfactual values $Q(1_i)$ and $Q(0_i)$ with basic event i certain or impossible, the Birnbaum importance $I_i^{\mathrm{B}} = \partial Q/\partial q_i = Q(1_i) - Q(0_i)$ measures sensitivity, the Fussell–Vesely importance $I_i^{\mathrm{FV}} = [Q_{\mathrm{sys}} - Q(0_i)]/Q_{\mathrm{sys}}$ and the risk reduction worth $\mathrm{RRW}_i = Q_{\mathrm{sys}}/Q(0_i)$ measure the improvement potential, and the risk achievement worth $\mathrm{RAW}_i =$

$Q(1_i)/Q_{\text{sys}}$ measures the vulnerability to the event's certain occurrence. Importance identifies leverage; decisions require consequences, feasibility and cost.

Component failure probability is not component importance: the controller, with the lowest unavailability in the cooling plant, ranks first on every measure, because it alone forms a first-order cut set. Rankings depend on the top event, on the parameters and on the independence assumptions, and a project that changes several basic events must be evaluated on the model rather than read off a ranking. Table 12.9 lists the chapter's terms, their meaning, expression and units, and the subsection that defines each.

Table 12.9: The terms of this chapter and where each is defined.

Term	Meaning	Expression	Units	Defined in
Basic-event probability	probability of a bottom event of the tree	q_i	–	12.1.1
Snapshot unavailability	unavailable at a random instant	q_i, Q_{sys}	–	12.1.1
System failure probability	top-event probability of the model	$Q_{\text{sys}} = Q(\mathbf{q})$	–	12.1.2
Counterfactual values	event i impossible or certain	$Q(0_i)$, $Q(1_i)$	–	12.1.2
Pivotal decomposition	Q linear in each q_i	$Q(0_i) + q_i\,[Q(1_i) - Q(0_i)]$	–	12.2.1
Birnbaum importance	sensitivity; probability event i is critical	$\partial Q/\partial q_i = Q(1_i) - Q(0_i)$	–	12.2.1
Critical (pivotal) event	its state decides the top event	equation (12.14)	–	12.2.3
Fussell–Vesely importance	fraction removed if event impossible	$[Q_{\text{sys}} - Q(0_i)]/Q_{\text{sys}}$	–	12.3.1
Risk achievement worth	factor of increase if event certain	$Q(1_i)/Q_{\text{sys}}$	–	12.3.3
Configuration risk	unavailability in a given equipment configuration	$Q(1_i)$ during an outage	–	12.3.4
Risk reduction worth	factor of reduction if event impossible	$Q_{\text{sys}}/Q(0_i) = 1/(1 - I_i^{\text{FV}})$	–	12.3.5
Relations	all four through I_i^{B}, q_i, Q_{sys}	equations (12.26)–(12.29)	–	12.4.4
Benefit of an action	reduction from a model change	$\Delta Q(a) = Q(\mathbf{q}) - Q(\mathbf{q}(a))$	–	12.5.2
Net benefit	consequences avoided minus cost	$B_{\text{net}}(a)$, equation (12.31)	R$	12.5.2
Structural importance	share of critical states, logic only	I_i^{ϕ}	–	12.6.2
Uncertainty importance	contribution to $\text{Var}(Q)$	$(I_i^{\text{B}})^2\,\text{Var}(q_i)$	–	12.6.2
Value of information	decision value of an observation	decision model	R$	12.6.3

Further reading. Rausand, Barros and Høyland, *System Reliability Theory*, third edition, for reliability importance measures and coherent systems; Kumamoto and Henley, *Probabilistic Risk Assessment and Management for Engineers and Scientists*, second edition, for component impor-

tance and quantitative risk assessment; U.S. Nuclear Regulatory Commission, NUREG/CR-3385, *Measures of Risk Importance and Their Applications*, the foundational study of RAW, RRW and related measures; U.S. Nuclear Regulatory Commission, *Probabilistic Risk Assessment*, on the broader role of quantitative risk models in engineering decisions.

12.8 Exercises

Exercise 12.1. (a) State which importance measure is the derivative of the system failure probability with respect to a basic-event probability. (b) A basic event has a Fussell–Vesely importance of 0.40 under the counterfactual definition (12.18); state what this means, and compute its RRW. (c) A basic event has RAW = 10; state what this means. (d) State the relation between RRW and the Fussell–Vesely importance, and whether they can rank two components differently for the same top event. (e) Explain why a project that improves three chillers can reduce the system failure probability more than a project that improves the single most important component. (Answers: (a) the Birnbaum importance, $\partial Q/\partial q_i$; (b) making the basic event impossible would eliminate 40 % of the baseline system failure probability, not that the component fails 40 % of the time, and RRW $= 1/(1-0.40) = 1.667$; (c) forcing the event to occur makes the system failure probability ten times the baseline, $Q(1_i)/Q_{\text{sys}} = 10$; (d) $\text{RRW}_i = 1/(1 - I_i^{\text{FV}})$, an increasing function, so they rank identically; (e) a project's benefit depends on which and how many parameters it changes, and must be evaluated from its actual changes to the system model, not from single-component rankings.)

Exercise 12.2. A cooling function fails if both of two redundant pumps A and B fail or if their common electrical supply E fails, $T = AB + E$, with independent basic events $q_A = q_B = 0.05$ and $q_E = 0.01$. (a) Compute Q_{sys}. (b) For pump A and for the supply E, compute $Q(0_i)$, $Q(1_i)$ and the four importance measures. (c) Verify equations (12.26) and (12.27) for pump A. (Answers: (a) $1 - (1 - 0.0025)(0.99) = 0.012475$; (b) pump A: $Q(0_A) = 0.01$, $Q(1_A) = 1 - (0.95)(0.99) = 0.0595$, $I_A^{\text{B}} = 0.0495$, $I_A^{\text{FV}} = 0.1984$, $\text{RAW}_A = 4.77$, $\text{RRW}_A = 1.248$; supply E: $Q(0_E) = 0.0025$, $Q(1_E) = 1$, $I_E^{\text{B}} = 0.9975$, $I_E^{\text{FV}} = 0.7996$, $\text{RAW}_E = 80.16$, $\text{RRW}_E = 4.99$; (c) $0.05 \times 0.0495/0.012475 = 0.1984$ and $1 + 0.95 \times 0.0495/0.012475 = 4.77$.)

Exercise 12.3. In the cooling plant of Table 12.2, suppose aging raises the unavailability of every chiller from 0.05 to 0.10, all else unchanged. (a) Compute Q_C and Q_{sys}. (b) Compute the Birnbaum importance of chiller 1 from the probability that it is critical, and check it against $Q(1_{C_1}) - Q(0_{C_1})$. (c) Compute the four measures for chiller 1 and for the controller, and compare the rankings with Table 12.5. (Answers: (a) $Q_C = 3(0.01) - 2(0.001) = 0.028$ and $Q_{\text{sys}} = 0.038971$; (b) $2(0.10)(0.90) \times 0.988713 = 0.177968$, equal to $0.199142 - 0.021174$; (c) chiller 1: $I^{\text{B}} = 0.17797$, $I^{\text{FV}} = 0.4567$, RAW = 5.11, RRW = 1.841; controller: $I^{\text{B}} = 0.97074$, $I^{\text{FV}} = 0.2491$, RAW = 25.66, RRW = 1.332; the chiller now ranks first on Fussell–Vesely and RRW, while the controller keeps the higher Birnbaum importance and RAW, so the rankings by different measures disagree.)

Exercise 12.4. For the cooling plant of Table 12.2, two actions are proposed: (i) replace both pumps with units of unavailability 0.01 each, at an illustrative cost of R$ 60 000; (ii) improve the controller from 0.01 to 0.008, at R$ 50 000. (a) For each, compute the new Q_{sys}, the reduction $\Delta Q(a)$, the reduction per R$ 100 000, and the hours of unavailability avoided per year. (b) Compute the first-order estimate $\sum_i I_i^{\text{B}} \, \Delta q_i$ for each, and explain why it is exact for one action and not for the other. (c) State whether the Birnbaum importance of the controller tells the operator how valuable a better estimate of q_K would be. (Answers: (a) pumps: 0.017669, 0.000786 (0.0786 points), 0.131 points per R$ 100 000, 6.9 h per year; controller: 0.016472, 0.001983 (0.1983 points), 0.397 points per R$ 100 000, 17.4 h per year; (b) pumps $2 \times 0.029473 \times 0.02 = 0.00118$, controller $0.991460 \times 0.002 = 0.001983$; the controller estimate

is exact because Q is linear in the single probability q_K, while the pump estimate overstates the benefit because the two pump probabilities change together and each pump's Birnbaum importance is proportional to the other's unavailability; (c) no; the value of information requires a decision model, and depends on how likely a better estimate is to change the decision.)

Part IV

Dependencies and Probabilistic Risk Assessment

Chapter 13

Common-Cause and Dependent Failures

13.1 Redundancy and the independence assumption

Chapter 12 ranked the components of a system by their contribution to its failure probability. That ranking, like every quantification in Chapters 9 to 12, assumed that the basic component failures were statistically independent. Engineered systems rarely operate in fully independent conditions. Two pumps may share an electrical supply, a cooling-water circuit, a controller, an operating load, a maintenance procedure and a room. Two transmission lines may be crossed by the same storm. Redundant sensors may share a calibration procedure or a software defect. Each of these relations creates a dependence between failure events, and a dependence can change the reliability of a redundant system by a large factor while leaving the reliability of every component unchanged. The question of this chapter is therefore

why a redundant system can be much less reliable than the reliabilities of its components suggest, and how the dependence among failure events is represented mathematically.

The chapter answers it in six steps. It first states independence as a property of the joint probability distribution and derives the limits and the correlation measure of dependence between two failure events. It then separates the mechanisms that create dependence and builds the explicit common-cause event model, with its consequence for the fault tree and for the value of adding redundant units. It treats a shared environment, in which failures are independent given the environment but dependent without it; failure propagation and load sharing, in which one failure changes the hazard of the survivors; and hazards that depend on the physical conditions and on the failure history. It closes with engineering practice and with the size and the uncertainty of the risk that dependence assumptions carry.

Jargon: dependent failures, common-cause failure (CCF), root cause, coupling factor, cascading failure, load sharing

Dependent failures are failure events whose joint probability is not the product of their individual probabilities. A *common-cause failure* (CCF) is the failure, or the functional unavailability, of two or more items from one shared cause within a time interval that matters for the system's mission; it is the most important class of dependent failures in redundant systems. The *root cause* is the most basic reason for the failure that, if corrected, would prevent its recurrence (a contaminated lubricant batch, a wrong procedure). A *coupling factor* (also called a coupling mechanism) is the property that makes several items susceptible to the same root cause: identical design, a common location, a shared support system, the same maintenance crew. A *cascading failure* (failure propagation) is a failure of one item that causes the failure of another. *Load sharing* is the arrangement in which surviving items take over the load of failed ones, so that their stress, and their hazard, rise after each failure.

Key idea

Redundancy protects against independent failures, but it may provide little protection against failures arising from shared causes: the failure probability of a redundant system is fixed by the joint distribution of the component failures, and the component failure probabilities alone do not determine it.

Remark 13.1. In this chapter the capital letters A, B and C name events: the failure of pump (or chiller) A, of pump B, and a common-cause event C. They are not the availability A, the consequence C or the capacity C of the notation chapter, none of which appears here as a symbol. The complement of an event is written with a bar, $\bar{A}$, and the indicator of an event as $\mathbf{1}_A$, equal to 1 when A occurs and 0 otherwise.

13.1.1 Two redundant cooling pumps

The cooling-water pump of the running case is now installed as a pair. Two pumps, A and B, are installed in parallel, and either one can deliver the required flow of $100\,\mathrm{m}^3/\mathrm{h}$. With the events

$$A = \{\text{pump A unavailable}\}, \qquad B = \{\text{pump B unavailable}\}$$

over a specified mission without repair, the pumping system fails only if both pumps are unavailable, and its failure event is the intersection

$$F_{\text{sys}} = A \cap B. \tag{13.1}$$

Figure 13.1(a) draws this success logic as the reliability block diagram of Chapter 9: two blocks in parallel between an input and an output terminal, the system working if a path of working blocks joins them. Suppose that each pump fails during the mission with probability $\Pr(A) = \Pr(B) = 0.05$ (illustrative). If the two failures are independent, the probability of the intersection is the product of the two probabilities,

$$\Pr(F_{\text{sys}}) = \Pr(A)\,\Pr(B) = 0.05^2 = 0.0025, \tag{13.2}$$

a system failure probability of $0.25\,\%$, twenty times smaller than that of one pump.

13.1.2 Shared causes and the joint failure probability

Suppose now that both pumps can be disabled by one shared electrical fault, by a contaminated batch of lubricant drawn from a common supply, or by an extreme condition in the pump room. The parallel diagram of Figure 13.1(a) does not change, and neither does the Boolean statement

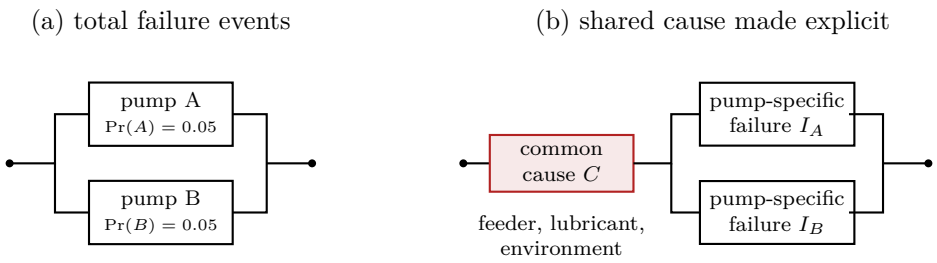


Figure 13.1: Two redundant cooling pumps, either sufficient. (a) The reliability block diagram on the total failure events A and B: nothing in it shows that the pumps share causes, and multiplying the two failure probabilities assumes that they share none. (b) The same system with the shared causes collected into one common-cause block C in series with the parallel pair of pump-specific failures I_A and I_B; this is the explicit common-cause model of Section 13.3.3.

of system failure: the system still fails exactly when both pumps are unavailable, $F_{\text{sys}} = A \cap B$. What changes is the probability of that intersection. The calculation of equation (13.2) is valid only if knowing that one pump has failed says nothing about the other, and a shared cause is precisely a reason for the failure of one pump to make the failure of the other more likely. The joint probability $\Pr(A \cap B)$ can then be substantially larger than the product $\Pr(A)\,\Pr(B)$ of the marginal (individual) probabilities.

This distinction is the foundation of dependent-failure modeling. The system logic, which says which combinations of component failures fail the system, and the probability model, which says how likely each combination is, are two separate parts of a reliability model. The logic is fixed by the functional architecture; the probability of the combinations is fixed by the joint distribution of the component failure events, and that distribution contains information that the marginal probabilities do not. The rest of the chapter states this information precisely, shows where it comes from physically, and shows how much it matters.

13.2 Mathematical dependence of failure events

This section states independence as a property of a joint distribution, works the conditional failure probability of the second pump, derives the bounds within which any joint failure probability must lie, and measures dependence by the correlation of the failure indicators.

13.2.1 Independence as a property of the joint distribution

Two events A and B are *statistically independent* if the probability of their intersection is the product of their probabilities,

$$\Pr(A \cap B) = \Pr(A)\,\Pr(B). \tag{13.3}$$

The conditional probability of B given A is defined, for $\Pr(A) > 0$, as

$$\Pr(B \mid A) = \frac{\Pr(A \cap B)}{\Pr(A)}, \tag{13.4}$$

the fraction of the cases in which A occurs that are also cases of B. Dividing equation (13.3) by $\Pr(A)$ gives the equivalent statement $\Pr(B \mid A) = \Pr(B)$: independence means that learning that A has occurred gives no information about whether B occurs. Rearranging equation (13.4) gives the multiplication rule, which holds with or without independence,

$$\Pr(A \cap B) = \Pr(A)\,\Pr(B \mid A), \tag{13.5}$$

and for dependent events the conditional probability $\Pr(B \mid A)$ differs from the marginal $\Pr(B)$.

Independence is a property of the joint distribution of the two events, not of the hardware. It does not require the components to be physically separate, and physical separation does not guarantee it: two pumps in different rooms are still dependent if they are serviced by the same crew with the same faulty procedure, or if both rooms overheat in the same heat wave. Whether two failure events are independent is a statement about the probability model, to be justified by the absence of shared causes, not inferred from a drawing.

13.2.2 Conditional failure probability of the second pump

Keep $\Pr(A) = 0.05$, and suppose that, given that pump A has failed, pump B fails with probability $\Pr(B \mid A) = 0.20$, because some of the causes that fail pump A also act on pump B (illustrative). By the multiplication rule (13.5),

$$\Pr(A \cap B) = 0.05 \times 0.20 = 0.01,$$

a system failure probability of 1 %. Table 13.1 compares the two models. Each row is one model; the columns give the conditional probability of the second failure, the joint failure probability, and its ratio to the independent value. The dependent model predicts four times the system failure probability of the independent one, yet the failure probability of pump A has not changed: what has changed is the probability of the joint failure.

Table 13.1: System failure probability of two redundant pumps, each with $\Pr(A) = \Pr(B) = 0.05$, under independence and under an illustrative dependence $\Pr(B \mid A) = 0.20$.

Model	$\Pr(B \mid A)$	$\Pr(A \cap B)$	Ratio to independent
Independent pumps	0.05	0.0025 (0.25 %)	1
Dependent pumps	0.20	0.0100 (1.00 %)	4

The denominator of a conditional probability is what gives it its meaning. To compute $\Pr(B \mid A)$ the attention is restricted to the cases in which A has occurred, and the question is what fraction of those cases are also cases of B. Figure 13.2 shows both steps for the dependent pump model. Panel (a) is the joint probability table: its columns are the two outcomes of pump A (failed, A, or not, $\bar{A}$), its rows the two outcomes of pump B, each cell holds the probability of one combination, and the margins hold the row and column totals. The four cells follow from the three given numbers: $\Pr(A \cap B) = 0.01$; $\Pr(A \cap \bar{B}) = \Pr(A) - \Pr(A \cap B) = 0.04$; $\Pr(\bar{A} \cap B) = \Pr(B) - \Pr(A \cap B) = 0.04$; and $\Pr(\bar{A} \cap \bar{B}) = 1 - 0.05 - 0.05 + 0.01 = 0.91$. The shaded column is the restricted sample space of the conditional probability. Panel (b) draws conditional probabilities as bars: each outlined bar is one conditioning event, rescaled to length one, and the filled part is the fraction of it in which B fails. Given A, B fails in $0.01/0.05 = 0.20$ of the cases in the dependent model, against 0.05 in the independent one. Given $\bar{A}$, B fails in $0.04/0.95 = 0.042$ of the cases, slightly less than its marginal 0.05: when the marginal probability is held fixed, a positive dependence raises the probability of the second failure after the first and lowers it after a success.

13.2.3 The Fréchet bounds

Dependence is not unlimited. With fixed marginal probabilities $p_A = \Pr(A)$ and $p_B = \Pr(B)$, the joint probability can take only values inside an interval, and the interval follows from two elementary facts. Since $A \cap B$ is contained in A and in B, its probability cannot exceed either, $\Pr(A \cap B) \leq \min(p_A, p_B)$. Since the union of the two events has probability at most one,

(a) joint probability table, dependent model

	A (A failed)	$\bar{A}$ (A works)	total
B	0.01	0.04	0.05
$\bar{B}$	0.04	0.91	0.95
total	0.05	0.95	1

shaded column: restricted sample space given A

(b) conditional failure probability of pump B

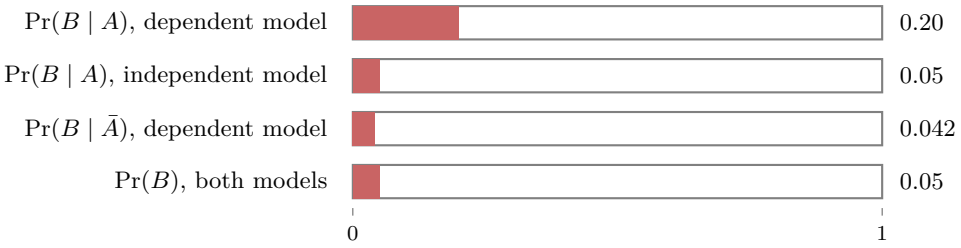


Figure 13.2: The dependent pump model with $\Pr(A) = \Pr(B) = 0.05$ and $\Pr(B \mid A) = 0.20$. (a) Joint probability table; the shaded column is the conditioning event A, and the darker cell is the joint failure $A \cap B$. (b) Conditional probabilities as fractions of the conditioning event, each bar rescaled to length one, compared with the independent model, in which $\Pr(B \mid A) = \Pr(B) = 0.05$.

$p_A + p_B - \Pr(A \cap B) = \Pr(A \cup B) \leq 1$, which gives $\Pr(A \cap B) \geq p_A + p_B - 1$; and a probability is never negative. Together,

$$\max(0,\ p_A + p_B - 1)\ \leq\ \Pr(A \cap B)\ \leq\ \min(p_A, p_B), \tag{13.6}$$

the *Fréchet bounds* (also Fréchet–Hoeffding bounds). Both are attained: the upper bound when the less probable event is contained in the more probable one, so that the two fail together whenever the less likely one fails; the lower bound, when $p_A + p_B \leq 1$, when the two events are mutually exclusive.

For the two pumps, $p_A = p_B = 0.05$ and

$$0\ \leq\ \Pr(A \cap B)\ \leq\ 0.05 :$$

with the same 5 % component failure probability, the system failure probability can lie anywhere between zero and 5 %. Table 13.2 lists four points of that interval, with the conditional probability $\Pr(B \mid A) = \Pr(A \cap B)/0.05$ and the correlation ρ_{AB} defined in the next subsection. Mutually exclusive failures make joint failure impossible; independence gives 0.25 %; the moderate positive dependence of the previous subsection gives 1 %; perfectly coincident failures, in which the pumps always fail together, give 5 %, the failure probability of one pump, so that the second pump adds nothing. The marginal probabilities therefore leave the system failure probability anywhere between zero and twenty times its independent value, and a redundant system cannot be quantified without information about the joint behavior of its components.

13.2.4 Correlation of the failure indicators

A single number that measures dependence is the correlation of the failure indicators. Let

$$X_A = \mathbf{1}_A, \qquad X_B = \mathbf{1}_B$$

Table 13.2: The same marginal failure probability $p_A = p_B = 0.05$ under four dependence structures. The numbers in the first column mark the points of Figure 13.3.

Dependence structure	$\Pr(A \cap B)$	$\Pr(B \mid A)$	ρ_{AB}	Ratio to independent
1. Mutually exclusive failures	0	0	-0.053	0
2. Independent failures	0.0025	0.05	0	1
3. Moderate positive dependence	0.01	0.20	0.158	4
4. Perfectly coincident failures	0.05	1	1	20

be the binary random variables equal to 1 when the pump fails and 0 otherwise. An indicator has expectation $\mathbb{E}[X_A] = 1 \cdot p_A + 0 \cdot (1 - p_A) = p_A$, and since $X_A^2 = X_A$, its variance is $\mathrm{Var}(X_A) = \mathbb{E}[X_A^2] - p_A^2 = p_A(1 - p_A)$. The product $X_A X_B$ equals 1 only when both pumps fail, so $\mathbb{E}[X_A X_B] = \Pr(A \cap B)$, and the covariance is

$$\mathrm{Cov}(X_A, X_B) = \mathbb{E}[X_A X_B] - \mathbb{E}[X_A]\,\mathbb{E}[X_B] = \Pr(A \cap B) - p_A p_B, \tag{13.7}$$

the excess of the joint probability over its independent value. Dividing by the two standard deviations gives the Pearson correlation of the indicators,

$$\rho_{AB} = \frac{\Pr(A \cap B) - p_A p_B}{\sqrt{p_A(1 - p_A)\, p_B(1 - p_B)}}, \tag{13.8}$$

which for two binary variables is known as the *phi coefficient*. In terms of the four cells of the joint table, $p_{11} = \Pr(A \cap B)$, $p_{10} = \Pr(A \cap \bar{B})$, $p_{01} = \Pr(\bar{A} \cap B)$ and $p_{00} = \Pr(\bar{A} \cap \bar{B})$, the numerator equals $p_{11}p_{00} - p_{10}p_{01}$, which can be verified by substituting $p_{10} = p_A - p_{11}$, $p_{01} = p_B - p_{11}$ and $p_{00} = 1 - p_A - p_B + p_{11}$. For the table of Figure 13.2(a), $\rho_{AB} = (0.01 \times 0.91 - 0.04 \times 0.04)/(0.05 \times 0.95) = 0.0075/0.0475 = 0.158$.

Solving equation (13.8) for the joint probability gives

$$\Pr(A \cap B) = p_A p_B + \rho_{AB} \sqrt{p_A(1 - p_A)\, p_B(1 - p_B)}, \tag{13.9}$$

and for equal marginal probabilities $p_A = p_B = p$,

$$\Pr(A \cap B) = p^2 + \rho_{AB}\, p(1 - p). \tag{13.10}$$

With $p = 0.05$ and $\rho_{AB} = 0.20$ (illustrative),

$$\Pr(A \cap B) = 0.0025 + 0.20 \times 0.05 \times 0.95 = 0.0025 + 0.0095 = 0.012,$$

a system failure probability of 1.2 %, 4.8 times the independent result.

The correlation of two failure events cannot take every value between -1 and 1. Substituting the Fréchet bounds (13.6) into equation (13.8) gives its feasible range. For equal marginals $p \le 1/2$, the upper bound $\Pr(A \cap B) = p$ gives $\rho_{AB} = (p - p^2)/(p(1 - p)) = 1$, and the lower bound $\Pr(A \cap B) = 0$ gives $\rho_{AB} = -p^2/(p(1 - p)) = -p/(1 - p)$; for $p \ge 1/2$ the lower bound is $-(1 - p)/p$. Rare failure events can therefore be strongly positively correlated but only weakly negatively: for $p = 0.05$ the correlation cannot fall below -0.0526. For unequal marginals even the upper limit is below one. A correlation coefficient quantifies dependence, but it does not explain its physical origin, and the sections that follow replace it, wherever possible, by a model of the mechanism.

13.2.5 Joint failure probability against correlation

Equation (13.10) changes the dependence structure while leaving the component failure probability fixed, and so isolates the effect of dependence on the redundant pair. Figure 13.3 draws it for $p = 0.05$. The horizontal axis is the correlation ρ_{AB} over its feasible range, from -0.0526 to 1; the vertical axis is the joint failure probability. The straight line is equation (13.10), the dashed line at 0.05 is the upper Fréchet bound, and the lower bound, zero, is the horizontal axis. The numbered points are the four cases of Table 13.2, and point 5 is the example with $\rho_{AB} = 0.20$. The line runs from the lower Fréchet bound at the smallest feasible correlation to the upper bound at $\rho_{AB} = 1$; the independent value, point 2, sits near the bottom of the range, so that almost the whole feasible interval lies above it.

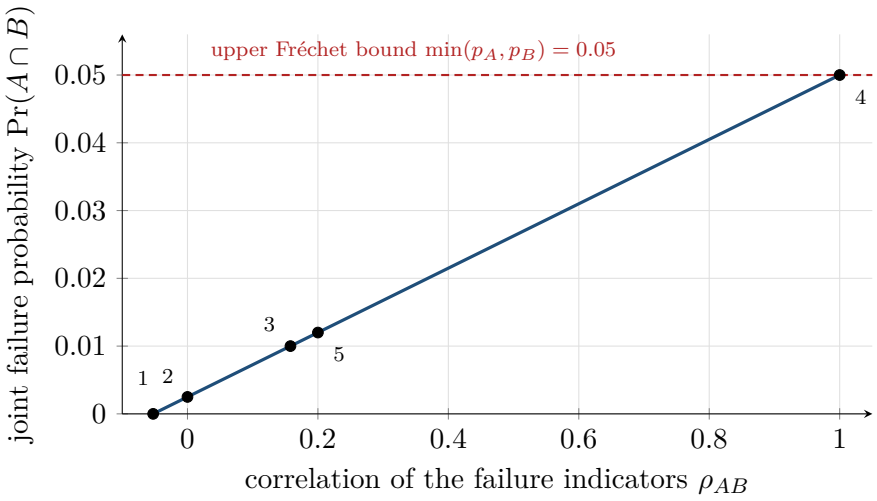


Figure 13.3: Joint failure probability of two units with equal marginal failure probability $p = 0.05$ against the correlation of their failure indicators, $\Pr(A \cap B) = p^2 + \rho_{AB}\, p(1-p)$, over the feasible range $-p/(1-p) \le \rho_{AB} \le 1$. Points: 1 mutually exclusive, 2 independent, 3 $\Pr(B \mid A) = 0.20$, 4 perfectly coincident, 5 $\rho_{AB} = 0.20$ (illustrative).

Table 13.3 extends the calculation to four component failure probabilities, $p = 0.01$, 0.05, 0.10 and 0.30, and six correlations, from the smallest feasible value to one. Each cell gives the joint failure probability of the pair and, in parentheses, its ratio to the independent value p^2. One cell shows the arithmetic: for $p = 0.01$ and $\rho_{AB} = 0.1$, $\Pr(A \cap B) = 0.0001 + 0.1 \times 0.01 \times 0.99 = 0.00109$, 10.9 times p^2. Dividing equation (13.10) by p^2 gives the ratio in closed form,

$$\frac{\Pr(A \cap B)}{p^2} = 1 + \rho_{AB}\,\frac{1-p}{p}, \tag{13.11}$$

which explains the trend down the columns and across the rows. A given correlation multiplies the joint failure probability more the rarer the component failures are: $\rho_{AB} = 0.1$ multiplies it by 1.23 when $p = 0.30$ and by 10.9 when $p = 0.01$. A correlation that looks small is therefore not a small effect for reliable components, and the independence assumption is least safe exactly where redundancy is most effective.

Principle 13.1. *Component failure probabilities alone do not determine the failure probability of a redundant system. The joint distribution of the failure events does, within the Fréchet bounds, and a fixed correlation of the failure indicators raises the joint failure probability by a factor* $1 + \rho_{AB}(1-p)/p$ *that grows as the components become more reliable.*

Table 13.3: Joint failure probability $\Pr(A \cap B) = p^2 + \rho_{AB}\, p(1-p)$ of two units with equal marginal failure probability p, and in parentheses its ratio to the independent value p^2. The first row is the smallest feasible correlation, $-p/(1-p)$, at which joint failure is impossible.

ρ_{AB}	$p = 0.01$	$p = 0.05$	$p = 0.10$	$p = 0.30$
minimum	0 ($\rho = -0.0101$)	0 ($\rho = -0.0526$)	0 ($\rho = -0.111$)	0 ($\rho = -0.429$)
0	0.0001 (1)	0.0025 (1)	0.0100 (1)	0.0900 (1)
0.1	0.00109 (10.9)	0.00725 (2.9)	0.0190 (1.9)	0.111 (1.23)
0.2	0.00208 (20.8)	0.0120 (4.8)	0.0280 (2.8)	0.132 (1.47)
0.5	0.00505 (50.5)	0.02625 (10.5)	0.0550 (5.5)	0.195 (2.17)
1	0.0100 (100)	0.0500 (20)	0.100 (10)	0.300 (3.33)

13.3 Types of dependence and the common-cause event model

Statistical dependence can arise from several physical mechanisms, and the remedy depends on which one is at work. This section classifies the mechanisms, fixes the meaning of common-cause failure, builds the explicit common-cause event model, and draws its consequences for the fault tree and for the value of adding redundant units.

13.3.1 Mechanisms of dependence

Reliability engineering distinguishes several mechanisms that make failure events statistically dependent. Table 13.4 lists six. Each row is one type; the second column gives its physical explanation and the third an example from the running cases. A common-cause failure is a shared cause that disables several components directly. Shared environmental exposure is a common varying environment to which several components respond. Failure propagation is the direct causation of one failure by another. Load sharing is the rise in stress on survivors after a failure. A functional dependency is the reliance of several components on one support item. A maintenance dependency is a procedure or an error repeated on several units.

Table 13.4: Mechanisms that create statistical dependence between failure events.

Dependency type	Physical explanation	Example
Common-cause failure	A shared cause disables several components	One contamination event damages two pumps
Shared environmental exposure	Components respond to the same varying environment	Extreme heat raises the failure propensity of several chillers
Failure propagation (cascading failure)	One component failure directly causes another	An electrical fault damages neighboring equipment
Load sharing	One failure increases the stress on the surviving components	The remaining pump operates at a higher load
Functional dependency	One component requires another to operate	Two pumps depend on one controller
Maintenance dependency	Common procedures or errors affect several units	An incorrect maintenance action performed on both redundant valves

The categories overlap. A shared electrical feeder is a functional dependency of both pumps: its failure stops both without either pump failing physically. A manufacturing defect in several pump motors is instead a common-cause failure of the motors themselves. The two cases can

produce the same loss of the pumping function, but they call for different remedies, a second feeder in the first case and a different motor supplier or an inspection program in the second, and a model that merges them hides the choice.

13.3.2 Terminology of common-cause failure

In probabilistic risk assessment (PRA), a common-cause failure denotes the failure or degradation of several components that are connected by a shared root cause and a coupling factor and that occur within a time window relevant to the mission. Three elements are therefore required: several affected components, a shared cause acting through a coupling factor, and a time relation among the failures. Common-cause failure data programs classify events by exactly these elements. The components that share a coupling factor and are analyzed together, the two pumps of this chapter for instance, form a *common-cause component group*.

A dependency is thus not necessarily a common-cause failure. A loss of the shared feeder is a dependency of the pumps on a support system, which a fault tree models as its own basic event; the failure of both motors from one defective winding batch is a common-cause failure of the motors. Load sharing and failure propagation are dependencies in which one failure causes or accelerates another, not failures from a shared cause. The distinction matters when the fault tree is built, because each dependency is represented differently, and when probabilities are estimated, because each draws on different data.

13.3.3 An explicit common-cause event model

Consider again the two pumps, now with their failure causes divided into two classes: pump-specific mechanisms, which act on one pump independently of the other, and one shared event that disables both. Define

$$\begin{aligned} I_A &= \{\text{independent failure of pump A}\},\\ I_B &= \{\text{independent failure of pump B}\},\\ C &= \{\text{a common cause disables both pumps}\}, \end{aligned}$$

with probabilities

$$\Pr(C) = c, \qquad \Pr(I_A) = \Pr(I_B) = q,$$

and assume that C, I_A and I_B are mutually independent. Pump A fails if its own mechanism acts or the common cause occurs, and similarly for pump B, so the total pump failure events are

$$A = I_A \cup C, \qquad B = I_B \cup C. \tag{13.12}$$

Figure 13.1(b) draws this structure: the common-cause block in series with the parallel pair of pump-specific failures.

The system fails when both pumps fail. Substituting equation (13.12) into equation (13.1) and applying the distributive law of Boolean algebra, $(X \cup C) \cap (Y \cup C) = (X \cap Y) \cup C$,

$$F_{\text{sys}} = (I_A \cup C) \cap (I_B \cup C) = C \cup (I_A \cap I_B). \tag{13.13}$$

There are two ways to lose both pumps: the common cause occurs, or both pumps fail from their independent mechanisms. The probability of the union follows from the addition rule, $\Pr(C \cup D) = \Pr(C) + \Pr(D) - \Pr(C \cap D)$ with $D = I_A \cap I_B$. By the assumed independence, $\Pr(D) = q^2$ and $\Pr(C \cap D) = cq^2$, so

$$Q_{\text{sys}} = c + q^2 - cq^2 = c + (1 - c)\, q^2. \tag{13.14}$$

The factor $(1 - c)$ removes the cases in which the common cause has already failed both pumps, which would otherwise be counted twice. The same steps give the marginal failure probability of one pump,

$$\Pr(A) = \Pr(I_A \cup C) = c + (1 - c)\, q. \tag{13.15}$$

13.3.4 Worked example of the common-cause model

Take $c = 0.01$ and $q = 0.05$ (illustrative). By equation (13.14),

$$Q_{\text{sys}} = 0.01 + 0.99 \times 0.05^2 = 0.01 + 0.002475 = 0.012475,$$

a system failure probability of 1.2475 %. Without the common-cause event the system failure probability is $q^2 = 0.0025$, so the shared-cause model predicts 4.99 times the value obtained from the pump-specific mechanisms alone, and the common cause contributes $0.01/0.012475 = 80\,\%$ of the total.

That comparison treats $q = 0.05$ as the probability of the pump-specific mechanisms only. The total probability that one pump fails is, by equation (13.15),

$$\Pr(A) = 0.01 + 0.99 \times 0.05 = 0.0595.$$

The point matters in practice. A pump failure probability estimated from field records counts every failure of a pump, whatever its cause, and so already includes the common-cause failures. Adding c on top of such a number counts the common-cause contribution twice; treating the total probability 0.0595 as if it were independent, $0.0595^2 = 0.00354$, omits the dependence and understates the system failure probability by a factor 3.5. A consistent model keeps the marginal fixed and divides it: if the field value is $\Pr(A) = 0.05$ and a common cause of probability $c = 0.01$ is identified within it, then $q = (0.05 - 0.01)/0.99 = 0.0404$ and $Q_{\text{sys}} = 0.01 + 0.99 \times 0.0404^2 = 0.0116$, 4.6 times the independent value 0.0025 at the same marginal. The partition of a total failure probability into an independent and a common-cause part is the subject of the beta-factor model in Chapter 14.

The common-cause model is one particular dependence structure of the kind treated in Section 13.2. Substituting equations (13.14) and (13.15) into equation (13.7) and expanding the square gives

$$\text{Cov}(X_A, X_B) = c + (1-c)q^2 - \big(c + (1-c)q\big)^2 = c(1-c)(1-q)^2,$$

which is positive whenever $0 < c < 1$. For $c = 0.01$ and $q = 0.05$ the covariance is 0.00893 and the correlation is $0.00893/(0.0595 \times 0.9405) = 0.160$, close to that of the conditional example of Section 13.2.2. The model replaces that number by a named event whose probability can be argued from its physical cause.

13.3.5 Common causes in the fault tree

Figure 13.4 draws the two models as fault trees, with the gate symbols of Chapter 10: the rectangle is the top event, the gate with a flat input side is an AND gate, the gate with a curved input side an OR gate, and the circles are basic events. Panel (a) is the tree of the total failure events, $T = A \cap B$, written $T = AB$ in the Boolean shorthand of fault-tree analysis. Panel (b) is the tree of equation (13.13): the top event, loss of pumping, occurs through an OR gate if the common cause C occurs or if the AND gate of the two independent failures is true, $T = C + I_A I_B$.

The two trees are logically equivalent, since $A = I_A + C$ and $B = I_B + C$ turn (a) into (b). They differ in what their basic events allow. In tree (a), the basic events A and B both contain C, and the standard quantification of an AND gate, the product of its input probabilities, is wrong for them. In tree (b), the common-cause mechanism appears explicitly as a basic event, every basic event is independent of the others, and the standard rules of Chapter 11 apply. The minimal cut sets of tree (b) are $\{C\}$, a single-event cut set, and $\{I_A, I_B\}$. Making the common cause explicit thus turns a redundant pair, which has no single-event cut set when its dependence is ignored, into a system with one: the shared cause is a single point of failure that the diagram of the total failure events concealed.

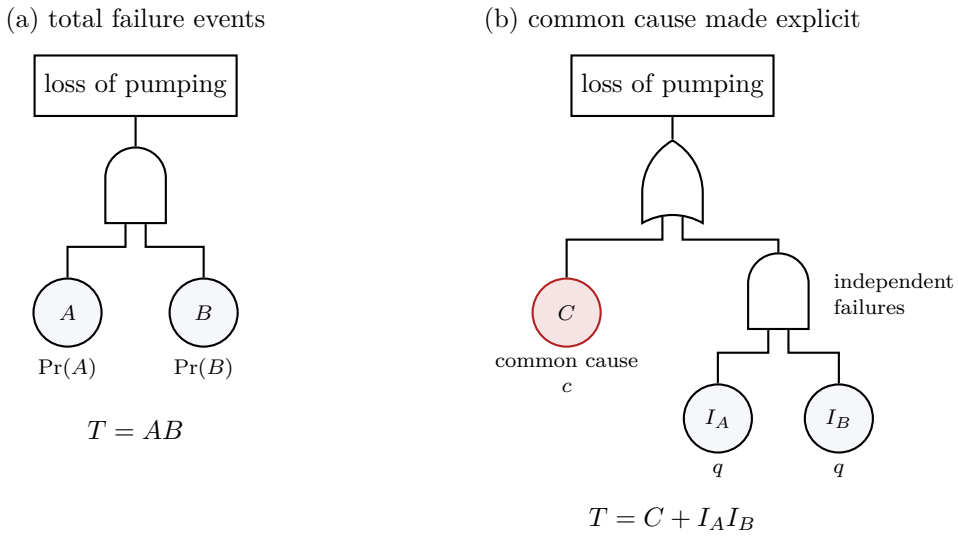


Figure 13.4: Fault trees for the loss of pumping. (a) On the total failure events A and B, which share the common cause and are therefore not independent. (b) With the common cause C as its own basic event; the basic events C, I_A and I_B are independent, so the top-event probability is $c + (1 - c)q^2$.

13.3.6 Redundancy saturation

Extend the model to n identical pumps, any one of which suffices, with one shared event C that disables all of them and independent pump-specific failures of probability q. The system fails if C occurs or if all n pump-specific failures occur, and the steps of equation (13.14) give

$$Q_n = c + (1 - c)\, q^n. \tag{13.16}$$

Since $0 \le q < 1$, the term q^n vanishes as n grows, and

$$\lim_{n \to \infty} Q_n = c.$$

Under the model, the system failure probability approaches the common-cause probability and never falls below it: a *common-cause floor*. The excess over the floor, $Q_n - c = (1 - c)q^n$, decreases geometrically by the factor q with each unit added.

Table 13.5 evaluates equation (13.16) for $c = 0.01$ and $q = 0.05$ and one to five pumps. The columns give the number of pumps, the independent contribution q^n, the system failure probability, and the share of it due to the common cause, c/Q_n. Figure 13.5 plots the same numbers on a logarithmic vertical axis against the number of pumps on the horizontal axis: the dots are Q_n, the dashed line the purely independent model q^n, and the dotted line the floor c. The second pump reduces the system failure probability from 5.95 % to 1.25 %, and the third to 1.012 %. Beyond that the benefit is negligible: the fourth and fifth pumps change the third significant figure only, while the independent model, which keeps falling by a factor of twenty per pump, predicts a failure probability of 3×10^{-7} for five pumps, more than four orders of magnitude below the floor. Adding identical units reduces the independent contribution but leaves the common-cause contribution untouched.

This is why diverse technology, physical separation, independent power supplies and independent controls can provide more reliability than another identical component. Each of these acts on c, which no number of identical units can reduce; an extra identical unit acts only on q^n, which is already small.

Table 13.5: Redundancy saturation: system failure probability of n redundant pumps, any one sufficient, with a common cause of probability $c = 0.01$ and pump-specific failure probability $q = 0.05$ (illustrative).

Number of pumps n	Independent part q^n	System failure probability Q_n	Common-cause share c/Q_n
1	0.05	0.059500 (5.9500 %)	16.8 %
2	0.0025	0.012475 (1.2475 %)	80.2 %
3	1.25×10^{-4}	0.010124 (1.0124 %)	98.8 %
4	6.25×10^{-6}	0.010006 (1.0006 %)	99.94 %
5	3.13×10^{-7}	0.010000 (1.0000 %)	99.997 %

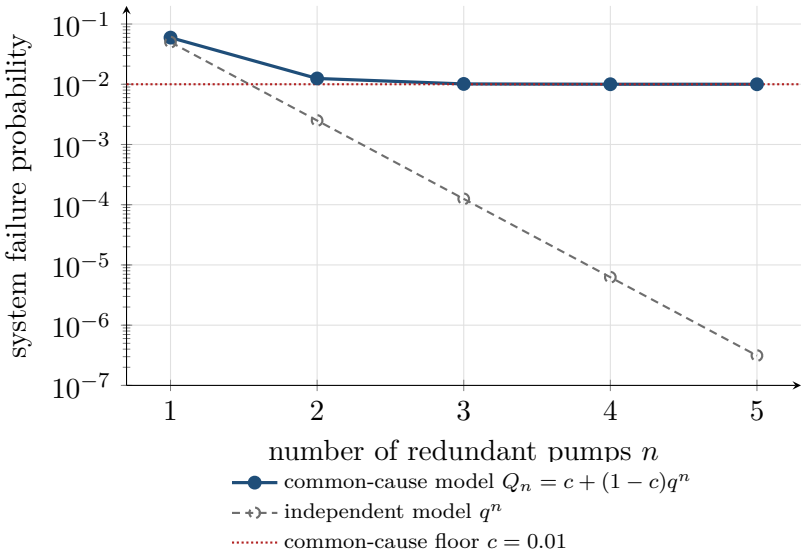


Figure 13.5: Redundancy saturation for $c = 0.01$ and $q = 0.05$ (illustrative). The common-cause model $Q_n = c + (1 - c)q^n$ levels off at the floor c from the third pump on, while the independent model q^n falls by a factor of twenty with each pump.

Principle 13.2. *Common-cause events defeat redundancy. With a shared cause of probability c, the failure probability of n redundant units any one of which suffices is $c + (1 - c)q^n$, which approaches c as n grows; past the first few units, reliability is improved by reducing c, not by adding identical units.*

13.4 Shared environmental conditions

A particularly important dependence arises when components respond to the same uncertain environment without any single event disabling them. This section shows that failures independent given the environment are dependent without it, computes the size of the effect, explains it with the law of total covariance, and applies it to transmission lines in a storm.

13.4.1 Conditional and unconditional independence

Two chillers of the data center plant are exposed to an ambient condition Z that is either normal or extreme,

$$Z \in \{\text{normal}, \text{extreme}\}, \qquad \Pr(Z = \text{extreme}) = 0.10, \qquad \Pr(Z = \text{normal}) = 0.90.$$

The failure probability of each chiller over the mission depends on the condition. Table 13.6 lists the illustrative values: each row is one condition, with its probability, the failure probability of one chiller in it, and the contributions of that condition to the marginal and to the joint failure probability computed in the next subsection. Let A and B be the failure events of chillers A and B, and assume that, within each condition, the two failures are independent:

$$\Pr(A \cap B \mid Z) = \Pr(A \mid Z)\ \Pr(B \mid Z). \tag{13.17}$$

This property is *conditional independence* given Z. It does not imply that A and B are independent once the condition is averaged out, because both conditional probabilities move with the same Z.

Table 13.6: Two chillers in a shared ambient condition (illustrative values) and the contribution of each condition to the marginal and the joint failure probabilities.

Condition z	$\Pr(Z = z)$	$\Pr(A \mid z) = \Pr(B \mid z)$	Contribution to $\Pr(A)$	Contribution to $\Pr(A \cap B)$
Normal	0.90	0.01	$0.90 \times 0.01 = 0.009$	$0.90 \times 0.01^2 = 0.00009$
Extreme	0.10	0.30	$0.10 \times 0.30 = 0.030$	$0.10 \times 0.30^2 = 0.009$
Total	1		$\Pr(A) = 0.039$	$\Pr(A \cap B) = 0.00909$

13.4.2 Marginal and joint failure probabilities

The marginal failure probability of either chiller follows from the law of total probability, which averages the conditional probability over the conditions,

$$\Pr(A) = \sum_{z} \Pr(A \mid Z = z)\ \Pr(Z = z) = 0.90 \times 0.01 + 0.10 \times 0.30 = 0.009 + 0.030 = 0.039,$$

and by symmetry $\Pr(B) = 0.039$. An analyst who took these marginal probabilities and assumed unconditional independence would compute

$$\Pr(A \cap B)_{\text{naive}} = 0.039^2 = 0.001521.$$

The correct joint probability applies the law of total probability to the intersection and then the conditional independence of equation (13.17):

$$\Pr(A \cap B) = \sum_{z} \Pr(A \cap B \mid Z = z)\, \Pr(Z = z) = \sum_{z} \Pr(A \mid z)\, \Pr(B \mid z)\, \Pr(z), \tag{13.18}$$

$$\Pr(A \cap B) = 0.90 \times 0.01^2 + 0.10 \times 0.30^2 = 0.00009 + 0.009 = 0.00909.$$

The shared-environment model gives a simultaneous failure probability of 0.909 %, against 0.1521 % under naive independence: 5.98 times larger, although the chillers were assumed independent in every condition. Almost all of the joint probability, 0.009 of 0.00909, comes from the extreme condition, which occurs one time in ten. The correlation of the failure indicators is, by equation (13.8), $(0.00909 - 0.001521)/(0.039 \times 0.961) = 0.202$.

13.4.3 The mechanism of environment-induced dependence

The dependence arises through the shared environmental variable. Extreme temperature makes both chillers more likely to fail, so the observation that one chiller has failed makes the extreme condition more plausible, and that raises the probability that the other fails. By Bayes' rule the probability of the extreme condition rises from 0.10 before the observation to

$$\Pr(Z = \text{extreme} \mid A) = \frac{0.10 \times 0.30}{0.039} = 0.769$$

after it, and the conditional failure probability of the second chiller becomes $\Pr(B \mid A) = 0.30 \times 0.769 + 0.01 \times 0.231 = 0.233$, equal to $\Pr(A \cap B)/\Pr(A) = 0.00909/0.039$, six times its marginal 0.039.

Figure 13.6 draws the structure. The box at the top is the ambient condition Z with its two states; the arrows are the influence of Z on the failure probability of each chiller; the two boxes below are the failure events. No arrow joins the two chillers: given Z, they are independent. The dashed line between them marks the dependence that appears when Z is not known, which the figure shows to be carried entirely by the common parent. A variable that drives several failure probabilities at once in this way is a *common stochastic driver*; weather, load and operating regime are the usual ones.

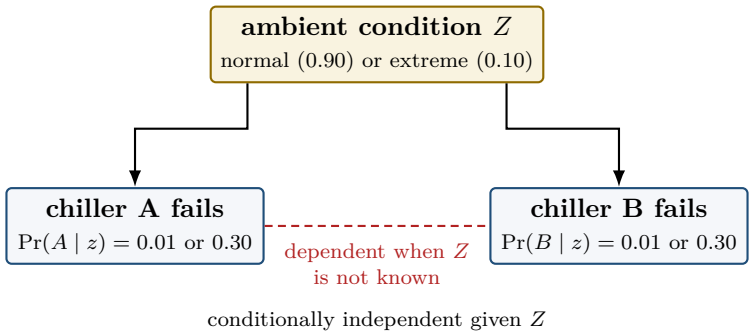


Figure 13.6: A shared environment as a common stochastic driver. The ambient condition Z influences the failure probability of both chillers; no direct link joins the chillers, which are independent given Z. The dashed line is the dependence induced when Z is averaged out (illustrative probabilities).

The *law of total covariance* states the mechanism in general. Its derivation takes two lines. By the tower property of expectation, $\mathbb{E}[X_A X_B] = \mathbb{E}\big[\mathbb{E}[X_A X_B \mid Z]\big]$, and $\mathbb{E}[X_A X_B \mid Z] = \operatorname{Cov}(X_A, X_B \mid Z) + \mathbb{E}[X_A \mid Z]\,\mathbb{E}[X_B \mid Z]$; likewise $\mathbb{E}[X_A] = \mathbb{E}\big[\mathbb{E}[X_A \mid Z]\big]$. Subtracting $\mathbb{E}[X_A]\mathbb{E}[X_B]$ from $\mathbb{E}[X_A X_B]$ gives

$$\operatorname{Cov}(X_A, X_B) = \mathbb{E}\big[\operatorname{Cov}(X_A, X_B \mid Z)\big] + \operatorname{Cov}\big(\mathbb{E}[X_A \mid Z],\, \mathbb{E}[X_B \mid Z]\big). \tag{13.19}$$

If the failures are conditionally independent, the first term is zero. The second term remains, and it is positive when the conditional failure probabilities $\mathbb{E}[X_A \mid Z] = \Pr(A \mid Z)$ and $\Pr(B \mid Z)$ move in the same direction with Z. For identical chillers both equal one function $p(Z)$, and the second term is the variance of $p(Z)$, never negative: $0.90 \times 0.10 \times (0.30 - 0.01)^2 = 0.007569$, which is exactly $0.00909 - 0.001521$. A shared environment that raises the failure probabilities of identical units together therefore always induces a positive dependence, and it induces none only if the failure probability does not vary with the environment. Components that respond in opposite directions, one more vulnerable in heat and the other in cold, would make the term negative. The decomposition is useful for any system exposed to common weather, load or operating conditions.

13.4.4 Transmission lines under severe weather

Two transmission lines of the running network case are exposed to a common storm, and the failure probability of each depends on the local wind. Let W be the storm intensity, a continuous random variable with probability density $p_W(w)$, and F_1, F_2 the failure events of the two lines. The law of total probability in its continuous form gives

$$\Pr(F_1 \cap F_2) = \int \Pr(F_1 \cap F_2 \mid W = w)\, p_W(w)\, \mathrm{d}w, \tag{13.20}$$

and if the lines are conditionally independent given the storm intensity,

$$\Pr(F_1 \cap F_2) = \int \Pr(F_1 \mid w)\ \Pr(F_2 \mid w)\, p_W(w)\, \mathrm{d}w. \tag{13.21}$$

Equation (13.18) is the same statement with two discrete conditions, and Chapter 2 used the same construction with two weather states for the current and voltage limits of one line.

Conditional independence given one variable may still be an incomplete model. A single storm intensity does not capture the spatial correlation of the wind along two routes, common structures such as a shared tower or right-of-way, the exposure of both routes to the same vegetation, protection actions that trip both lines, or the propagation of one outage to the other. A richer model conditions on a vector of shared variables, for example

$$\mathbf{Z} = (W,\ \text{wind direction},\ \text{soil condition},\ \text{vegetation exposure}),$$

and integrates over its joint density. Each failure probability given $\mathbf{Z}$ can then come from a structural model of the tower and conductor, a meteorological model of the storm and an electrical model of the network. This conditioning on shared physical variables is the bridge between classical reliability engineering, which treats failure probabilities as given numbers, and physics-based system risk analysis, which computes them from the state of the system and its environment.

> **Principle 13.3.** *Conditional independence does not imply unconditional independence. Failures that are independent given the weather, the load or the operating condition become dependent when that condition is uncertain, with a covariance equal to the covariance of their conditional failure probabilities.*

13.5 Failure propagation, load sharing and conditional hazards

The dependences so far came from shared causes. A different mechanism operates when the failure of one component changes the physical conditions experienced by another. This section introduces that dynamic dependence, derives the reliability of two load-sharing pumps, and generalizes the result to hazards that depend on the operating conditions and on the failure history.

13.5.1 Dependence through failure propagation

Several familiar events change the conditions of the surviving components. When one pump fails, the demand on its redundant partner rises. When a transmission line trips, the current redistributes onto the other lines. When a chiller fails, the load on the remaining chillers increases. When a battery module becomes unavailable, the current and the thermal duty of the remaining modules rise. In each case the first failure can alter the probability of the subsequent failures, through a changed load, a changed temperature or a changed current.

This is a *dynamic dependence*: the dependence acts through the order and the timing of the failures, not through a shared cause present from the start. A static joint probability at one instant does not describe it fully, because the probability that the second unit fails before the end of the mission depends on how long it has to run under the increased stress, and so on the time of the first failure. The model has to follow the system through its states as the failures occur.

13.5.2 Load sharing between two pumps

Two pumps operate in parallel, and either can maintain the required flow alone. While both operate, each has a constant hazard

$$\lambda_0 = 0.0001\,\mathrm{h}^{-1}.$$

When one fails, the survivor must carry the whole flow at a higher load, and its hazard rises to

$$\lambda_1 = 0.0005\,\mathrm{h}^{-1},$$

five times the shared-load value (illustrative). There is no repair during the mission. The quantity sought is the probability that the pumping system survives a mission of $H = 1000\,\mathrm{h}$ without losing both pumps.

Figure 13.7 draws the states of the system. Each box is a state, labeled by the number of surviving pumps: state 2, both pumps operating; state 1, one pump operating; state 0, both failed, which is system failure. The arrows are the transitions, labeled with their rates. From state 2 the system moves to state 1 when either pump fails; the two pumps each fail at rate λ_0, and the time to the first of two independent exponential failures is exponential with the sum of the rates, so the transition rate is $2\lambda_0$. From state 1 it moves to state 0 at the survivor's raised rate λ_1. With no repair there are no arrows back. Let $P_k(t)$ be the probability that the system is in state k at time t; initially both pumps operate,

$$P_2(0) = 1, \qquad P_1(0) = P_0(0) = 0.$$

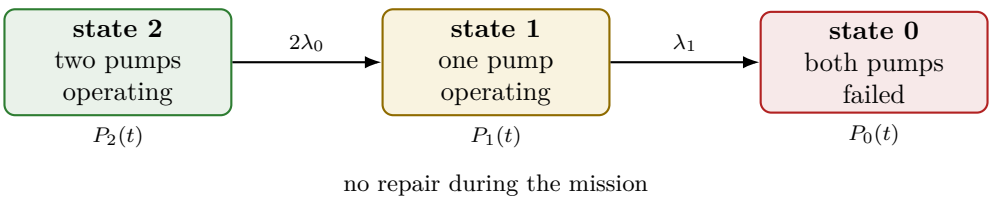


Figure 13.7: States of two load-sharing pumps, labeled by the number of surviving pumps. Both pumps operate at hazard λ_0 each, so the first failure occurs at rate $2\lambda_0$; the survivor then operates at the raised hazard λ_1. State 0 is system failure; there is no repair.

13.5.3 State probabilities

Both pumps surviving. The system remains in state 2 until the first failure, whose time is exponential with rate $2\lambda_0$, so

$$P_2(t) = \mathrm{e}^{-2\lambda_0 t}. \tag{13.22}$$

Exactly one pump surviving. The system is in state 1 at time t if the first failure occurred at some time u between 0 and t and the survivor has not failed in the remaining time $t-u$. The density of the first failure at time u is $2\lambda_0\,\mathrm{e}^{-2\lambda_0 u}$, and the survivor, now at hazard λ_1, survives the remaining interval with probability $\mathrm{e}^{-\lambda_1(t-u)}$. Summing over all times of the first failure,

$$P_1(t) = \int_0^t 2\lambda_0\,\mathrm{e}^{-2\lambda_0 u}\,\mathrm{e}^{-\lambda_1(t-u)}\,\mathrm{d}u. \tag{13.23}$$

Taking $\mathrm{e}^{-\lambda_1 t}$ out of the integral leaves $\int_0^t \mathrm{e}^{(\lambda_1-2\lambda_0)u}\,\mathrm{d}u = \big(\mathrm{e}^{(\lambda_1-2\lambda_0)t}-1\big)/(\lambda_1-2\lambda_0)$, and multiplying back gives, for $\lambda_1 \neq 2\lambda_0$,

$$P_1(t) = \frac{2\lambda_0}{\lambda_1 - 2\lambda_0}\big(\mathrm{e}^{-2\lambda_0 t} - \mathrm{e}^{-\lambda_1 t}\big). \tag{13.24}$$

If $\lambda_1 = 2\lambda_0$, the integrand of equation (13.23) is the constant $2\lambda_0\,\mathrm{e}^{-2\lambda_0 t}$, and

$$P_1(t) = 2\lambda_0 t\,\mathrm{e}^{-2\lambda_0 t}, \tag{13.25}$$

which is also the limit of equation (13.24) as $\lambda_1 \to 2\lambda_0$. Differentiating equations (13.22) and (13.24) confirms that they satisfy $\mathrm{d}P_2/\mathrm{d}t = -2\lambda_0 P_2$ and $\mathrm{d}P_1/\mathrm{d}t = 2\lambda_0 P_2 - \lambda_1 P_1$: the probability flows into state 1 at rate $2\lambda_0$ from state 2 and out of it at rate λ_1. These are the state equations of a Markov model, which a later chapter on Markov models derives in general.

13.5.4 System reliability and comparison with no load sharing

The system functions while at least one pump survives, so its reliability is

$$R_{\text{system}}(t) = P_2(t) + P_1(t) = \mathrm{e}^{-2\lambda_0 t} + \frac{2\lambda_0}{\lambda_1 - 2\lambda_0}\big(\mathrm{e}^{-2\lambda_0 t} - \mathrm{e}^{-\lambda_1 t}\big). \tag{13.26}$$

For $t = H = 1000\,\text{h}$, $2\lambda_0 t = 0.2$ and $\lambda_1 t = 0.5$, so

$$P_2 = \mathrm{e}^{-0.2} = 0.81873, \qquad P_1 = \frac{0.0002}{0.0003}\big(\mathrm{e}^{-0.2} - \mathrm{e}^{-0.5}\big) = \tfrac{2}{3}(0.81873 - 0.60653) = 0.14147,$$

$$R_{\text{system}}(1000\,\text{h}) = 0.81873 + 0.14147 = 0.96020.$$

The pumping system survives the mission with probability 96.02 %.

If the survivor kept its original hazard λ_0, the two pumps would be independent and identical, each with reliability $\mathrm{e}^{-\lambda_0 t}$, and the redundant pair would have

$$R_{\text{independent}}(t) = 1 - \big(1 - \mathrm{e}^{-\lambda_0 t}\big)^2, \tag{13.27}$$

which at 1000 h is $1-(1-\mathrm{e}^{-0.1})^2 = 1-0.095163^2 = 0.99094$. Setting $\lambda_1 = \lambda_0$ in equation (13.26) recovers this result, $\mathrm{e}^{-2\lambda_0 t} - 2(\mathrm{e}^{-2\lambda_0 t} - \mathrm{e}^{-\lambda_0 t}) = 2\mathrm{e}^{-\lambda_0 t} - \mathrm{e}^{-2\lambda_0 t}$, which is a check on the derivation. The mission failure probability is $1 - 0.96020 = 0.03980$ with load sharing, against $1 - 0.99094 = 0.00906$ without: the load-sharing dependence multiplies it by 4.4.

Figure 13.8 plots both reliability functions over 10 000 h. The horizontal axis is the mission time, the vertical axis the probability; the solid line is $R_{\text{system}}(t)$ with load sharing, the dashed line the independent pair, and the dotted line the probability $P_1(t)$ that the system is running on one pump. The marks are the values at 1000 h. The two reliability curves start together, since the first

failure occurs at the same rate in both models, and separate as the probability of running on one stressed pump grows; $P_1(t)$ peaks at 0.217 at $t = \ln(\lambda_1/2\lambda_0)/(\lambda_1 - 2\lambda_0) = \ln 2.5/0.0003 = 3054\,\mathrm{h}$ and then declines, because the lone pump fails quickly. Table 13.7 lists values at six mission times. Its columns are the mission time, the two state probabilities, the two system reliabilities and the ratio of the failure probabilities. The ratio is close to five for short missions, where it approaches λ_1/λ_0, because a short mission fails only if the second failure follows the first quickly, and that is exactly what the raised hazard makes λ_1/λ_0 times more likely; it falls towards one for long missions, in which both models eventually lose both pumps.

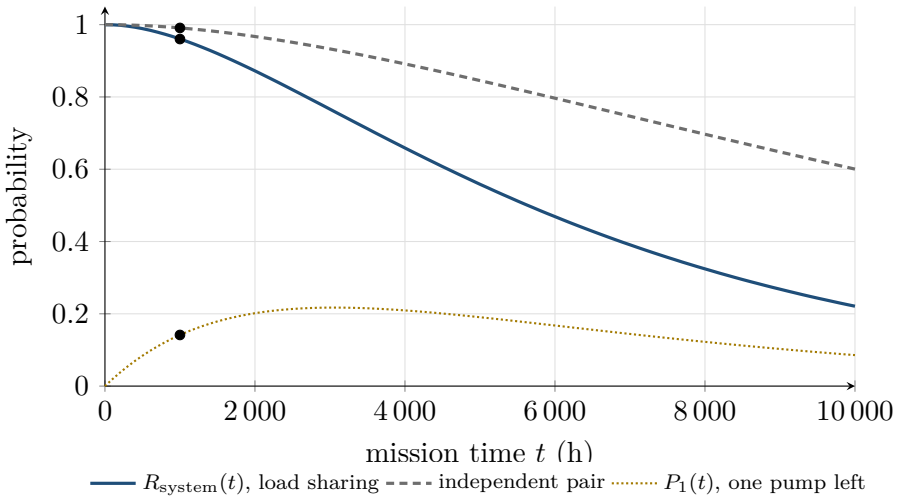


Figure 13.8: Reliability of two redundant pumps with load sharing, $R_{\mathrm{system}}(t) = \mathrm{e}^{-2\lambda_0 t} + [2\lambda_0/(\lambda_1 - 2\lambda_0)](\mathrm{e}^{-2\lambda_0 t} - \mathrm{e}^{-\lambda_1 t})$, and without, $1 - (1 - \mathrm{e}^{-\lambda_0 t})^2$, for $\lambda_0 = 0.0001\,\mathrm{h}^{-1}$ and $\lambda_1 = 0.0005\,\mathrm{h}^{-1}$ (illustrative), with the probability $P_1(t)$ of running on one pump. Marks at $t = 1000\,\mathrm{h}$.

Table 13.7: Two redundant pumps with load sharing ($\lambda_0 = 0.0001\,\mathrm{h}^{-1}$, $\lambda_1 = 0.0005\,\mathrm{h}^{-1}$) and without it (both at λ_0), for several mission times; no repair.

t (h)	$P_2(t)$	$P_1(t)$	$R_{\mathrm{system}}(t)$	$R_{\mathrm{independent}}(t)$	Failure-probability ratio
100	0.98020	0.01931	0.99951	0.99990	4.93
500	0.90484	0.08402	0.98886	0.99762	4.68
1000	0.81873	0.14147	0.96020	0.99094	4.40
2000	0.67032	0.20163	0.87195	0.96714	3.90
5000	0.36788	0.19053	0.55841	0.84518	2.85
10 000	0.13534	0.08573	0.22107	0.60042	1.95

The two pumps may be identical, and their original failure mechanisms independent; the dynamic behavior of the system still couples their failures. The coupling shows in the failure events themselves. By symmetry each pump is equally likely to be the survivor in state 1, so the probability that pump A has failed by 1000 h is $1-(P_2+P_1/2) = 1-(0.81873+0.07073) = 0.11054$, larger than its independent value 0.09516 because a pump that outlives its partner runs at λ_1. The product of the two marginals is $0.11054^2 = 0.01222$, against the joint failure probability 0.03980: the failure events are dependent, with a correlation of 0.28. Load sharing thus raises both the marginal and the joint failure probability, the joint much more.

A static reliability block diagram that assigns each pump one constant reliability misses the stress increase that follows the first failure. The system-level model must represent how load redistribution changes the future hazards. The situation is the same as in transmission

contingency analysis, where the loss of one component moves the remaining system into a different operating regime, with different flows and different margins.

13.5.5 Condition-dependent hazards

The load-sharing example used two constant hazards, λ_0 and λ_1. In real systems the hazard varies continuously with the operating conditions. A general condition-dependent formulation writes the hazard of component i as

$$\lambda_i\big(t \mid \mathcal{H}_t,\ \mathbf{z}(t)\big), \tag{13.28}$$

the instantaneous failure rate of component i at time t given $\mathcal{H}_t$, the history of the system up to time t, including the previous failures and their times, and $\mathbf{z}(t)$, the operating conditions such as current, temperature, pressure or mechanical stress. The variables in $\mathbf{z}(t)$ are called *covariates*; one standard form multiplies a baseline hazard by a function of them, the proportional-hazards model. The hazard of a transmission component, for example, may depend on

$$\lambda_i(t,\ T_i,\ I_i,\ W_i),$$

with T_i the component temperature, I_i its current and W_i the relevant weather variables; that of a pump on

$$\lambda_i(t,\ Q_i,\ \Delta p_i,\ T_i),$$

with Q_i its flow, Δp_i its pressure rise and T_i its operating temperature. The essential point is that these conditions can change when another component fails.

Failure history. The simplest history dependence is the load-sharing model itself. If pump A fails at time t_A, the hazard of pump B becomes

$$\lambda_B(t) = \begin{cases} \lambda_0, & t < t_A, \\ \lambda_1, & t \geq t_A, \end{cases} \tag{13.29}$$

and the probability that B fails before the end of the mission depends on when A failed: an early failure of A leaves B a long time at the raised hazard, a late one a short time. Given $t_A < H$, the probability that B survives to H is $\mathrm{e}^{-\lambda_0 t_A - \lambda_1 (H - t_A)}$, which runs from $\mathrm{e}^{-\lambda_1 H} = 0.607$ for $t_A = 0$ to $\mathrm{e}^{-\lambda_0 H} = 0.905$ for $t_A = H$ in the example. A single unconditional failure probability for B cannot describe this interaction; it is the average of these conditional values over the distribution of t_A.

13.5.6 Probabilistic physical coupling in a network

Consider three transmission lines between the same two areas. If line 1 trips, the power it carried redistributes over lines 2 and 3, in proportions fixed by the network equations of topology, injections and operating controls. Let $\mathbf{x}(t)$ be the operating state (the bus voltages and angles, from which the line currents follow), $\mathbf{s}(t)$ the topology (the status of every line), and $\mathbf{u}(t)$ the operating inputs (generation and demand). The electrical state is obtained from the network equations,

$$\mathbf{F}\big(\mathbf{x}(t),\ \mathbf{s}(t),\ \mathbf{u}(t)\big) = \mathbf{0}, \tag{13.30}$$

where the bold $\mathbf{F}$ is the vector of power-balance equations, not a cumulative failure probability. The hazard of line 2 may then depend on its resulting loading or thermal state,

$$\lambda_2(t) = h_2\big(\mathbf{x}(t),\ \mathbf{s}(t)\big), \tag{13.31}$$

with h_2 a function relating the loading of line 2 to its failure rate. When line 1 fails, the topology jumps from $\mathbf{s}^-$ to $\mathbf{s}^+$; the solution of equation (13.30) changes; and so may the hazard of line 2. With three identical lines sharing 240 MW, for instance, each carries 80 MW; after line 1 trips, lines 2 and 3 carry 120 MW each, a 50 % increase in loading (illustrative). The chain of effects is

1. line 1 outage;
2. change of the network topology, $\mathbf{s}^- \to \mathbf{s}^+$;
3. change of the power flows and line loadings, through equation (13.30);
4. change of the future failure hazards, through equation (13.31).

A reliability model of such a system may therefore need to re-solve the physical model each time the component states change, which is how cascading outages are analyzed.

Increased loading does not by itself imply an increased failure probability. A line loaded to 120 MW within its thermal rating may be no more likely to fail than at 80 MW; the hazard rises only if the increased stress drives a failure mechanism, such as conductor overheating and sag towards vegetation, or a protection relay set close to the new current. The function h_2 in equation (13.31) therefore requires a credible physical or empirical model connecting the relevant stress to the failure mechanism, in the sense of Chapter 4.

> **Principle 13.4.** *Failures can change the probability of subsequent failures. Load redistribution and physical interaction make the hazard of a component depend on the history of the system and on its operating state, and the reliability model must then follow the system through its states instead of assigning each component one fixed probability.*

13.6 Dependent failures in engineering practice and quantitative risk

This section turns from the mathematics to practice: the typical common causes, the defenses against them, and the ways a fault tree can carry a dependence. It then shows how strongly a small dependence assumption changes the computed risk of a highly redundant system, how uncertain that assumption is, why the data are scarce, and how the results bear on a design decision for the chiller plant.

13.6.1 Typical common causes

Common-cause failures arise from three broad categories of shared conditions. *Environmental exposure*: flooding, lightning, extreme temperature, wind, contamination and other environmental conditions can affect several components together. *Shared controls and infrastructure*: a common controller, communication network, power supply or protective device creates one failure pathway through several nominally redundant units. *Human and organizational coupling*: shared maintenance procedures, incorrect test configurations, common calibration errors or defective replacement parts from one batch produce related failures in several units.

A common cause need not produce failures at the same instant. What counts is whether the affected components become unavailable within the time window that matters for the system's mission. One maintenance error, for example, may leave latent defects in both redundant pumps, which are discovered only when each pump is demanded, possibly hours or days apart; the resulting failures are still a common-cause concern, because both defects were present before the demand. The classification of an event as a common-cause failure accordingly rests on its root cause, its coupling factor, the defenses that were in place and failed, and the timing of the failures.

13.6.2 Defenses against dependent failures

Different dependence mechanisms require different engineering responses. Table 13.8 pairs seven dependencies with possible defenses: each row names a dependency and the measures that weaken it. Independent feeders and separated supply paths remove a shared electrical supply; redundant controllers with independent control functions remove a common controller; physical separation, shielding and diverse locations reduce shared environmental exposure; design diversity, independent qualification and alternate suppliers reduce manufacturing defects; independent verification, staggered maintenance and written procedures reduce maintenance errors; capacity margin, load management and protective controls limit load-sharing stress; and measurement diversity, independent calibration and plausibility checks reduce a shared sensor bias.

Table 13.8: Dependencies and possible defenses against them.

Dependency	Possible defense
Shared electrical supply	Independent feeders, properly separated supply paths
Common controller	Redundant controllers with appropriately independent control functions
Shared environmental exposure	Physical separation, shielding, diverse locations
Manufacturing defect	Design diversity, independent qualification, alternate suppliers
Maintenance error	Independent verification, staggered maintenance, procedures
Load-sharing stress	Capacity margin, load management, protective controls
Shared sensor bias	Measurement diversity, independent calibration, plausibility checks

The goal of these defenses is not to add components but to reduce the mechanisms that make redundant components fail together; in the model of Section 13.3.3, to reduce c. Diversity helps, but diverse components can still share vulnerable power, software, communication or physical infrastructure: two pumps of different makes in the same flooded basement fail together. Every proposed defense should therefore be evaluated against the specific failure mechanisms and coupling factors it is meant to break.

13.6.3 Dependence in the fault tree

The ordinary fault tree of two redundant pumps has the top event $T = AB$. When A and B are the total failure events, this Boolean expression remains correct under dependence, since it states only that the system fails when both pumps fail. Its quantification as $\Pr(T) = \Pr(A)\,\Pr(B)$ does not, and four approaches replace it.

- **Joint probability.** Estimate $\Pr(A \cap B)$ directly, from field data on joint failures, from a conditional probability $\Pr(B \mid A)$, or from a joint statistical model, and assign it to the AND gate.
- **Explicit shared-cause events.** Decompose the total failures into common-cause and component-specific events, $T = C + I_A I_B$, as in Figure 13.4(b), and assign probabilities to the underlying events with justified independence or conditional-dependence assumptions.
- **Conditional environmental models.** Introduce the shared variables $\mathbf{Z}$ and compute

$$\Pr(A \cap B) = \int \Pr(A \cap B \mid \mathbf{Z} = \mathbf{z})\, p(\mathbf{z})\, \mathrm{d}\mathbf{z},$$

with $p(\mathbf{z})$ the joint density of the shared variables, as in Section 13.4.

- **Dynamic state models.** When one failure changes the subsequent hazards, represent the state transitions and their timing explicitly, as in Section 13.5.2.

The four approaches are not interchangeable. An explicit common-cause event is useful when the shared initiating mechanism can be identified and quantified. A conditional model is useful when the environmental or operating drivers are observable, since the dependence can then be computed from measured conditions instead of being estimated as a parameter. A dynamic model is required when the order of the events and the redistribution of load matter. When the common cause cannot be named but its effect is known from data, a parametric common-cause model assigns a share of each component's failure probability to the common cause; the beta-factor and alpha-factor models of Chapter 14 are of this kind.

13.6.4 Small dependence assumptions and large risk differences

A system has three redundant components, any one of which is sufficient, each with an independent failure probability $q = 0.01$ (illustrative). With no shared causes,

$$Q_{\text{sys}} = q^3 = 10^{-6},$$

one failure per million missions in probability terms. Introduce a common-cause event of probability $c = 10^{-4}$, itself a rare event, one-hundredth of the component failure probability. By equation (13.16),

$$Q_{\text{sys}} = c + (1 - c)\, q^3 = 0.0001 + 0.9999 \times 0.000001 = 1.01 \times 10^{-4},$$

101 times the independent value; the common cause contributes 99 % of it. The more effective independent redundancy becomes, the smaller q^n is, and the more the residual risk is dominated by even rare shared causes. Assuming independence is therefore most dangerous precisely in the most highly redundant systems.

13.6.5 Uncertainty about dependence

Suppose the existence of a common cause is accepted but its probability c is not known. Instead of fixing one value, c is treated as an uncertain parameter, an epistemic uncertainty in the vocabulary of the notation chapter. For the model $Q_n(c, q) = c + (1 - c)q^n$ the sensitivities to the two parameters are the partial derivatives

$$\frac{\partial Q_n}{\partial c} = 1 - q^n, \qquad \frac{\partial Q_n}{\partial q} = n(1 - c)\, q^{n-1}. \tag{13.32}$$

If q^n is very small, $\partial Q_n/\partial c \approx 1$: the system failure probability moves almost one for one with c. The sensitivity to q contains q^{n-1} and becomes tiny when q is small and n large. For $c = 10^{-4}$, $q = 0.01$ and $n = 3$, $\partial Q_n/\partial c = 1 - 10^{-6} = 0.999999$ and $\partial Q_n/\partial q = 3 \times 0.9999 \times 10^{-4} = 3.0 \times 10^{-4}$. In relative terms, a 1 % change in c changes Q_n by 0.99 %, and a 1 % change in q changes it by 0.03 %. Halving q to 0.005 lowers Q_n from 1.0100×10^{-4} to 1.0012×10^{-4}, less than 1 %. In a highly redundant system, further refinement of the component failure probabilities therefore has little value while the common-cause probability remains poorly known.

Table 13.9 shows the consequence for an illustrative range $c \in [10^{-4}, 10^{-3}]$ with $q = 0.01$ and $n = 3$. Each row is one value of c; the columns give the system failure probability, its ratio to the independent value 10^{-6}, and the common-cause share. At the two ends of the range,

$$Q_{\text{sys}}(0.0001) = 1.010 \times 10^{-4}, \qquad Q_{\text{sys}}(0.001) = 1.001 \times 10^{-3},$$

a factor of 9.9, nearly an order of magnitude, from the dependence assumption alone. That uncertainty can change an engineering decision. The interval itself is hypothetical: a defensible analysis needs evidence for the distribution or the bounds assigned to c, and the estimation of c from limited data, and of its uncertainty, is the subject of Chapter 14 and, with prior distributions, of the methods of Chapter 8.

Table 13.9: System failure probability of three redundant components ($q = 0.01$) for several values of the common-cause probability c (illustrative).

c	$Q_{\text{sys}} = c + (1-c)q^3$	Ratio to $q^3 = 10^{-6}$	Common-cause share
0	1.000×10^{-6}	1	0
1×10^{-4}	1.010×10^{-4}	101	99.0 %
3×10^{-4}	3.010×10^{-4}	301	99.7 %
1×10^{-3}	1.001×10^{-3}	1001	99.9 %

13.6.6 Common-cause data

Estimating independent component failure probabilities is already difficult, as Chapter 7 showed. Estimating common-cause probabilities is harder, because simultaneous or causally related failures are rare, so that even a large fleet records few of them. A failure database that supports common-cause estimation must preserve, for each event, which components were affected; when the failures occurred; whether the components were exposed to the same initiating condition; whether the failure mechanisms were related; what coupling factors existed; which defenses succeeded or failed; and whether the failures were latent or occurred on demand.

The classification depends on the mechanism and the mission context, not on the calendar. Recording that two pumps failed in the same year does not establish a common-cause event, since two independent failures can fall in one year; and observing failures at different times does not rule one out, since a shared latent defect can be revealed at two separate demands. Common-cause data programs for probabilistic risk assessment collect and classify exactly these features of each event, so that the events can be counted consistently for the estimation methods of the next chapter.

13.6.7 A fourth chiller or a second water supply

The data center plant has three chillers, each rated 500 kW, and the facility requires at least 900 kW of cooling, so two chillers are needed ($2 \times 500 = 1000$ kW) and one is not enough: the plant is a 2-out-of-3 system. All three chillers draw on the same cooling-water source. The operator proposes adding a fourth identical chiller. Whether this improves reliability substantially is settled by four observations. The fourth chiller does increase the number of independent chiller failures that can be tolerated, from one to two, provided its capacity can be used. It does not remove the shared cooling-water dependency. A single failure of the common source can disable all four chillers, since every chiller requires it. A second cooling-water supply, or a cooling technology independent of the source, may therefore give a larger improvement.

The right evaluation compares the top-event probabilities of the alternative designs, shared-cause vulnerabilities included. Table 13.10 does so with illustrative numbers: each chiller fails independently during the mission with probability $q = 0.05$, the cooling-water source fails with probability 0.01, and a second source, if installed, is independent of the first and can supply every chiller alone, so that the supply fails with probability $0.01^2 = 10^{-4}$. The chiller group of the 2-out-of-3 plant fails when at least two of three chillers fail, with probability $3q^2(1-q)+q^3 = 0.007125+0.000125 = 0.00725$; with four chillers, of which two are still required,

it fails when at least three fail, with probability $4q^3(1-q) + q^4 = 0.000475 + 0.00000625 = 0.000481$. With the supply failure as the common cause of probability c_{w}, the top-event probability is $c_{\mathrm{w}} + (1 - c_{\mathrm{w}})\, Q_{\mathrm{chillers}}$, as in equation (13.14). For the present plant this is $0.01 + 0.99 \times 0.00725 = 0.01718$. The rows of the table are the four designs; the columns give the supply failure probability, the chiller group failure probability and the top-event probability.

Table 13.10: Top-event probability of four chiller-plant designs: chillers of 500 kW, 900 kW required, chiller failure probability 0.05 each, cooling-water source failure probability 0.01 each, all independent (illustrative).

Design	Supply failure c_{w}	Chiller group failure	Top event Q_{sys}
Three chillers, one source (present)	0.01	0.00725	0.01718
Four chillers, one source	0.01	0.000481	0.01048
Three chillers, two sources	0.0001	0.00725	0.00735
Four chillers, two sources	0.0001	0.000481	0.00058

The fourth chiller lowers the top-event probability from 0.0172 to 0.0105, which is close to the floor 0.01 set by the single source; the second source lowers it to 0.0073, by more than the fourth chiller does; the two together lower it to 0.0006, thirty times below the present plant, because each removes the contribution the other leaves. With these numbers the second supply is the better single investment, and the fourth chiller pays off only after the shared supply has been addressed. The ranking depends on the assumed numbers, and on the assumption that the two sources are truly independent: a second source drawn from the same intake would share its causes and give much less. Counting redundant chillers answers none of these questions; comparing top-event probabilities with the shared causes represented answers all of them. The next chapter, Chapter 14, gives the standard parametric models with which the probability of such shared causes is estimated from limited data.

13.7 What to retain

Definition 13.1 (Dependent failures)**.** The failure probability of a redundant system depends on both the behavior of its components and the relations among them: it is determined by the joint distribution of the component failure events, which shared causes, shared environments, load sharing and failure propagation make different from the product of the marginal probabilities. With a common cause of probability c, the failure probability of n redundant units, any one sufficient, is $c + (1 - c)q^n$, which never falls below c.

Table 13.11 lists the terms of the chapter, with their meaning, their units and the subsection that defines each.

Table 13.11: Terms of the chapter and where each is defined.

Term	Meaning	Units	Defined in
Statistical independence	$\Pr(A \cap B) = \Pr(A)\Pr(B)$, equivalently $\Pr(B \mid A) = \Pr(B)$; a property of the joint distribution	–	13.2.1
Conditional probability	$\Pr(B \mid A) = \Pr(A \cap B)/\Pr(A)$; multiplication rule $\Pr(A \cap B) = \Pr(A)\Pr(B \mid A)$	–	13.2.1
Joint failure probability	probability that both units fail; fails a redundant pair	–	13.2.2
Fréchet bounds	$\max(0, p_A + p_B - 1) \le \Pr(A \cap B) \le \min(p_A, p_B)$	–	13.2.3
Phi coefficient ρ_{AB}	correlation of the failure indicators, eq. (13.8); feasible range set by the marginals	–	13.2.4
Dependent failures	failure events whose joint probability is not the product of their probabilities	–	13.3.1
Common-cause failure (CCF)	failure of several items from a shared root cause through a coupling factor, within a relevant time window	–	13.3.2
Coupling factor	property that makes several items susceptible to the same root cause	–	13.3.2
Common-cause component group	items sharing a coupling factor, analyzed together	–	13.3.2
Explicit common-cause model	$A = I_A \cup C$, $B = I_B \cup C$; $Q_{\text{sys}} = c + (1-c)q^2$	–	13.3.3
Common-cause floor	$\lim_{n\to\infty} Q_n = c$; redundancy saturation	–	13.3.6
Conditional independence	$\Pr(A \cap B \mid Z) = \Pr(A \mid Z)\Pr(B \mid Z)$; does not imply independence	–	13.4.1
Common stochastic driver	shared variable (weather, load) that moves several failure probabilities together	–	13.4.3
Law of total covariance	$\mathrm{Cov} = \mathbb{E}[\mathrm{Cov}(\cdot \mid Z)] + \mathrm{Cov}(\mathbb{E}[\cdot \mid Z], \mathbb{E}[\cdot \mid Z])$	–	13.4.3
Failure propagation (cascading failure)	one failure causes or accelerates another; a dynamic dependence	–	13.5.1
Load sharing	survivors take over the load and their hazard rises, $\lambda_0 \to \lambda_1$	h^{-1}	13.5.2
State probabilities $P_k(t)$	probability that k units survive at time t	–	13.5.3
Condition-dependent hazard	$\lambda_i(t \mid \mathcal{H}_t, \mathbf{z}(t))$, hazard given history and covariates	h^{-1}	13.5.5
Defenses	measures that break coupling factors: separation, diversity, independent supplies, staggered maintenance	–	13.6.2

Further reading. Rausand, Barros and Høyland, *System Reliability Theory*, third edition, for dependent failures, common causes and system reliability. Kumamoto and Henley, *Probabilistic Risk Assessment and Management for Engineers and Scientists*, second edition, for dependent events and PRA. U.S. Nuclear Regulatory Commission, *Common-Cause Failure Database and NUREG/CR-4780 background*, for the definitions and the development of common-cause failure methods. *NUREG/CR-6268, Common-Cause Failure Database and Analysis System*, for the classification, the event data and the quantification methodology. NASA, *Common Cause Failures Dominate and Defeat Redundancy*, for why common-cause failure can dominate the risk of redundant systems.

13.8 Exercises

Exercise 13.1. Answer each part in one or two sentences, referring to the definitions of the chapter. (a) Two pumps each have a 5 % marginal failure probability. Is their joint failure probability necessarily 0.25 %? (No, only if their failures are independent: the product rule requires independence, and shared causes or other dependencies change the joint probability.) (b) Two chillers are independent conditional on the ambient temperature. Are they necessarily independent without conditioning? (No: a shared random environment can induce marginal dependence even when conditional independence holds.) (c) What is the main limitation of adding many redundant components when all share one common-cause event? (The common-cause contribution creates a residual failure floor: $Q_n = c + (1 - c)q^n$ approaches c as n grows.) (d) A pump fails and its redundant partner must operate under higher stress. What kind of dependency is this? (A load-sharing dependency: the first failure changes the remaining component's operating conditions and its subsequent hazard.) (e) Which approach is most appropriate when the order of component failures changes the subsequent failure rates? (A dynamic state-transition model, which can represent changing hazards, event histories and transitions between system configurations.)

Exercise 13.2. (a) Two redundant pumps have marginal failure probabilities $p_A = 0.04$ and $p_B = 0.06$ over a mission, and $\Pr(B \mid A) = 0.25$. Compute the joint failure probability and its ratio to the value under independence, the conditional probability $\Pr(A \mid B)$, and the probability $\Pr(A \cup B)$ that at least one pump fails. Check the joint probability against the Fréchet bounds, compute the phi coefficient, and compute the largest and smallest correlations these marginals allow; explain why the largest is below one. (b) Two transmission lines are exposed to a storm that is calm with probability 0.85, moderate with probability 0.12 and severe with probability 0.03; in these three conditions each line fails during the event with probability 0.002, 0.03 and 0.25, independently given the condition (illustrative). Compute the marginal failure probability of one line, the joint failure probability under naive independence and under the shared-storm model, their ratio, the conditional probability $\Pr(F_2 \mid F_1)$, the posterior probability of the severe condition given that line 1 has failed, and the correlation of the two failure indicators. Verify the covariance with the law of total covariance, equation (13.19).

Exercise 13.3. A group of identical pumps, any one sufficient, has pump-specific failure probability $q = 0.04$ and a common cause of probability $c = 0.002$ (illustrative). (a) Compute the total failure probability of one pump. (b) Compute Q_n for $n = 1$, 2, 3 and 4, and the common-cause share of each. (c) An analyst treats the total pump failure events as independent with the probability of (a). Compute the analyst's system failure probability for $n = 2$ and $n = 3$ and the factor by which it understates the correct value. (d) Another analyst takes the probability of (a), which came from field records of all pump failures, as q and adds the common cause on top. Compute that analyst's Q_2 and explain the error. (e) Compute the correlation of two pump failure indicators from $\mathrm{Cov}(X_A, X_B) = c(1 - c)(1 - q)^2$. (f) For $n = 3$, evaluate the two sensitivities of equation (13.32) and state which parameter is worth estimating more precisely.

Exercise 13.4. Two load-sharing pumps, either sufficient, each have hazard $\lambda_0 = 0.0002\,\mathrm{h}^{-1}$ while both run; the survivor's hazard rises to $\lambda_1 = 0.0004\,\mathrm{h}^{-1}$; there is no repair. (a) Since $\lambda_1 = 2\lambda_0$, use equation (13.25); derive it from equation (13.23) and show that the system reliability is $\mathrm{e}^{-x}(1 + x)$ with $x = 2\lambda_0 t$. (b) Compute P_2, P_1 and R_{system} at 1000 h, and compare the mission failure probability with that of two independent pumps at λ_0. (c) Repeat (b) with $\lambda_1 = 0.001\,\mathrm{h}^{-1}$. (d) For the case of (b), compute the probability that pump A has failed by 1000 h, compare the product of the two marginal failure probabilities with the joint failure probability, and compute the correlation of the failure indicators.

Chapter 14

Beta-Factor and Alpha-Factor Models

14.1 Common-cause failure in redundant groups and its parameter models

Chapter 13 showed that the independence assumption can underestimate the failure probability of a redundant system by orders of magnitude, because one shared cause can disable several of its components at once. A probabilistic risk assessment therefore needs parameters that state how much of the failure behavior of a group of similar components comes from shared causes, and how many components a shared cause disables. The question of this chapter is

how the failure probability of a redundant group is computed when one cause can disable several of its components, which parameters summarize that dependence, and what information a simplified parameter model discards.

The two classical answers are the beta-factor model, which splits the failure rate of each component into an independent part and a part due to events that disable the whole group, and the alpha-factor model, which also represents events that disable only some members of the group. The chapter first separates the rates and the probabilities that the two models are built from, then derives each model, works examples with two, three and four components, compares the two models, and ends with the estimation of their parameters from failure records.

Jargon: common-cause failure, common-cause component group, coupling factor, root cause, multiplicity, beta factor, alpha factor, basic event, staggered testing

A *common-cause failure* (CCF) is the failure of two or more components from one shared cause within a short interval, which defeats the independence that redundancy assumes. A *common-cause component group* is the set of components, usually identical in design, function, environment and maintenance, that are judged susceptible to the same shared causes and are analyzed together. The *root cause* is the condition that starts a failure (a contaminated lubricant, a calibration error); the *coupling factor*, or coupling mechanism, is the shared property that lets one root cause reach several components at once (the same lubricant batch, the same procedure, the same power supply). The *multiplicity* of a failure event is the number of components of the group it disables. The *beta factor* β is the fraction of a component's failure rate that is attributed to events disabling the whole group. The *alpha factor* α_k is the fraction of failure events of the group that disable exactly k of its components. A *basic event* is the lowest-level event of a fault tree to which a probability is assigned; in CCF modeling, the failure of one specified subset of the group from a common cause is a basic event. *Staggered testing* tests the components of a standby group at evenly spaced times; *non-staggered testing* tests them together.

Key idea

A beta factor partitions the failure rate of one component; an alpha factor partitions the failure events of a group. Neither is a probability of system failure: that probability follows only after the parameters are converted into event rates or basic-event probabilities and combined with the system's success logic over a stated mission.

This section states the effect of a shared cause on a redundant group and places the two parameter models among the models of dependence.

14.1.1 Redundancy under the independence assumption

Consider a group of three identical components, any one of which can perform the required function, and let q be the probability that one component fails during a mission. If the three failures are independent, the group fails only when all three fail, with probability

$$Q_{\text{independent}} = q^3, \tag{14.1}$$

and for $q = 0.01$ the result is 10^{-6}. Four such components give $(0.01)^4 = 10^{-8}$. Each added component multiplies the group failure probability by a further factor q, so that under independence redundancy appears to drive the failure probability toward zero.

The calculation ignores the possibility that the components fail together. Redundant pumps or chillers usually share a maintenance procedure, identical parts from the same manufacturing lot, a common electrical supply or control system, and the same environment. Each of these is a coupling factor, and a single root cause that reaches the group through it can disable two, three or all of the components in one event. A single event of that kind, with a probability of the order of 10^{-3} over the mission, dominates a result of 10^{-6} or 10^{-8}, whatever the number of components.

14.1.2 Common-cause parameter models

Chapter 13 modeled a shared cause explicitly: a common event of probability c that disables all n components, and independent component-specific failures of probability q, gave the group failure probability

$$Q_{\text{sys}} = c + (1 - c)\, q^n. \tag{14.2}$$

That model leaves open where the values of c and q come from, and whether a shared event always disables all n components. Common-cause parameter models answer both points. They express c and q, or their generalizations, as fractions of a component's total failure rate or failure probability, which is estimated from ordinary component records, and they estimate the fractions themselves from records of events that affected several components. The beta-factor model uses one fraction; the alpha-factor model uses one fraction for each possible multiplicity $k = 1, \ldots, m$ in a group of m components. Both are standard in probabilistic risk assessment practice.

The book's distinction between the physical system, the model and the evidence applies directly. The physical system has root causes and coupling factors: a lubricant batch, a shared cooling-water header, a maintenance crew. The model replaces them by a few parameters and an assumed event process. The evidence is a set of failure records in which each event has been classified by the components it affected. A beta or alpha factor is a summary of that evidence under that model; it does not identify the physical mechanism, and a change in the design changes the mechanism, not the factor.

14.2 Failure rates and failure probabilities

Common-cause parameters are defined either on rates or on probabilities, and errors arise when one is substituted for the other. This section defines the rates, the probabilities per demand and the mission failure probabilities used in the rest of the chapter, and the condition under which a probability can be approximated by a rate times a duration.

14.2.1 Rates: the component failure rate and the group-event frequency

Table 14.1 lists the quantities of the chapter. The first two are rates. The *total component failure rate* λ_{T} is the constant hazard of one component of the group, counting every event that fails it, whether the event affects that component alone or several components together. It is a rate in h^{-1}, and under the exponential model of Chapter 6 it is estimated as the number of failures of the component type divided by the accumulated operating time of all components of that type. The *group-event frequency* r_k is the rate, in h^{-1}, of events that disable exactly k components of the group, whichever k components they are; it is estimated as the number of such events divided by the operating time of the group. The two rates count different things: λ_{T} counts component failures seen from one component, r_k counts events seen from the group. One event of multiplicity four adds one to the count behind r_4 and one failure to each of the four components.

14.2.2 Probabilities per demand and snapshot unavailability

Standby equipment, such as an emergency cooling pump or a backup generator, is idle most of the time and must start and run when a demand arrives. Its reliability is stated as a *failure probability per demand* q_{d}: the probability that it fails to start or to run when called on. It is dimensionless and is estimated as the number of failures on demand divided by the number of demands, tests included. It is a probability attached to an event, not a rate attached to a duration.

A standby component can also fail while idle and remain failed, without anyone knowing, until it is tested or demanded. Its *snapshot unavailability* $q_{\mathrm{U}}(t)$ is the probability that it is in a failed state at the instant t. If it fails in standby at a constant rate λ and the failure is revealed only by a test performed every τ hours, the unavailability u hours after a test is $1 - \mathrm{e}^{-\lambda u}$, and

Table 14.1: Rates and probabilities used in common-cause modeling. The first two are rates; the others are dimensionless probabilities.

Quantity	Symbol	Interpretation	Units
Total component failure rate	λ_{T}	constant hazard of one component, from all events that fail it	h^{-1}
Group-event frequency	r_k	rate of events disabling exactly k components of the group	h^{-1}
Failure probability per demand	q_{d}	probability that a standby component fails to perform when demanded	–
Snapshot unavailability	$q_{\mathrm{U}}(t)$	probability that a component is in a failed state at the instant t	–
Mission failure probability	$Q(t)$	probability that the defined failure occurs by time t	–

its average over the test interval is the standard result

$$\bar{q}_{\mathrm{U}} = \frac{1}{\tau}\int_0^{\tau}\left(1-\mathrm{e}^{-\lambda u}\right)\mathrm{d}u = 1-\frac{1-\mathrm{e}^{-\lambda\tau}}{\lambda\tau} \approx \frac{\lambda\tau}{2} \quad (\lambda\tau \ll 1). \tag{14.3}$$

For $\lambda = 10^{-4}\,\mathrm{h}^{-1}$ and a monthly test, $\tau = 720\,\mathrm{h}$, the exact average is $1 - 0.069469/0.072 = 0.03515$ and the approximation gives 0.036 (illustrative values). The test interval enters the result directly, which is why the testing arrangement of a group matters to its common-cause contribution (Section 14.6.4).

14.2.3 Mission failure probability and the rare-event approximation

For an item with constant hazard λ, the probability that it fails by time t is the cumulative failure probability of the exponential model,

$$Q(t) = 1 - \mathrm{e}^{-\lambda t}. \tag{14.4}$$

Expanding the exponential, $Q(t) = \lambda t - (\lambda t)^2/2 + \cdots$, so that

$$Q(t) \approx \lambda t \quad \text{when } \lambda t \ll 1, \tag{14.5}$$

the rare-event approximation. It always overestimates, and the relative error is about $\lambda t/2$. Table 14.2 quantifies it: at $\lambda t = 0.01$ the approximation is 0.5 % high, at $\lambda t = 0.1$ it is 5 % high, and at $\lambda t = 1$ it gives a probability of 1 where the exact value is 0.632.

Table 14.2: Exact mission failure probability $1 - \mathrm{e}^{-\lambda t}$ against the rare-event approximation λt (equations (14.4) and (14.5)).

λt	$1 - \mathrm{e}^{-\lambda t}$	relative error of λt
0.001	0.000999500	0.05 %
0.01	0.00995017	0.50 %
0.05	0.0487706	2.5 %
0.1	0.0951626	5.1 %
0.2	0.181269	10.3 %
0.5	0.393469	27.1 %
1	0.632121	58.2 %

The distinction matters for common-cause analysis because the parameters are estimated from counts of events over exposure, which are rates or per-demand frequencies, while engineering decisions need finite-time system failure probabilities or snapshot unavailabilities. A relation that is exact between rates, such as a split of a component failure rate into two parts, becomes only approximate between the corresponding mission probabilities. Section 14.3.5 works this out for the beta factor.

14.3 The beta-factor model

The beta-factor model is the simplest common-cause parameter model. This section defines it, states its assumptions, derives the failure probability of a redundant group under it, recomputes the two-, three- and four-pump example, and shows why the group failure probability is not the beta factor times a component's mission failure probability.

14.3.1 Definition and the partition of the component failure rate

Consider a common-cause component group of m identical components. The total failure rate of one component is split into two parts,

$$\lambda_{\mathrm{T}} = \lambda_{\mathrm{I}} + \lambda_{\mathrm{C}}, \tag{14.6}$$

where λ_{I} is the rate of component-specific (independent) failures and λ_{C} is the rate of a shared event that disables the entire group. The beta factor is the common-cause fraction of the total,

$$\beta = \frac{\lambda_{\mathrm{C}}}{\lambda_{\mathrm{T}}}, \qquad 0 \le \beta \le 1, \tag{14.7}$$

so that

$$\lambda_{\mathrm{I}} = (1-\beta)\,\lambda_{\mathrm{T}}, \qquad \lambda_{\mathrm{C}} = \beta\,\lambda_{\mathrm{T}}. \tag{14.8}$$

Both parts are rates in h^{-1}; β is dimensionless. With $\beta = 0.10$, ten percent of the modeled failure rate of each component comes from events that disable all members of the group, and ninety percent from failures that affect that component alone.

Remark 14.1. In this chapter β denotes the common-cause beta factor of equation (14.7), a fraction of a failure rate. It is not the Weibull shape parameter, which the rest of the book writes β as well; the Weibull model does not appear in this chapter.

14.3.2 Assumptions and interpretation

The standard beta-factor model rests on four assumptions:

1. all components of the group are susceptible to the same shared event;
2. a common-cause event disables the entire group, never a part of it;
3. the remaining component-specific failures are independent of one another and of the common-cause event;
4. the total failure rate of a component can be meaningfully partitioned into the two contributions.

The model is deliberately restrictive. It is not a general model of statistical dependence: it has one shared mechanism, and that mechanism always has multiplicity m. An event that disables exactly two of four components has no place in it.

The value $\beta = 0.10$ is a statement about rates. It does not mean that there is a 10 % probability that all components fail during a mission; that probability depends on λ_{T} and on the mission duration, and for the example below it is about 0.1 %. It does not mean that the failure indicators of two components have a correlation coefficient of 0.10. Nor does it mean

that 10 % of the failure events in a particular observed sample involved the whole group: that fraction is an event fraction, and Section 14.6.3 shows that it differs from β.

Common mistake

Reading $\beta = 0.1$ as a 10 % probability of losing the whole group. The beta factor partitions a failure rate; the probability of a group-wide failure over a mission follows only after the rate $\beta\lambda_{\mathrm{T}}$ is combined with the mission duration.

14.3.3 Failure probability of a redundant group

Consider a group of m components of which any one is sufficient for the required function, under these assumptions:

- all m components are operating at $t = 0$;
- each component fails independently at the constant rate λ_{I};
- a common-cause event occurs at the constant rate λ_{C}, independently of the component-specific failures, and disables all m components;
- a failed component remains failed: there is no repair during the mission.

Figure 14.1 draws the model as a reliability block diagram. The block on the left stands for the common-cause event, which the group survives only if it does not occur; it is in series with a parallel arrangement of the m component blocks, each of which stands for the independent failure process of one component. The diagram is read as in Chapter 9: the group functions when there is a path from left to right through functioning blocks.

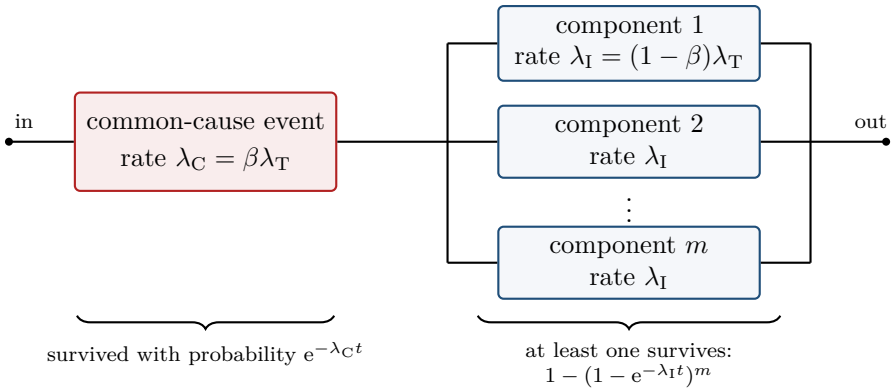


Figure 14.1: Reliability block diagram of the beta-factor model for a group of m components of which one is required. The common-cause event is a block in series with the parallel group of independent component failure processes; the group reliability is the product of the two braced probabilities (equation (14.9)).

The group survives to t if, and only if, no common-cause event occurs in $[0, t]$ and at least one component survives its own independent failure process. The common-cause event is a Poisson process of rate λ_{C}, so the probability of no occurrence in $[0, t]$ is $\mathrm{e}^{-\lambda_{\mathrm{C}}t}$. One component avoids its independent failure with probability $\mathrm{e}^{-\lambda_{\mathrm{I}}t}$, so it fails independently by t with probability $1 - \mathrm{e}^{-\lambda_{\mathrm{I}}t}$, all m components fail independently with probability $(1 - \mathrm{e}^{-\lambda_{\mathrm{I}}t})^m$, and at least one survives with the complement. The two events are independent by assumption, so the group reliability is the product

$$R_{\mathrm{system}}(t) = \mathrm{e}^{-\lambda_{\mathrm{C}}t}\left[1 - \left(1 - \mathrm{e}^{-\lambda_{\mathrm{I}}t}\right)^m\right], \tag{14.9}$$

and the group failure probability, with equation (14.8) substituted, is

$$Q_{\mathrm{sys}}(t) = 1 - \mathrm{e}^{-\beta\lambda_{\mathrm{T}}t}\left[1 - \left(1 - \mathrm{e}^{-(1-\beta)\lambda_{\mathrm{T}}t}\right)^m\right]. \tag{14.10}$$

It is written $Q_m(t)$ where the group size must be shown. The result is exact for the stated exponential (Poisson) event-process model; it is not exact for real equipment whose assumptions have not been checked.

Equation (14.10) is the explicit model of equation (14.2) with its two probabilities expressed through the rates: writing $c = 1 - \mathrm{e}^{-\lambda_C t}$ and $q = 1 - \mathrm{e}^{-\lambda_I t}$, $1 - (1-c)(1-q^m) = c + (1-c)\,q^m$. The beta-factor model therefore supplies what the explicit model left open: c and q both follow from one estimated total rate λ_T, one fraction β and the mission duration. The partition is exact at the level of a single component: a component fails by t if either process reaches it, with probability $1 - \mathrm{e}^{-\lambda_C t}\mathrm{e}^{-\lambda_I t} = 1 - \mathrm{e}^{-\lambda_T t}$, the same as before the split.

14.3.4 Two, three and four pumps

Take a group of emergency cooling pumps of which one is sufficient, with the illustrative values

$$\lambda_T = 10^{-4}\,\mathrm{h}^{-1}, \qquad \beta = 0.10, \qquad t = 100\,\mathrm{h}.$$

By equation (14.8), $\lambda_I = 9 \times 10^{-5}\,\mathrm{h}^{-1}$ and $\lambda_C = 10^{-5}\,\mathrm{h}^{-1}$, so that $\lambda_C t = 0.001$ and $\lambda_I t = 0.009$, and equation (14.10) becomes

$$Q_m(100\,\mathrm{h}) = 1 - \mathrm{e}^{-0.001}\left[1 - \left(1 - \mathrm{e}^{-0.009}\right)^m\right]. \tag{14.11}$$

For two pumps, $\mathrm{e}^{-0.001} = 0.9990005$ and $1 - \mathrm{e}^{-0.009} = 0.0089596$, so

$$Q_2 = 1 - 0.9990005 \times \left(1 - 0.0089596^2\right) = 1 - 0.9990005 \times 0.99991973 = 0.00107969,$$

or 0.10797 %. The comparison model assigns the same total rate λ_T to each pump but no common cause, so $Q_{\text{independent}} = (1 - \mathrm{e}^{-\lambda_T t})^m$ with $1 - \mathrm{e}^{-0.01} = 0.0099502$; for two pumps it is 9.9006×10^{-5}, or 0.00990 %. Table 14.3 gives both models for two, three and four pumps.

Table 14.3: Failure probability of a group of m pumps of which one is required, over $t = 100\,\mathrm{h}$, with $\lambda_T = 10^{-4}\,\mathrm{h}^{-1}$ per pump (illustrative). The independent model has no common cause; the beta-factor model has $\beta = 0.10$ (equation (14.11)).

Pumps m	Independent only	Beta-factor model	Ratio
2	9.9006×10^{-5}	1.07969×10^{-3}	10.9
3	9.8512×10^{-7}	1.00022×10^{-3}	1015
4	9.8022×10^{-9}	9.99507×10^{-4}	1.02×10^{5}
$m \to \infty$	0	9.99500×10^{-4}	–

Figure 14.2 draws the table. The horizontal axis is the number of pumps; the vertical axis is the group failure probability on a logarithmic scale from 10^{-10} to 10^{-1}. Gray bars are the independent model and red bars the beta-factor model; the dashed line is the limit $1 - \mathrm{e}^{-\lambda_C t}$ derived below. The gray bars fall by a factor of about one hundred per added pump; the red bars barely change. In percent, the independent model gives 0.00990, 0.0000985 and 0.000000980 %, the beta-factor model 0.10797, 0.10002 and 0.09995 %.

The independent model predicts that each added pump reduces the failure probability a hundredfold. The beta-factor model predicts a lower limit. As $m \to \infty$ the term $(1 - \mathrm{e}^{-\lambda_I t})^m$ tends to zero, because its base is less than one, and

$$\lim_{m \to \infty} Q_m(t) = 1 - \mathrm{e}^{-\lambda_C t}, \tag{14.12}$$

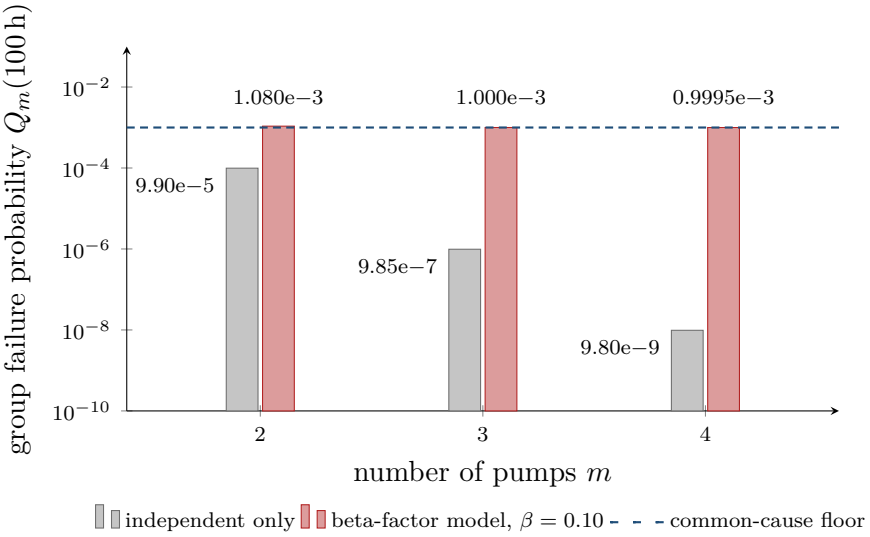


Figure 14.2: Failure probability of a group of m pumps, one required, over 100 h: independent failures with $\lambda_T = 10^{-4}\,\mathrm{h}^{-1}$, $(1-\mathrm{e}^{-\lambda_T t})^m$, against the beta-factor model of equation (14.10) with $\beta = 0.10$; illustrative values (Table 14.3). The common-cause floor is $1 - \mathrm{e}^{-\lambda_C t} = 9.995 \times 10^{-4}$.

which is $1 - \mathrm{e}^{-0.001} = 9.995 \times 10^{-4}$, or 0.09995 %, in the example. This is the common-cause floor: no number of identical pumps that share the modeled mechanism removes it, because the event disables all of them. The four-pump group, at 9.99507×10^{-4}, is already within 7×10^{-9} of the floor. To first order, equation (14.10) reads $Q_m \approx \lambda_C t + (\lambda_I t)^m$, a common-cause term that does not depend on m plus an independent term that falls with m; for two pumps it gives $0.001 + 0.000081 = 0.001081$ against the exact 0.0010797.

Principle 14.1. *In a redundant group with a common cause, adding identical components reduces only the independent contribution. The group failure probability approaches the probability of the shared event, the common-cause floor, which only a change that breaks the coupling factor can lower.*

14.3.5 The system failure probability is not beta times the mission failure probability

The partition of equation (14.8) holds between rates. The corresponding mission probabilities are, for one component,

$$Q_T(t) = 1 - \mathrm{e}^{-\lambda_T t},$$

and for the group-wide event,

$$Q_C(t) = 1 - \mathrm{e}^{-\beta\lambda_T t}.$$

Their ratio is

$$\frac{Q_C(t)}{Q_T(t)} = \frac{1 - \mathrm{e}^{-\beta\lambda_T t}}{1 - \mathrm{e}^{-\lambda_T t}}, \tag{14.13}$$

which is not β in general: $Q_C(t) \neq \beta\, Q_T(t)$. For short missions both numerator and denominator are close to their rare-event approximations, $\beta\lambda_T t$ and $\lambda_T t$, and the ratio tends to β; to first order in $\lambda_T t$ the ratio is $\beta[1 + (1-\beta)\lambda_T t/2]$. As the mission lengthens the component is almost certain to fail by some cause, $Q_T \to 1$, while Q_C keeps growing, so the ratio rises toward 1.

Figure 14.3 plots equation (14.13). The horizontal axis is the dimensionless exposure $\lambda_T t$, from 0 to 5; for $\lambda_T = 10^{-4}\,\mathrm{h}^{-1}$ it spans missions from 0 to 50 000 h. The vertical axis is the ratio.

The solid curves are $\beta = 0.10$ and $\beta = 0.20$; the dotted horizontal lines are the values β that a multiplication by β would give. For the 100 h pump mission, $\lambda_T t = 0.01$, the ratio is 0.1005 and the product βQ_T is a good approximation: 9.95×10^{-4} against the exact 9.995×10^{-4}. At $\lambda_T t = 1$ (10 000 h) the ratio is 0.1505: $\beta Q_T = 0.0632$ against $Q_C = 0.0952$, an underestimate by a third. At $\lambda_T t = 5$ the ratio is 0.396.

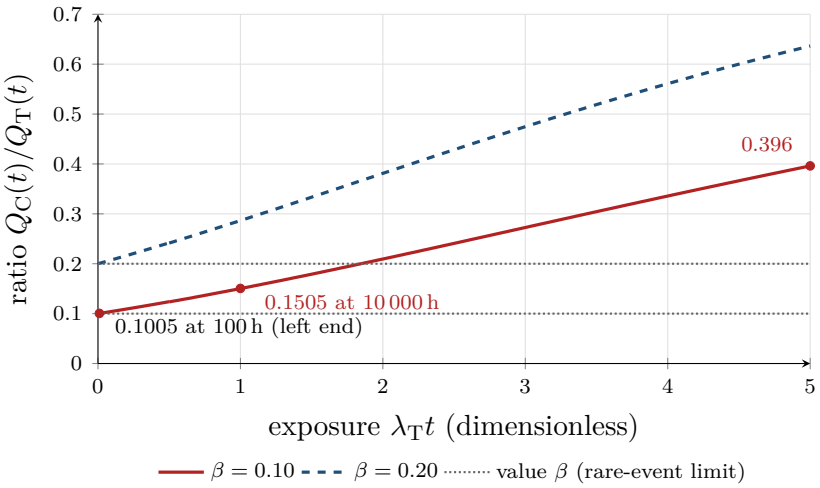


Figure 14.3: Ratio of the mission probability of the group-wide common-cause event to a component's total mission failure probability, $Q_C/Q_T = (1 - e^{-\beta x})/(1 - e^{-x})$ with $x = \lambda_T t$ (equation (14.13)), for $\beta = 0.10$ and 0.20. The ratio equals β only in the limit of short missions. Marked: $x = 0.01$, 1 and 5, which are 100, 10 000 and 50 000 h for $\lambda_T = 10^{-4}\,\mathrm{h}^{-1}$.

Multiplying a component failure probability by β is therefore a rare-event approximation, adequate when the probabilities are small and wrong by a growing margin for long missions; there the exponential expressions must be used. The same caution applies to the per-demand form of the model used for standby equipment. There the beta-factor model is often written directly on probabilities,

$$q_I = (1 - \beta)\, q_T, \qquad q_C = \beta\, q_T, \tag{14.14}$$

with q_T the total failure probability per demand of one component, q_I the probability of its independent failure and q_C the probability of the group-wide failure on that demand. This is a parameterization of the risk model, not an identity that follows from probability theory. If the two failure events were independent with these probabilities, the probability that the component fails would be the probability of their union, $q_C + q_I - q_C q_I = q_T - \beta(1 - \beta)\, q_T^2$, slightly less than q_T. The discrepancy is negligible when q_T is small, but the denominator of the probability (per demand, per test, per mission), the testing regime and the group definition must be the ones under which β was estimated.

14.4 The alpha-factor model

The alpha-factor model replaces the single group-wide event of the beta-factor model by events of every multiplicity from one to m. This section defines the alpha factors, derives event rates and basic-event probabilities from them, gives an exact evaluation of the resulting event model over a mission, and works examples with two, three and four components.

14.4.1 Multiplicity of failure events

The beta-factor model distinguishes two kinds of event: the independent failure of one component and a common-cause event that disables the whole group. In a group of four chillers, a disturbance

of the shared cooling-water supply can disable exactly two of them; a maintenance error can affect three controllers of four; an electrical transient can damage a subset of identical cards. None of these is a single independent failure or a failure of the full group, and the beta-factor model has no place for them.

The alpha-factor model represents the multiplicity of failure events. For a common-cause component group of m components, the alpha factor of order k is the conditional probability

$$\alpha_k = \Pr(\text{exactly } k \text{ components affected} \mid \text{a failure event of the group}), \qquad k = 1, \ldots, m, \tag{14.15}$$

with

$$\sum_{k=1}^{m} \alpha_k = 1. \tag{14.16}$$

The factor α_1 is the fraction of events that affect one component only, which are the independent failures; α_2 to α_m are the fractions of common-cause events of each multiplicity. In a four-component group, α_1, α_2, α_3 and α_4 are the fractions of group failure events that affect exactly one, two, three and all four components.

Figure 14.4 shows the four multiplicities of a four-pump group. Each row is one kind of event; the four squares are the pumps, red where the event disables the pump and green where it does not. The third column gives the number of distinct subsets of that size, $\binom{4}{k}$, and the last two columns state whether each model represents the event. The beta-factor model represents only the first and last rows; the alpha-factor model represents all four.

event	pumps A B C D	subsets	beta factor	alpha factor
one pump fails (α_1)		4	yes (λ_{I})	yes
two pumps fail (α_2)		6	not separately	yes
three pumps fail (α_3)		4	not separately	yes
all four fail (α_4)		1	yes (λ_{C})	yes



Figure 14.4: Failure-event multiplicities in a group of four pumps. Each row is one kind of event, drawn for one of its subsets; the multiplicity is the number of pumps the event disables. The beta-factor model represents only single failures and failures of the whole group; the alpha-factor model represents every multiplicity.

The alpha factors are fractions of *events*, not fractions of failed components. A single event that disables four chillers counts as one group event and as four component failures. This distinction is the origin of the multiplicity weights in every formula that converts alpha factors into component quantities.

14.4.2 Group-event rates from the alpha factors

The alpha factors give the shares of events of each multiplicity but not how often events occur. To obtain rates, take an explicit, symmetric continuous-time event model: failure events of the group occur as a Poisson process of rate r (in h^{-1}); each event independently receives a multiplicity k with probability α_k; and, given k, every subset of k components is equally likely to be the affected one. A Poisson process whose events are sorted at random into classes splits into independent Poisson processes, one per class (the standard thinning property). Events of multiplicity k therefore occur at the rate

$$r_k = r\,\alpha_k. \tag{14.17}$$

The mean number of components affected by one event is the mean multiplicity

$$\bar{k} = \sum_{k=1}^{m} k\,\alpha_k, \tag{14.18}$$

so the group produces component failures at the expected rate $r\bar{k}$. The total failure rate of one specified component follows by symmetry. An event of multiplicity k contains that component with probability k/m, since it chooses k of the m components at random, so the rate of events that reach the component is

$$\lambda_{\mathrm{T}} = \sum_{k=1}^{m} r_k \frac{k}{m} = \frac{r\bar{k}}{m}. \tag{14.19}$$

Solving for r and substituting in equation (14.17),

$$r = \frac{m\,\lambda_{\mathrm{T}}}{\bar{k}}, \qquad r_k = \frac{m\,\lambda_{\mathrm{T}}\,\alpha_k}{\bar{k}}. \tag{14.20}$$

There are $\binom{m}{k}$ distinct subsets of k components, all equally likely, so the rate of an event that disables one particular subset of size k is

$$r_{k,\mathrm{subset}} = \frac{r_k}{\binom{m}{k}} = \frac{m\,\lambda_{\mathrm{T}}\,\alpha_k}{\bar{k}\,\binom{m}{k}}. \tag{14.21}$$

The binomial identity $\binom{m}{k} = \frac{m}{k}\binom{m-1}{k-1}$, that is $m/\binom{m}{k} = k/\binom{m-1}{k-1}$, gives the equivalent form

$$r_{k,\mathrm{subset}} = \frac{k\,\alpha_k}{\bar{k}\,\binom{m-1}{k-1}}\,\lambda_{\mathrm{T}}. \tag{14.22}$$

A check: a given component belongs to $\binom{m-1}{k-1}$ subsets of size k, so the subset rates that reach it add up to $\sum_k \binom{m-1}{k-1} r_{k,\mathrm{subset}} = \sum_k k\alpha_k\lambda_{\mathrm{T}}/\bar{k} = \lambda_{\mathrm{T}}$, and the share of a component's failure rate due to events of multiplicity k is $k\alpha_k/\bar{k}$. These are rates of the explicit Poisson model just defined; the classical alpha-factor model is also stated in terms of probabilities per demand, with two testing conventions, which the next subsection derives.

14.4.3 Basic-event probabilities under non-staggered and staggered testing

For standby equipment the classical formulation works with probabilities per demand. Let Q_{T} be the total failure probability per demand of one component, and $Q_k^{(m)}$ the probability per demand of the basic event in which one specified subset of k components of a group of m fails from a common cause ($k = 1$ is the independent failure of one specified component). Since a component belongs to $\binom{m-1}{k-1}$ subsets of size k, consistency requires, to first order, $Q_{\mathrm{T}} = \sum_k \binom{m-1}{k-1} Q_k^{(m)}$. The expression of $Q_k^{(m)}$ through the alpha factors depends on how the failure events were observed, which for standby equipment means how the group was tested.

Non-staggered testing. All m components are tested together, N_{D} times. Let n_k be the number of events of multiplicity k found, $n = \sum_k n_k$, so that $\alpha_k = n_k/n$ and $\bar{k} = \sum_k k n_k/n$. Each test is one opportunity for each of the $\binom{m}{k}$ subsets of size k, so $Q_k^{(m)} = n_k/[\binom{m}{k} N_{\mathrm{D}}]$; and each test is one demand on each of the m components, which together failed $\sum_k k n_k = n\bar{k}$ times, so $Q_{\mathrm{T}} = n\bar{k}/(mN_{\mathrm{D}})$. Eliminating n and N_{D},

$$Q_{k,\mathrm{NS}}^{(m)} = \frac{k\,\alpha_k}{\bar{k}\,\binom{m-1}{k-1}}\,Q_{\mathrm{T}}, \tag{14.23}$$

the algebraic form of the subset rate (14.22) with λ_{T} replaced by Q_{T}.

Staggered testing. Each component is tested N_{D} times at its own times, evenly offset from the others, and a test that finds a failure is followed by tests of the other components, so each event is discovered once, at a test of one of its members. There are mN_{D} tests, and each is an opportunity for the $\binom{m-1}{k-1}$ subsets of size k that contain the tested component, so $Q_k^{(m)} = n_k/[m\binom{m-1}{k-1}N_{\mathrm{D}}]$; the probability that the tested component is found failed counts each event once, $Q_{\mathrm{T}} = n/(mN_{\mathrm{D}})$. Eliminating as before,

$$Q_{k,\mathrm{S}}^{(m)} = \frac{\alpha_k}{\binom{m-1}{k-1}}\, Q_{\mathrm{T}}. \tag{14.24}$$

Both expressions satisfy the consistency condition: the sum $\sum_k \binom{m-1}{k-1} Q_k^{(m)}$ is $\sum_k k\alpha_k Q_{\mathrm{T}}/\bar{k} = Q_{\mathrm{T}}$ for equation (14.23) and $\sum_k \alpha_k Q_{\mathrm{T}} = Q_{\mathrm{T}}$ for equation (14.24). They differ by the factor $k/\bar{k}$.

These are the established basic-event expressions of the classical alpha-factor model, here derived by counting in an idealized test program. They require the alpha factors and Q_{T} to be estimated or defined under the same testing convention: under the staggered convention, Q_{T} counts events, not component failures. They must not be substituted indiscriminately into the continuous-time event model of Section 14.4.2, which was derived separately and only shares the algebraic form of the non-staggered expression. Section 14.6.4 compares the two conventions numerically.

14.4.4 Exact evaluation of the alpha-factor event model

The continuous-time model of Section 14.4.2 can be evaluated exactly over a mission, including several events in one mission and overlaps between their affected subsets. Assume, as there, that events of each subset occur as independent homogeneous Poisson processes, and that there is no repair. Let S be a particular set of s components. The set survives the mission if no event reaches any of its members. An event of multiplicity k misses S only when its subset lies entirely outside S; the fraction of size-k subsets that do so is $\binom{m-s}{k}/\binom{m}{k}$, which is zero when $k > m - s$. The rate of events that reach at least one member of S is therefore

$$h_s = \sum_{k=1}^{m} r_k \left[1 - \frac{\binom{m-s}{k}}{\binom{m}{k}}\right], \qquad h_0 = 0, \tag{14.25}$$

and, as the superposition of independent Poisson processes is a Poisson process,

$$\Pr(\text{all components in } S \text{ survive to } t) = \mathrm{e}^{-h_s t}. \tag{14.26}$$

For $s = 1$ the bracket equals k/m and $h_1 = \lambda_{\mathrm{T}}$, as equation (14.19) requires; for $s = m$ every event reaches the set and $h_m = r$.

Let $N_{\mathrm{surv}}(t)$ be the number of surviving components. The probability that exactly the components of a particular set of size j survive and the other $m - j$ have failed follows from equation (14.26) by inclusion and exclusion over the failed components: start from the probability that the j survive, subtract the probabilities that they and one more survive, add back those with two more, and so on. Multiplying by the $\binom{m}{j}$ sets of size j, which are equally likely by symmetry,

$$\Pr(N_{\mathrm{surv}}(t) = j) = \binom{m}{j} \sum_{v=0}^{m-j} (-1)^v \binom{m-j}{v} \mathrm{e}^{-h_{j+v}t}. \tag{14.27}$$

A system that requires at least a functioning components, a a-out-of-m system, fails with probability

$$Q_{\mathrm{sys}}(t) = \sum_{j=0}^{a-1} \Pr(N_{\mathrm{surv}}(t) = j). \tag{14.28}$$

This is the exact finite-time failure probability of the marked-Poisson shock model. The classical alpha-factor method itself specifies relative event multiplicities and basic-event probabilities; the Poisson model added here is one explicit dynamic interpretation that permits exact mission calculations, not a requirement of alpha-factor analysis. All the exact values below were computed with equations (14.25)–(14.28) and checked against a direct solution of the corresponding Markov model over the 2^m sets of failed components.

14.4.5 Example A: two components and the relation to the beta factor

Take $m = 2$ components A and B, $\lambda_{\mathrm{T}} = 10^{-4}\,\mathrm{h}^{-1}$, and $(\alpha_1, \alpha_2) = (0.90, 0.10)$: ninety percent of group failure events affect one pump and ten percent affect both. The mean multiplicity is

$$\bar{k} = 1(0.90) + 2(0.10) = 1.10,$$

and equation (14.20) gives

$$r = \frac{2 \times 10^{-4}}{1.10} = 1.81818 \times 10^{-4}\,\mathrm{h}^{-1}, \qquad r_1 = 1.63636 \times 10^{-4}\,\mathrm{h}^{-1}, \qquad r_2 = 1.81818 \times 10^{-5}\,\mathrm{h}^{-1}.$$

There are two single-component events, A only and B only, each at the rate $r_{\{A\}} = r_{\{B\}} = r_1/2 = 8.18182 \times 10^{-5}\,\mathrm{h}^{-1}$, and one two-component event at $r_{\{A,B\}} = r_2 = 1.81818 \times 10^{-5}\,\mathrm{h}^{-1}$. The model consists of three independent event processes, which is the structure of the beta-factor model with $m = 2$. With one pump required, the system fails over a 100 h mission if the A-and-B event occurs or if both single-component processes occur:

$$Q_2(t) = 1 - \mathrm{e}^{-r_2 t}\left[1 - \left(1 - \mathrm{e}^{-r_{\{A\}} t}\right)^2\right].$$

With $r_2 t = 0.00181818$ and $r_{\{A\}} t = 0.00818182$, $\mathrm{e}^{-r_2 t} = 0.99818347$ and $1 - \mathrm{e}^{-r_{\{A\}} t} = 0.00814844$, so

$$Q_2(100\,\mathrm{h}) = 1 - 0.99818347 \times \left(1 - 0.00814844^2\right) = 0.001882806,$$

about 0.1883 %. Equation (14.27) gives the same value.

The beta factor of this group is not α_2. The equivalent beta factor is the fraction of one component's failure rate due to events that disable the whole group,

$$\beta_{\mathrm{equiv}} = \frac{r_2}{\lambda_{\mathrm{T}}} = \frac{2\alpha_2/\bar{k}}{1} = \frac{2\alpha_2}{\alpha_1 + 2\alpha_2}, \tag{14.29}$$

where the middle step uses $r_2 = 2\lambda_{\mathrm{T}}\alpha_2/\bar{k}$ from equation (14.20). Substituting, $\beta_{\mathrm{equiv}} = 0.20/1.10 = 0.1818$. Although only 10 % of group events affect both components, those events supply 18.18 % of each component's failure rate, because each of them is a failure of two components. The beta-factor model with $\beta = 0.1818$ and $\lambda_{\mathrm{T}} = 10^{-4}\,\mathrm{h}^{-1}$ reproduces $Q_2(100\,\mathrm{h}) = 0.001882806$ exactly, since for two components the two models have the same three event processes.

14.4.6 Example B: three components with two required

Take $m = 3$, $\lambda_{\mathrm{T}} = 10^{-4}\,\mathrm{h}^{-1}$, and $(\alpha_1, \alpha_2, \alpha_3) = (0.80, 0.15, 0.05)$. The mean multiplicity is $\bar{k} = 0.80 + 2(0.15) + 3(0.05) = 1.25$ and the total group-event rate is $r = 3 \times 10^{-4}/1.25 = 2.4 \times 10^{-4}\,\mathrm{h}^{-1}$. Table 14.4 lists the event rates by multiplicity, the rate of one particular subset, and the share of a component's failure rate, $k\alpha_k/\bar{k}$. For a particular pair, such as A and B, $r_{\{A,B\}} = 3.60 \times 10^{-5}/\binom{3}{2} = 1.20 \times 10^{-5}\,\mathrm{h}^{-1}$, the same as the rate of the triple event.

The system is the data center cooling plant of the book's running case, three chillers of which two are required, here with the illustrative rates of the table and a mission of 100 h.

Table 14.4: Event rates of Example B: $m = 3$, $\lambda_{\mathrm{T}} = 10^{-4}\,\mathrm{h}^{-1}$, $(\alpha_1, \alpha_2, \alpha_3) = (0.80, 0.15, 0.05)$, $\bar{k} = 1.25$ (equations (14.20) and (14.22)).

Components affected k	α_k	r_k (h^{-1})	subsets	$r_{k,\mathrm{subset}}$ (h^{-1})	share $k\alpha_k/\bar{k}$
1	0.80	1.92×10^{-4}	3	6.40×10^{-5}	0.64
2	0.15	3.60×10^{-5}	3	1.20×10^{-5}	0.24
3	0.05	1.20×10^{-5}	1	1.20×10^{-5}	0.12

It fails when two or more chillers are unavailable. An event of multiplicity two or three is immediately sufficient; their combined rate is $r_2 + r_3 = 4.8 \times 10^{-5}\,\mathrm{h}^{-1}$, and for a short mission the direct contribution is approximately

$$Q_{\mathrm{direct}}(100\,\mathrm{h}) \approx (r_2 + r_3)\, t = 0.0048,$$

or 0.48 %. Two single-component events on different chillers also fail the system, and several events can overlap. Figure 14.5 draws the corresponding fault tree, as in Chapter 10. The top event is an OR gate over seven inputs: the four common-cause basic events, one per affected subset (the circles labeled AB, AC, BC and ABC), and three AND gates, each over the independent basic events of two chillers. Below each circle is its mission probability, $1-\mathrm{e}^{-r_{k,\mathrm{subset}}t}$. The seven inputs are the minimal cut sets of the system (Chapter 11): four of order one and three of order two.

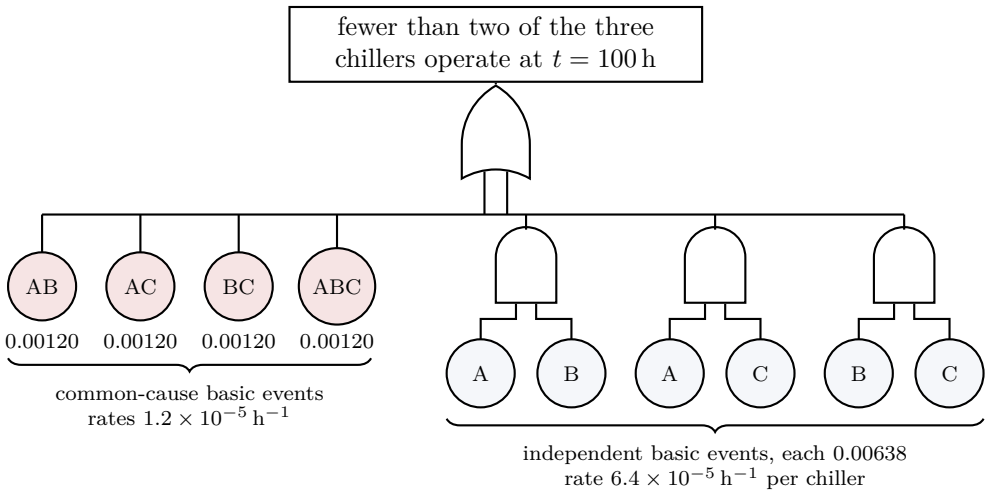


Figure 14.5: Fault tree of the 2-out-of-3 chiller plant under the alpha-factor model of Example B (Table 14.4), mission 100 h, illustrative values. Red circles are common-cause basic events of the subsets AB, AC, BC and ABC; white circles are independent failures of one chiller, each appearing in two AND gates. Numbers are mission probabilities $1 - \mathrm{e}^{-r_{k,\mathrm{subset}}t}$.

The rare-event approximation over the minimal cut sets adds their probabilities: $4 \times 0.00119928 + 3 \times 0.00637956^2 = 0.00479712 + 0.00012210 = 0.0049192$. The exact evaluation uses equation (14.25). For $s = 1$, $h_1 = \lambda_{\mathrm{T}} = 10^{-4}\,\mathrm{h}^{-1}$. For $s = 2$, a single event misses a given pair with probability $1/3$, a pair event never misses it (the only pair disjoint from a pair would need four components), and the triple event never does, so $h_2 = 1.92 \times 10^{-4}(2/3) + 3.6 \times 10^{-5} + 1.2 \times 10^{-5} = 1.76 \times 10^{-4}\,\mathrm{h}^{-1}$. For $s = 3$, $h_3 = r = 2.4 \times 10^{-4}\,\mathrm{h}^{-1}$. With $\mathrm{e}^{-0.01} = 0.99004983$,

$\mathrm{e}^{-0.0176} = 0.98255398$ and $\mathrm{e}^{-0.024} = 0.97628571$, equation (14.27) gives

$$\begin{aligned}
\Pr(N_{\text{surv}} = 3) &= \mathrm{e}^{-h_3 t} = 0.97628571, \\
\Pr(N_{\text{surv}} = 2) &= 3\left[\mathrm{e}^{-h_2 t} - \mathrm{e}^{-h_3 t}\right] = 0.01880480, \\
\Pr(N_{\text{surv}} = 1) &= 3\left[\mathrm{e}^{-h_1 t} - 2\,\mathrm{e}^{-h_2 t} + \mathrm{e}^{-h_3 t}\right] = 0.00368278, \\
\Pr(N_{\text{surv}} = 0) &= 1 - 3\,\mathrm{e}^{-h_1 t} + 3\,\mathrm{e}^{-h_2 t} - \mathrm{e}^{-h_3 t} = 0.00122672,
\end{aligned}$$

and by equation (14.28) with $a = 2$,

$$Q_{\text{2-out-of-3}}(100\,\mathrm{h}) = 0.00368278 + 0.00122672 = 0.00490949,$$

or 0.49095 %. The direct multi-component events, 0.0048, capture most but not all of it; the remaining 0.00011 comes mainly from pairs of single-component events on different chillers. The rare-event sum over the minimal cut sets, 0.0049192, is 0.2 % high. An alpha-factor model must therefore be combined with the system's Boolean failure logic, not read off a single multiplicity.

14.4.7 Example C: four components with one required

Take $m = 4$, $\lambda_{\mathrm{T}} = 10^{-4}\,\mathrm{h}^{-1}$, and $(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (0.75, 0.15, 0.07, 0.03)$. The mean multiplicity is

$$\bar{k} = 1(0.75) + 2(0.15) + 3(0.07) + 4(0.03) = 1.38,$$

and the group-event rate is $r = 4 \times 10^{-4}/1.38 = 2.89855 \times 10^{-4}\,\mathrm{h}^{-1}$. Table 14.5 lists the event rates. A component's failure rate now comes 54 % from independent failures and 46 % from events of multiplicity two to four.

Table 14.5: Event rates of Example C: $m = 4$, $\lambda_{\mathrm{T}} = 10^{-4}\,\mathrm{h}^{-1}$, $(\alpha_1, \ldots, \alpha_4) = (0.75, 0.15, 0.07, 0.03)$, $\bar{k} = 1.38$.

Components affected k	α_k	r_k (h^{-1})	subsets	$r_{k,\text{subset}}$ (h^{-1})	share $k\alpha_k/\bar{k}$
1	0.75	2.17391×10^{-4}	4	5.43478×10^{-5}	0.5435
2	0.15	4.34783×10^{-5}	6	7.24638×10^{-6}	0.2174
3	0.07	2.02899×10^{-5}	4	5.07246×10^{-6}	0.1522
4	0.03	8.69565×10^{-6}	1	8.69565×10^{-6}	0.0870

Suppose any one component can sustain the required function, so that the system fails only when all four are unavailable. A single four-component event is sufficient, and over 100 h it contributes approximately $r_4 t = 8.69565 \times 10^{-6} \times 100 = 0.000869565$. That is not the total, because there are other pathways to the loss of all four:

- a three-component event and a failure of the fourth component;
- two two-component events on complementary pairs;
- a two-component event and two single-component failures;
- four separate single-component failures;
- overlapping combinations of several events.

The exact evaluation sums them all. Equation (14.25) gives $h_1 = 1.00000 \times 10^{-4}$, $h_2 = 1.73913 \times 10^{-4}$, $h_3 = 2.35507 \times 10^{-4}$ and $h_4 = 2.89855 \times 10^{-4}\,\mathrm{h}^{-1}$, and equation (14.27) gives the distribution of survivors at 100 h: 0.971431 for four, 0.0211755 for three, 0.00444441 for two, 0.00206170 for one and 0.000887799 for none. With $a = 1$,

$$Q_{\text{1-out-of-4}}(100\,\mathrm{h}) = \Pr(N_{\text{surv}} = 0) = 0.000887799,$$

or 0.08878 %, of which the direct four-component event supplies 98 %.

Table 14.6 collects the three examples. They use different alpha factors and different success criteria, so they are not a controlled comparison of group sizes, and they do not show that one architecture is intrinsically better than another. They show how the multiplicity of events interacts with the success logic of the system: in the two-component and four-component groups with one unit required, only the events that disable every unit fail the system directly, while in the 2-out-of-3 plant every common-cause event does.

Table 14.6: Exact mission failure probabilities of the three alpha-factor examples, $\lambda_{\mathrm{T}} = 10^{-4}\,\mathrm{h}^{-1}$, $t = 100\,\mathrm{h}$; illustrative values. The direct term is the rare-event probability of the common-cause events that fail the system by themselves.

Group	Requirement	Alpha factors	Direct term	Exact Q_{sys}
2 components	at least 1 operates	$(0.90, 0.10)$	0.001818	0.0018828 (0.18828 %)
3 components	at least 2 operate	$(0.80, 0.15, 0.05)$	0.004800	0.0049095 (0.49095 %)
4 components	at least 1 operates	$(0.75, 0.15, 0.07, 0.03)$	0.000870	0.0008878 (0.08878 %)

14.5 Beta against alpha

The two models differ in the multiplicities they can represent, in the data they need, and in how their parameters are converted. This section compares them, derives the conversion between them, shows that an alpha factor is not a joint failure probability, and tabulates the effect of the beta factor on groups of two, three and four units.

14.5.1 What each model can represent

Table 14.7 sets the two models side by side for a group of four components. Both represent single-component failures and failures of the entire group. Only the alpha-factor model represents events of intermediate multiplicity, which the beta-factor model either omits or must assign to one of its two kinds of event. The price is more parameters and more detailed data: the beta-factor model needs only the aggregate common-cause share of the failures, while the alpha-factor model needs each recorded event classified by how many components it affected.

The beta-factor model is a restricted multiplicity model. It allows only single-component events and events of the whole group, which in alpha factors is

$$\alpha_2 = \cdots = \alpha_{m-1} = 0.$$

Published beta and alpha factors can rest on different denominators and estimation procedures, so a numerical value is never transferred from one model to the other without the conversion of the next subsection.

14.5.2 Converting between beta and alpha

In the restricted model, the beta factor is the share of a component's failure rate due to whole-group events. By the share formula of Section 14.4.2, events of multiplicity m supply $m\alpha_m/\bar{k}$ of it, and with only α_1 and α_m nonzero $\bar{k} = \alpha_1 + m\alpha_m$, so

$$\beta = \frac{m\,\alpha_m}{\alpha_1 + m\,\alpha_m}, \tag{14.30}$$

Table 14.7: The beta-factor and alpha-factor models compared, for a common-cause group of four components.

Property	Beta-factor model	Alpha-factor model
Single-component failures	yes	yes
Two-component CCF in a four-component group	not separately	yes
Three-component CCF in a four-component group	not separately	yes
Entire-group CCF	yes	yes
Main parameters	β and the total component failure rate or probability	$\alpha_1, \ldots, \alpha_m$ (summing to one) and the total rate or probability
Complexity	low	higher
Data required	aggregate CCF contribution	multiplicity of each failure event
Group-size sensitivity	restricted	flexible
Typical use	screening, simplified PRA, limited data	detailed PRA of redundant systems

under the symmetric event-rate interpretation. For $m = 2$ this is equation (14.29). In general α_m does not equal β: the conversion needs the multiplicity weight m because alpha factors count group events, while β partitions the per-component failure rate. For $\alpha_m = 0.10$, 0.05 and 0.03 in groups of two, three and four, equation (14.30) gives $\beta = 0.1818$, 0.1364 and 0.1101.

When the alpha factors of intermediate order are not zero, no single beta factor reproduces the model, and the choice of conversion decides the answer. Example C illustrates the two obvious choices for its 1-out-of-4 system. Lumping every multi-component event into a whole-group event gives $\beta = \sum_{k\geq 2} k\alpha_k/\bar{k} = 0.63/1.38 = 0.4565$ and, by equation (14.10), $Q_4(100\,\mathrm{h}) = 0.004555$, five times the exact 0.000888, because events that leave two or three pumps running are treated as fatal. Keeping only the whole-group share, $\beta = 4\alpha_4/\bar{k} = 0.0870$, and treating the rest as independent, gives 0.000869, slightly below the exact value, because the partial events that combine into a loss of all four are dropped. The alpha-factor result lies between the two, and only the alpha-factor model places it.

14.5.3 Alpha factors are not joint failure probabilities

In Example B, $\alpha_3 = 0.05$ means that five percent of the modeled failure events of the group affect all three components. It does not mean that the probability that all three fail is 0.05, and it does not mean that the pairwise correlation of the component failure indicators is 0.05. The probability that all three chillers have failed by 100 h is $\Pr(N_{\mathrm{surv}} = 0) = 0.00123$ (Section 14.4.6), forty times smaller than α_3, and it would be different again for another mission duration or another total rate.

An alpha factor is a conditional probability: it describes the multiplicity of an event given that an event has occurred. To obtain an unconditional component or system failure probability, three further pieces of information are needed: the rate at which events occur (or the probability per demand), the mission duration or the test interval, and the distribution of the affected subsets, which the symmetric model takes as uniform. Alpha factors alone do not establish the probability of system failure.

Common mistake

Reading an alpha factor as the probability of a multiple failure. The factor α_k is the fraction of failure events that have multiplicity k, conditional on an event; the probability that k specified components fail over a mission is $1 - \mathrm{e}^{-r_{k,\mathrm{subset}}t}$, with $r_{k,\mathrm{subset}}$ from equation (14.22).

14.5.4 Redundancy against the beta factor: a worked table

The effect of the beta factor on the value of redundancy can be tabulated from equation (14.10). The model is the one of Section 14.3.3: all components start operating, any one is sufficient, the total failure rate of each component is fixed at $\lambda_\mathrm{T} = 10^{-4}\,\mathrm{h}^{-1}$, and the beta factor partitions it between independent and whole-group events. Table 14.8 gives the group failure probability for $m = 2$, 3 and 4 and seven values of β from 0 to 0.40, for missions of 100 h and 1000 h. The row $\beta = 0$ is the independent model.

Table 14.8: Failure probability of a group of m components, one required, under the beta-factor model (equation (14.10)) with $\lambda_\mathrm{T} = 10^{-4}\,\mathrm{h}^{-1}$; illustrative values. The last column is the common-cause floor $1 - \mathrm{e}^{-\beta\lambda_\mathrm{T}t}$ for $t = 1000$ h.

	$t = 100$ h			$t = 1000$ h			
β	$m = 2$	$m = 3$	$m = 4$	$m = 2$	$m = 3$	$m = 4$	floor
0	9.901e−5	9.851e−7	9.802e−9	9.056e−3	8.618e−4	8.201e−5	0
0.01	1.970e−4	1.010e−4	1.000e−4	9.875e−3	1.836e−3	1.078e−3	9.995e−4
0.02	2.951e−4	2.009e−4	2.000e−4	1.070e−2	2.810e−3	2.074e−3	1.998e−3
0.05	5.892e−4	5.007e−4	4.999e−4	1.316e−2	5.728e−3	5.055e−3	4.988e−3
0.10	1.080e−3	1.000e−3	9.995e−4	1.728e−2	1.058e−2	1.000e−2	9.950e−3
0.20	2.061e−3	1.999e−3	1.998e−3	2.560e−2	2.025e−2	1.984e−2	1.980e−2
0.40	4.028e−3	3.992e−3	3.992e−3	4.247e−2	3.940e−2	3.922e−2	3.921e−2

One entry shows the arithmetic. For $m = 3$, $\beta = 0.05$ and $t = 1000$ h, the exposures are $\lambda_\mathrm{C}t = 0.05 \times 10^{-4} \times 1000 = 0.005$ and $\lambda_\mathrm{I}t = 0.095$, so $\mathrm{e}^{-0.005} = 0.995012$, $1 - \mathrm{e}^{-0.095} = 0.090627$, and

$$Q_3 = 1 - 0.995012 \times \left(1 - 0.090627^3\right) = 1 - 0.995012 \times 0.999256 = 0.005728.$$

The independent term, $0.090627^3 = 0.000744$, is now an eighth of the total; the rest is the floor 0.004988.

Figure 14.6 plots the 1000 h columns as continuous functions of β. The horizontal axis is the beta factor from 0 to 0.4; the vertical axis is the group failure probability on a logarithmic scale. The three solid curves are two, three and four components; the dashed curve is the common-cause floor, which does not depend on m and lies almost under the four-component curve for β above about 0.05; the dots mark $\beta = 0.10$.

Three trends follow from the table and the figure. First, the benefit of an added component shrinks as β grows. At $\beta = 0$ and 1000 h, four components are 110 times less likely to fail than two; at $\beta = 0.01$ the ratio is 9.2, at $\beta = 0.05$ it is 2.6, at $\beta = 0.10$ it is 1.7, and at $\beta = 0.40$ it is 1.08. Second, already at $\beta = 0.01$ the three- and four-component groups at 100 h sit on the floor, 1.010×10^{-4} and 1.000×10^{-4} against the independent 9.851×10^{-7} and 9.802×10^{-9}: a common-cause share of one percent raises the failure probability of a four-pump group ten thousandfold. Third, the floor grows with the mission duration in proportion to $\beta\lambda_\mathrm{T}t$, while the independent term of an m-unit group grows as $[(1-\beta)\lambda_\mathrm{T}t]^m$; a longer mission therefore leaves more room for redundancy to act, which is why the curves separate more at 1000 h than at 100 h. The residual failure probability of a highly redundant group is dominated by the group-wide event, and adding components becomes progressively less valuable as β increases.

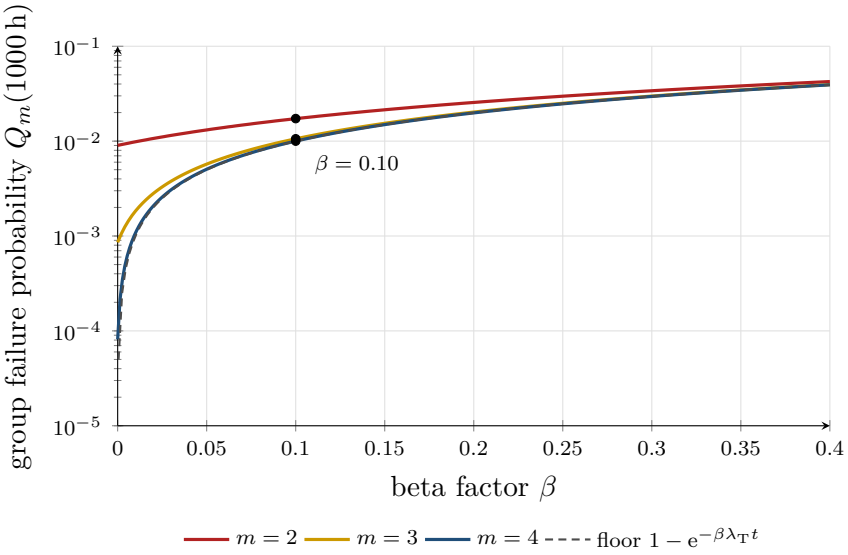


Figure 14.6: Failure probability of a group of m components, one required, over $t = 1000\,\mathrm{h}$ under the beta-factor model, $Q_m = 1 - \mathrm{e}^{-\beta\lambda_{\mathrm{T}}t}[1 - (1 - \mathrm{e}^{-(1-\beta)\lambda_{\mathrm{T}}t})^m]$ (equation (14.10)) with $\lambda_{\mathrm{T}} = 10^{-4}\,\mathrm{h}^{-1}$, against β; illustrative values (Table 14.8). The dashed curve is the common-cause floor.

14.6 Estimating the parameters from data

The parameters of both models are estimated from records of failure events, classified by the components they affected. This section states what the records must contain, derives the estimators of the alpha and beta factors with their statistical uncertainty, compares the two testing conventions, and closes with the choice of model in practice.

14.6.1 What must be recorded

Suppose a facility operates four identical emergency cooling pumps. A conventional maintenance database records, for each failed pump, the equipment identifier, the time and the failure mode: Pump A at 1500 h with a bearing seizure, Pump B at 2200 h with a motor failure (illustrative). That is enough to estimate λ_{T} but not the common-cause parameters, which depend on the relations among failures: whether the two pumps shared a maintenance procedure, whether the same lubricant contamination was present, whether a common electrical transient occurred, whether both pumps were actually exposed and susceptible, whether the second failure was caused by the first (a cascade, not a common cause), and whether the failures were found together or at different inspections.

A record useful for common-cause analysis therefore contains:

- the set of affected components and the common-cause component group to which they belong;
- failure times, demands and inspection times;
- shared environmental conditions and operating modes;
- the root cause and the coupling factor;
- whether each failure was complete, partial or latent (present but undiscovered);
- the testing schedule and the observation exposure.

The group definition is especially important. A common-cause component group should comprise components that are susceptible to a meaningful set of shared mechanisms, not every component of the same general equipment category. Two identical pumps in different buildings, with different crews and supplies, may share a manufacturer and little else.

14.6.2 Alpha-factor estimators and their uncertainty

Suppose that 100 correctly classified failure events have been observed in a group of four pumps, with the multiplicities of the second column of Table 14.9. If events are independently sampled and equally observable, the counts $(n_1, \ldots, n_m)$ are multinomial with probabilities $(\alpha_1, \ldots, \alpha_m)$, and the maximum likelihood estimate is the observed fraction,

$$\hat{\alpha}_k = \frac{n_k}{\sum_{j=1}^{m} n_j}, \tag{14.31}$$

as maximizing $\ell(\boldsymbol{\alpha}) = \sum_k n_k \ln \alpha_k$ subject to $\sum_k \alpha_k = 1$ with a Lagrange multiplier shows (Chapter 7). For the counts of the table, $\hat{\boldsymbol{\alpha}} = (0.75, 0.15, 0.07, 0.03)$, the alpha factors of Example C. The estimator describes the observed multiplicity distribution. It is biased if events of different multiplicities have different probabilities of being detected, recorded or classified correctly: if single failures are found at routine inspections while multiple failures stay hidden until a demand on the system, the recorded counts understate the higher multiplicities.

Table 14.9: Illustrative multiplicity counts of 100 failure events in a group of four pumps; maximum likelihood estimates with their standard errors $\sqrt{\hat{\alpha}_k(1-\hat{\alpha}_k)/n}$; and the posterior under a uniform Dirichlet$(1,1,1,1)$ prior, with its mean and 90 % credible interval.

Components affected k	events n_k	$\hat{\alpha}_k$	std. error	posterior parameter	posterior mean	90 % interval
1	75	0.750	0.043	76	0.731	(0.657, 0.799)
2	15	0.150	0.036	16	0.154	(0.100, 0.215)
3	7	0.070	0.026	8	0.077	(0.039, 0.124)
4	3	0.030	0.017	4	0.038	(0.013, 0.074)
total	100	1		104	1	

The estimates are uncertain, and most uncertain for the high multiplicities that matter most. Only three of the 100 events affected all four pumps; $\hat{\alpha}_4 = 0.03$ is a sparse count, and its standard error, $\sqrt{0.03 \times 0.97/100} = 0.017$, is more than half the estimate. A Bayesian treatment, as in Chapter 8, represents this uncertainty by a distribution. The natural prior for a vector of nonnegative fractions that sum to one is the Dirichlet distribution,

$$\boldsymbol{\alpha} \sim \text{Dirichlet}(\gamma_1, \ldots, \gamma_m), \qquad p(\boldsymbol{\alpha}) \propto \prod_{k=1}^{m} \alpha_k^{\gamma_k - 1},$$

with positive parameters γ_k. Multiplying by the multinomial likelihood $\prod_k \alpha_k^{n_k}$ adds the counts to the exponents, so the posterior given the data D is again Dirichlet,

$$\boldsymbol{\alpha} \mid D \sim \text{Dirichlet}(\gamma_1 + n_1, \ldots, \gamma_m + n_m), \tag{14.32}$$

with posterior mean

$$\mathbb{E}[\alpha_k \mid D] = \frac{\gamma_k + n_k}{\sum_j (\gamma_j + n_j)}. \tag{14.33}$$

Each marginal of a Dirichlet distribution is a beta distribution, $\alpha_k \mid D \sim \text{Beta}(\gamma_k + n_k, \sum_j (\gamma_j + n_j) - \gamma_k - n_k)$, from which credible intervals follow.

With the uniform prior Dirichlet$(1,1,1,1)$ and the counts of Table 14.9, the posterior is Dirichlet$(76, 16, 8, 4)$ and the posterior means are $76/104 = 0.731$, $16/104 = 0.154$, $8/104 = 0.077$ and $4/104 = 0.038$: the prior pulls each estimate slightly toward $1/4$, most visibly for the rare multiplicities. Figure 14.7 draws the posterior marginals of α_2, α_3 and α_4. The horizontal axis

is the value of the alpha factor; the vertical axis is the posterior density; the short vertical ticks at the bottom mark the maximum likelihood estimates. The density of α_4 is concentrated below 0.1 but spreads over a factor of about five, between 0.013 and 0.074 at 90 % credibility. The calculation still assumes complete, equally observable and correctly classified events from one homogeneous population; incomplete observation, testing differences, uncertain classification and heterogeneity between plants require a more elaborate likelihood.

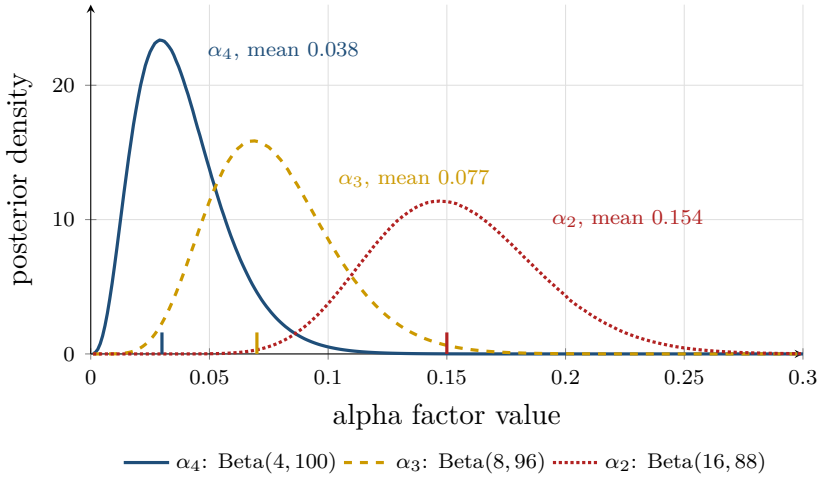


Figure 14.7: Posterior marginal densities of α_2, α_3 and α_4 for the counts of Table 14.9 under a uniform Dirichlet prior: $\alpha_k \mid D \sim \text{Beta}(1+n_k, 104-1-n_k)$, with density $x^{a-1}(1-x)^{b-1}/B(a,b)$ (equation (14.32)); illustrative counts. Ticks at the bottom mark the maximum likelihood estimates 0.15, 0.07 and 0.03.

14.6.3 Beta-factor estimation

The beta factor is a share of component failures, not of events, and the estimator must count accordingly. Consider a two-component group with the counts of Example A: 90 single-component failure events and 10 double-component events. There are 100 failure events but

$$90(1) + 10(2) = 110$$

affected-component failures, of which 20 belong to double events. The fraction of component failures attributable to events that disable the whole group is

$$\hat{\beta} = \frac{10(2)}{110} = 0.1818, \tag{14.34}$$

while the fraction of group events that affect both components is $\hat{\alpha}_2 = 10/100 = 0.10$. Hence $\hat{\beta} \neq \hat{\alpha}_2$; the difference is the multiplicity weighting, and the two estimates are related by equation (14.30). Dividing 10 by 100 estimates an alpha-type quantity, not the beta factor. For a group of m components with only single and whole-group events the estimator is $\hat{\beta} = mn_m/(n_1 + mn_m)$. A real estimation must also account for the operating exposure or the number of demands behind the counts, the group definition, the failure modes included, and the observation process, as for the alpha factors.

14.6.4 Staggered against non-staggered testing

Components of a standby group can be tested together or individually at different times. Consider three standby pumps, each periodically tested to see whether it starts on demand.

Under non-staggered testing all three are tested at roughly the same time, and a common latent defect that appears between tests remains undetected in all of them until that scheduled test. Under staggered testing they are tested at evenly spaced times, and a shared latent defect is found at the next test of any of them, after which the others are checked. The arrangement affects how long multiple failures remain undetected, hence the effective unavailability of the group, and how the failure records must be read when the parameters are estimated.

The effect on the latent exposure can be quantified with the standby model of equation (14.3). With a test interval τ, a whole-group event occurring at a random time waits on average $\tau/2$ for detection under non-staggered testing, and on average $\tau/(2m)$ under staggered testing, since some component is tested every τ/m. The time-averaged unavailability contributed by that event is about $\lambda_{\mathrm{C}}\tau/2$ and $\lambda_{\mathrm{C}}\tau/(2m)$ respectively. For $\lambda_{\mathrm{C}} = 10^{-5}\,\mathrm{h}^{-1}$, a monthly test ($\tau = 720\,\mathrm{h}$) and four pumps, these are 3.6×10^{-3} and 9.0×10^{-4} (illustrative values). Testing can therefore change the system unavailability even when it does not change the physical rate of the shared defect. This matters for standby safety equipment, emergency generators and protective systems.

On the estimation side, the conventions lead to the two basic-event expressions of Section 14.4.3, equations (14.23) and (14.24). Table 14.10 evaluates both for the alpha factors of Example C with an illustrative $Q_{\mathrm{T}} = 10^{-3}$ per demand. The ratio of the two is $k/\bar{k}$: the non-staggered expression gives a smaller independent term and larger common-cause terms, by a factor $4/1.38 = 2.90$ for the whole-group event. The table holds the alpha factors and Q_{T} fixed only to display the algebra; in practice each set of inputs is estimated under its own convention, and the table is not a statement about which testing scheme is safer. The expressions require alpha factors and Q_{T} estimated or defined consistently with the testing convention, and neither may be used where the other's conventions hold.

Table 14.10: Basic-event probabilities per demand of one specified subset of k components in a group of four, from the alpha factors of Example C with $\bar{k} = 1.38$ and $Q_{\mathrm{T}} = 10^{-3}$ (illustrative), under the non-staggered and staggered expressions (equations (14.23) and (14.24)). The last row checks $\sum_k \binom{3}{k-1} Q_k^{(4)} = Q_{\mathrm{T}}$.

k	α_k	$\binom{3}{k-1}$	non-staggered $Q_{k,\mathrm{NS}}^{(4)}$	staggered $Q_{k,\mathrm{S}}^{(4)}$	ratio $k/\bar{k}$
1	0.75	1	5.4348×10^{-4}	7.5000×10^{-4}	0.725
2	0.15	3	7.2464×10^{-5}	5.0000×10^{-5}	1.449
3	0.07	3	5.0725×10^{-5}	2.3333×10^{-5}	2.174
4	0.03	1	8.6957×10^{-5}	3.0000×10^{-5}	2.899
$\sum_k \binom{3}{k-1} Q_k^{(4)}$			1.0000×10^{-3}	1.0000×10^{-3}	

14.6.5 Model selection in practice

The beta-factor model is a useful starting point when data are limited and the group-wide failure mechanism is the principal concern. The alpha-factor model becomes more valuable when common-cause events of intermediate multiplicity can materially affect system success, as in a 2-out-of-3 or 2-out-of-4 arrangement. Neither model accounts for everything. A more physical model is required when failure probabilities depend on common weather, shared electrical loading, temperature, maintenance history or the propagation of one failure to another. Table 14.11 pairs typical situations with a useful starting approach.

A beta or alpha factor is a compact summary of dependent failure behavior, and it does not explain the physical mechanism. An alpha-factor model may estimate that 3 % of group

Table 14.11: Starting approaches to dependent failures by situation.

Situation	Useful starting approach
Two redundant components, little CCF data	beta-factor model
Three or more components where partial-group failures matter	alpha-factor model
Known shared electrical or control dependency	explicit shared-event fault tree
Failures depend on common weather or loading conditions	conditional physical or statistical model
One failure increases the hazards of the survivors	dynamic reliability or load-sharing model
Several possible causes with uncertain structure	Bayesian model comparison or sensitivity analysis

failure events affect all four chillers of a plant without revealing whether those events arise from electrical transients, contamination of the shared cooling water, maintenance errors or a common software defect. The distinction matters when risk-reduction measures are selected: when physical information about an important shared cause is available, it is better to model that cause explicitly than to rely on an aggregate factor.

Consider four redundant cooling pumps whose failure probability is dominated by the whole-group event, as in Table 14.3. One option installs a fifth identical pump that shares the same electrical and control infrastructure; under the beta-factor model with the same parameters, the five-pump group fails over 100 h with probability 9.995002×10^{-4} against 9.995066×10^{-4} for four, an improvement of six parts in a million, because the shared event still disables every pump. The other option separates the pumps into two independently powered and controlled groups of two. If, and only if, the redesign removes every coupling between the groups, each group fails with the two-pump probability 1.0797×10^{-3} and the system with $(1.0797 \times 10^{-3})^2 = 1.17 \times 10^{-6}$; any coupling that remains between the groups restores a floor at its own rate. The benefit of the second option cannot be claimed merely by assuming smaller beta or alpha factors: it requires justified new event rates, group definitions and dependencies. In the same way, a four-chiller data center may gain more from independent cooling-water supplies than from a fifth chiller, depending on the relative probabilities of independent failures, partial common-cause events and complete common-cause events. That is a problem of model comparison and decision, and the correct design follows from the modified system model, not from a parameter label alone.

The models of this chapter quantify how a system fails; the next chapter turns to what follows an initiating event, through event trees and accident sequences.

14.7 What to retain

Definition 14.1 (Beta-factor and alpha-factor models)**.** The beta factor $\beta = \lambda_C/\lambda_T$ is the share of a component's failure rate due to events that disable its whole common-cause group; the alpha factor α_k is the fraction of the group's failure events that disable exactly k components. Beta is simpler but assumes that every common-cause event is group-wide; alpha factors describe partial-group multiplicities. Both require clear event definitions, consistent failure-rate or probability conventions, and data that represent the mechanisms under investigation, and neither is a system failure probability until it is combined with event rates, the mission or test interval, and the system's success logic.

Rates and probabilities must not be mixed: a partition that is exact between rates is only a

rare-event approximation between mission probabilities. Event fractions and component-failure fractions differ by the multiplicity weight. Adding identical components cannot lower the group failure probability below the common-cause floor. Testing and observation shape both the effective unavailability and the estimates, and sparse multiple-failure data call for uncertainty analysis, not false precision. Table 14.12 lists the chapter's terms, their meaning and units, and the subsection that defines each.

Table 14.12: The terms of this chapter and where each is defined.

Term	Meaning	Expression	Units	Defined in
Total component failure rate	hazard of one component from all events	$\lambda_{\mathrm{T}} = \lambda_{\mathrm{I}} + \lambda_{\mathrm{C}}$	h^{-1}	14.2.1
Group-event frequency	rate of events of multiplicity k	$r_k = r\alpha_k$	h^{-1}	14.2.1
Failure probability per demand	failure when called on	q_{d}	–	14.2.2
Snapshot unavailability	failed at an instant	$\bar{q}_{\mathrm{U}} \approx \lambda\tau/2$	–	14.2.2
Mission failure probability	failure by t	$Q(t) = 1 - \mathrm{e}^{-\lambda t}$	–	14.2.3
Beta factor	common-cause share of the rate	$\beta = \lambda_{\mathrm{C}}/\lambda_{\mathrm{T}}$	–	14.3.1
Beta-factor group failure	one of m required	equation (14.10)	–	14.3.3
Common-cause floor	limit of many identical units	$1 - \mathrm{e}^{-\lambda_{\mathrm{C}} t}$	–	14.3.4
Alpha factor	fraction of events of multiplicity k	α_k, $\sum_k \alpha_k = 1$	–	14.4.1
Mean multiplicity	components per event	$\bar{k} = \sum_k k\alpha_k$	–	14.4.2
Group-event rate	all events of the group	$r = m\lambda_{\mathrm{T}}/\bar{k}$	h^{-1}	14.4.2
Specific-subset rate	one subset of size k	$r_k/\binom{m}{k}$	h^{-1}	14.4.2
Basic-event probability	subset of k per demand	equations (14.23), (14.24)	–	14.4.3
Shock-model survival	set of s survives	$\mathrm{e}^{-h_s t}$	–	14.4.4
Equivalent beta	beta of a restricted alpha model	$m\alpha_m/(\alpha_1 + m\alpha_m)$	–	14.5.2
Alpha estimator	observed event fraction	$\hat{\alpha}_k = n_k/\sum_j n_j$	–	14.6.2
Dirichlet posterior	uncertainty of the alpha factors	Dirichlet$(\gamma_k + n_k)$	–	14.6.2
Beta estimator	share of component failures	$mn_m/(n_1 + mn_m)$	–	14.6.3
Staggered testing	components tested at offset times	latent exposure $\tau/(2m)$	–	14.6.4

Further reading. Mosleh, Rasmuson and Marshall, *Guidelines on Modeling Common-Cause Failures in Probabilistic Risk Assessment*, NUREG/CR-5485, the foundational reference for the beta, alpha and other common-cause parameter models; IAEA-TECDOC-648, *Procedures for Conducting Common Cause Failure Analysis in Probabilistic Safety Assessment*, for implementation, parameter estimation and testing assumptions; NRC/INL, *Causal CCF Parameter Estimations*, for the established formulas and their conventions; IAEA-TECDOC-1922, *Common Cause Failures in Digital Instrumentation and Control Systems*, for common-cause methods and terminology; NASA, *Common-Cause Failure Modeling with Alpha Factors*, for practical alpha-factor calculation formulas.

14.8 Exercises

Exercise 14.1. (a) In a rate-based beta-factor model, state what $\beta = 0.1$ represents. (b) In a four-component alpha-factor model, state what α_3 represents. (c) Explain why the conversion from alpha factors to a beta factor requires multiplicity weighting. (d) State what limits the effectiveness of adding identical redundant components in a beta-factor model. (e) State when a richer alpha-factor model is useful. (f) Explain why the testing arrangement matters for alpha-factor unavailability calculations. (Answers: (a) that 10 % of the modeled per-component failure rate comes from group-wide common-cause events, not a 10 % probability of losing all components or a correlation of 0.1; (b) the fraction of failure events that affect exactly three components; (c) because a double failure counts as two affected-component failures but one group event; (d) the group-wide common-cause process, which sets the floor $1 - \mathrm{e}^{-\lambda_C t}$; (e) when events disabling an intermediate number of components materially affect system failure; (f) because testing determines when latent multiple failures are discovered and how long they persist, and the basic-event expressions differ between staggered and non-staggered testing.)

Exercise 14.2. A group of standby cooling pumps has $\lambda_T = 2 \times 10^{-4}\,\mathrm{h}^{-1}$ per pump and $\beta = 0.05$; one pump is required over a mission of 500 h. (a) Compute λ_I and λ_C. (b) Compute the group failure probability for two and for three pumps by equation (14.10), and the independent-only values with the same λ_T. (c) Compute the common-cause floor. (d) Compare $Q_C(500\,\mathrm{h})$ with $\beta\, Q_T(500\,\mathrm{h})$. (Answers: (a) $1.9 \times 10^{-4}\,\mathrm{h}^{-1}$ and $10^{-5}\,\mathrm{h}^{-1}$; (b) 0.013160 and 0.0057282, against $(1-\mathrm{e}^{-0.1})^2 = 0.0090559$ and 0.00086178; (c) $1-\mathrm{e}^{-0.005} = 0.0049875$; (d) $Q_C = 0.0049875$ against $0.05 \times 0.095163 = 0.0047581$, so the product underestimates by 4.6 %.)

Exercise 14.3. The chiller plant of Example B is re-estimated as $(\alpha_1, \alpha_2, \alpha_3) = (0.90, 0.07, 0.03)$ with $\lambda_T = 10^{-4}\,\mathrm{h}^{-1}$; two of three chillers are required over 100 h. (a) Compute $\bar{k}$, r and r_1, r_2, r_3. (b) Compute the rate of one particular subset of each size and check that the rates reaching one chiller add up to λ_T. (c) Compute the direct approximation $(r_2 + r_3)t$ and the rare-event sum over the minimal cut sets, and compare with the exact value 0.0028386 from equation (14.27). (d) For a two-component group with $\alpha_1 = 0.8$ and $\alpha_2 = 0.2$, compute the equivalent beta factor. (Answers: (a) $\bar{k} = 1.13$, $r = 2.6549 \times 10^{-4}\,\mathrm{h}^{-1}$, $r_k = 2.3894 \times 10^{-4}$, 1.8584×10^{-5} and $7.9646 \times 10^{-6}\,\mathrm{h}^{-1}$; (b) 7.9646×10^{-5}, 6.1947×10^{-6} and $7.9646 \times 10^{-6}\,\mathrm{h}^{-1}$, and $7.9646 \times 10^{-5} + 2 \times 6.1947 \times 10^{-6} + 7.9646 \times 10^{-6} = 1.0000 \times 10^{-4}\,\mathrm{h}^{-1}$; (c) 0.0026549, and $3 \times 0.00061928 + 0.00079614 + 3 \times 0.0079330^2 = 0.0028428$, against the exact 0.0028386; (d) $\beta = 0.4/1.2 = 1/3$, not 0.2.)

Exercise 14.4. A group of three standby pumps has recorded 48 failure events: 40 affecting one pump, 6 affecting two and 2 affecting all three. (a) Compute the maximum likelihood estimates of the alpha factors and the standard error of $\hat{\alpha}_3$. (b) With a uniform Dirichlet$(1,1,1)$ prior, give the posterior, the posterior means, and the 90 % credible interval of α_3, whose marginal is Beta$(3, 48)$. (c) With $Q_T = 2 \times 10^{-3}$ per demand and the estimates of (a), compute the basic-event probability of one specified subset of each size under the non-staggered and the staggered expressions. (d) In a separate two-pump group with 46 single and 4 double events, compute $\hat{\beta}$ and $\hat{\alpha}_2$. (Answers: (a) $(0.8333, 0.1250, 0.0417)$, standard error 0.029; (b) Dirichlet$(41, 7, 3)$, means $(0.8039, 0.1373, 0.0588)$, interval $(0.017, 0.121)$; (c) with $\bar{k} = 1.2083$, non-staggered 1.379×10^{-3}, 2.069×10^{-4} and 2.069×10^{-4}, staggered 1.667×10^{-3}, 1.250×10^{-4} and 8.333×10^{-5}; (d) $\hat{\beta} = 8/54 = 0.148$ and $\hat{\alpha}_2 = 4/50 = 0.080$.)

Chapter 15

Event Trees and Accident Sequences

15.1 Event trees and the progression of an initiating event

Chapters 10 to 14 studied fault trees, minimal cut sets, importance measures and common-cause failures. Those methods answer one question: which combinations of events can cause a defined system failure. Event tree analysis starts from the other end. It takes an event that has already occurred, a loss of utility supply or the trip of a transmission line, and asks what can happen next. A loss of utility supply at a data center does not by itself shut the servers down: the uninterruptible power supply can carry the load, the standby generator can start, and operators can restore the supply. A transmission-line trip does not by itself shed load: protection, redispatch of generation and emergency control schemes can keep the disturbance from becoming an interruption of service. The question of this chapter is

once an initiating event occurs, how the success or failure of protection functions, operator actions and recovery determines whether the event remains a minor disturbance or develops into a serious system failure, and how often each outcome is expected.

An event tree organizes the possible progressions as mutually exclusive sequences of successes and failures of those functions. Because each sequence begins with an initiating event that has a frequency and ends in an outcome that has a consequence, the event tree is the step of a probabilistic risk assessment that connects initiating-event frequencies to operational consequences.

Jargon: initiating event, event tree, heading, branch, accident sequence, end state, success criterion

An *initiating event* is an event that disturbs the normal operation of a system and requires a response from its protection or safety functions; it begins the sequence being analyzed. An *event tree* is an inductive logic diagram that starts from one initiating event and develops, from left to right, the possible responses of the system; *event tree analysis* (ETA) is its construction and quantification. Each column of the tree is a *heading* (also called a function event): a protection function, safety system, barrier or operator action whose success or failure is asked at that point. At each heading the path divides into *branches*, conventionally success upward and failure downward. An *accident sequence*, or simply sequence, is one path from the initiating event through a specified outcome at each heading it crosses. The *end state* is the outcome category at the end of a sequence (service maintained, brief interruption, load shedding). The *success criterion* of a heading is the measurable condition that defines its success, such as "emergency power supplies the IT load within 10 s and for the whole outage". A *probabilistic risk assessment* (PRA) combines initiating events, event trees, fault trees and consequence models into a quantitative statement of risk.

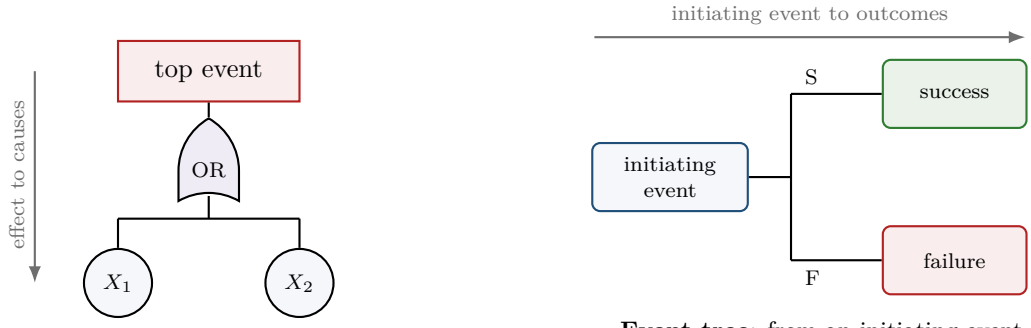
Key idea

An event tree partitions what can follow one initiating event into mutually exclusive sequences. The probability of each sequence is the product of branch probabilities conditioned on everything that happened before; multiplied by the initiating-event frequency it becomes an expected number of occurrences per year, which can then be weighted by the sequence's consequence.

This section contrasts the event tree with the fault tree and defines the quantity that every event tree starts from, the frequency of its initiating event.

15.1.1 Fault trees and event trees

Figure 15.1 sets the two diagrams side by side. On the left is a fault tree: the box at the top is the top event, the undesirable result; the gate beneath it is an OR gate; the circles are two basic events X_1 and X_2, either of which produces the top event. The analysis reads from the top down, from a defined failure back to the combinations of causes that produce it, which makes it deductive. On the right is an event tree: the box on the left is the initiating event; the line divides at one heading into a success branch (upper box) and a failure branch (lower box). The analysis reads from left to right, from a disturbance forward to its possible outcomes, which makes it inductive.



Fault tree: from an undesirable result backward to its causes (deductive)

Event tree: from an initiating event forward to its possible outcomes (inductive)

Figure 15.1: The fault tree (left) reasons backward from a defined top event through logic gates to basic events; the event tree (right) reasons forward from an initiating event through the success (S) or failure (F) of a function to its outcomes. The two are complementary parts of a probabilistic risk assessment.

The two methods complement one another. A fault tree can compute the probability that a standby generator fails to supply emergency power; an event tree can use that probability as one of its branches to evaluate the consequences that follow a loss of utility supply. Section 15.5 develops the link. Together, initiating events, event trees, fault trees and consequence models are the structure of a probabilistic risk assessment: the fault trees explain why a function fails, and the event tree explains what that failure does to the progression of the initiating event.

15.1.2 The initiating event: frequency and probability

An initiating event is the event that begins the sequence being analyzed. Typical initiating events in the systems of this book are the loss of utility electrical supply, the trip of a transmission line, the failure of a cooling-water pump, the loss of chilled-water supply, a battery thermal alarm, and the failure of a process-control system. Each one is defined precisely enough that its

occurrences can be counted in the operating records, and each one is analyzed by its own event tree, or by one tree shared by a group of initiating events that demand the same response.

Two quantities describe how often an initiating event I occurs, and they must be kept apart. Let $N_I(H)$ be the number of occurrences in a period of length H, measured in calendar years. The *initiating-event probability* over the period is

$$\Pr(N_I(H) \geq 1), \tag{15.1}$$

the probability that I occurs at least once during $[0, H]$; it is dimensionless and has a meaning only with its period stated. The *initiating-event frequency* is the expected number of occurrences per unit time,

$$f_{\mathrm{IE}} = \frac{\mathbb{E}[N_I(H)]}{H}, \tag{15.2}$$

in events per year, yr^{-1}; for a stationary process the ratio does not depend on H. It is estimated from operating records as the number of occurrences observed divided by the exposure, the total time over which they could have occurred: four losses of utility supply in ten years of records give $f_{\mathrm{IE}} = 0.4\,\mathrm{yr}^{-1}$. Estimation from sparse records is the subject of Chapter 7. A frequency of $f_{\mathrm{IE}} = 0.20\,\mathrm{yr}^{-1}$ means an average of 0.20 initiating events per year, or one every five years on average. It does not mean that an event occurs every five years.

The relation between the two quantities needs a model of the occurrence times. The standard one is the homogeneous Poisson process: occurrences are independent of one another and happen at a constant rate f_{IE}, so that the probability of one occurrence in a short interval Δt is $f_{\mathrm{IE}}\,\Delta t$ to first order. Let $P_0(t)$ be the probability of no occurrence in $[0, t]$. No occurrence in $[0, t + \Delta t]$ requires none in $[0, t]$ and none in the next Δt, so $P_0(t + \Delta t) = P_0(t)\,(1 - f_{\mathrm{IE}}\,\Delta t)$ to first order; letting $\Delta t \to 0$ gives $\mathrm{d}P_0/\mathrm{d}t = -f_{\mathrm{IE}}\,P_0$ with $P_0(0) = 1$, whose solution is $P_0(t) = \mathrm{e}^{-f_{\mathrm{IE}} t}$. This is the derivation of the exponential reliability function, equation (3.4), with the initiating event in place of the failure: the time to the first initiating event is exponential with mean $1/f_{\mathrm{IE}}$. The probability of at least one occurrence in $[0, H]$ is the complement,

$$\Pr(N_I(H) \geq 1) = 1 - \mathrm{e}^{-f_{\mathrm{IE}} H}, \tag{15.3}$$

and the number of occurrences has the Poisson distribution $\Pr(N_I(H) = n) = (f_{\mathrm{IE}} H)^n\,\mathrm{e}^{-f_{\mathrm{IE}} H}/n!$, whose mean is $f_{\mathrm{IE}} H$, in agreement with equation (15.2).

For $f_{\mathrm{IE}} = 0.20\,\mathrm{yr}^{-1}$ and $H = 1\,\mathrm{yr}$,

$$\Pr(N_I(1\,\mathrm{yr}) \geq 1) = 1 - \mathrm{e}^{-0.20} = 1 - 0.818731 = 0.181269,$$

or 18.13 %, not 20 %. The difference is the probability mass of years with two or more events: $\Pr(N_I = 1) = 0.20\,\mathrm{e}^{-0.20} = 0.163746$ and $\Pr(N_I \geq 2) = 1 - 0.818731 - 0.163746 = 0.017523$. The expected number counts every occurrence, the probability only whether there was at least one; in general $\mathbb{E}[N_I] = \sum_{n \geq 1} \Pr(N_I \geq n) \geq \Pr(N_I \geq 1)$. Over five years the probability of at least one event is $1 - \mathrm{e}^{-1} = 0.6321$, and the probability of a five-year period with none is $\mathrm{e}^{-1} = 0.3679$, although one event is expected.

Because $\mathrm{e}^{-x} \geq 1 - x$ for every $x \geq 0$, equation (15.3) gives

$$\Pr(N_I(H) \geq 1) \leq f_{\mathrm{IE}} H, \qquad \Pr(N_I(H) \geq 1) \approx f_{\mathrm{IE}} H \quad \text{when } f_{\mathrm{IE}} H \ll 1, \tag{15.4}$$

with an error of about $(f_{\mathrm{IE}} H)^2/2$. The approximation is conservative. At $f_{\mathrm{IE}} H = 0.20$ it overstates the probability by 0.0187, about 10 %; at $f_{\mathrm{IE}} H = 0.01$ it gives 0.01 against the exact 0.00995, a difference of 0.5 %. The distinction between an annual frequency and an annual probability runs through the rest of the chapter.

Common mistake

Reading the reciprocal of a frequency as a recurrence interval. A frequency of $0.02\,\mathrm{yr}^{-1}$ does not mean one event every 50 years: under the Poisson model the probability of at least one occurrence in 50 years is $1 - \mathrm{e}^{-1} = 0.63$, and two or more occurrences in that period are possible.

15.2 Event-tree branching and sequence probabilities

This section builds event trees from one heading to several, derives the probability of a sequence as a product of conditional branch probabilities, states when the branches may be treated as independent, and proves that the sequence probabilities of a complete tree sum to one.

15.2.1 The simplest event tree

Suppose the utility supply is lost and one standby generator must take the load. Let G be the event that the generator starts and fulfills its required function, and $\bar{G}$ its complement. Given the initiating event, the tree has two branches, G and $\bar{G}$, with probabilities

$$\Pr(G \mid I) = p_G, \qquad \Pr(\bar{G} \mid I) = 1 - p_G, \qquad \Pr(G \mid I) + \Pr(\bar{G} \mid I) = 1. \tag{15.5}$$

The last equality is the fundamental branching rule: the branches at a heading are mutually exclusive and exhaust the possible outcomes of that heading. A branch probability is a probability per demand, conditional on the initiating event having occurred; it is not a rate. For the generator it combines the probability of failing to start on demand with the probability of failing to run for the duration of the outage, both judged against the success criterion of the heading.

If the generator succeeds with probability $\Pr(G \mid I) = 0.95$, then $\Pr(\bar{G} \mid I) = 0.05$. With $f_{\mathrm{IE}} = 0.20\,\mathrm{yr}^{-1}$, the expected frequency of the sequence "utility supply lost, generator fails" is

$$f_{I\bar{G}} = f_{\mathrm{IE}} \Pr(\bar{G} \mid I) = 0.20 \times 0.05 = 0.01\,\mathrm{yr}^{-1}.$$

The multiplication follows from an expected count. Number the initiating events of a period $1, 2, \ldots, N_I(H)$ and suppose that each one, independently of the others and of how many there are, is followed by a generator failure with probability $\Pr(\bar{G} \mid I)$. The number of generator-failure sequences is the sum of $N_I(H)$ indicators, and its expectation is $\mathbb{E}[N_I(H)]\,\Pr(\bar{G} \mid I) = f_{\mathrm{IE}}\,H\,\Pr(\bar{G} \mid I)$; dividing by H gives the frequency. Under the Poisson model the generator-failure sequences themselves form a Poisson process with rate $0.01\,\mathrm{yr}^{-1}$ (a standard property, the thinning of a Poisson process), so the probability of at least one such sequence in a year is $1 - \mathrm{e}^{-0.01} = 0.00995$. The value $0.01\,\mathrm{yr}^{-1}$ is an expected occurrence rate, one sequence per hundred years on average, and not a prediction that the sequence occurs exactly once in a century nor the probability that it occurs in a given year.

15.2.2 Multiple protection functions

A data center usually has several protection functions against the loss of utility supply: an uninterruptible power supply (UPS), which carries the IT load from its batteries the instant the supply is lost; a standby generator, which establishes sustained emergency power after it starts and its transfer switch closes; and a load-management or graceful shutdown function, which sheds noncritical load or shuts the IT equipment down in an orderly way when the first two are insufficient. After the initiating event these functions succeed or fail in different combinations, and the event tree enumerates the combinations in the order in which the functions are demanded.

Figure 15.2 is a simplified tree with the first two functions as headings. The box on the left is the initiating event; the column headings across the top are the UPS and the generator; at each heading the path divides into a success branch (S, upward) and a failure branch (F, downward); the boxes on the right are the four end states, green where the IT supply is maintained and red where it is not. Reading the paths from the top: both functions succeed and power is maintained; the UPS succeeds and the generator fails, so the IT load rides through on the batteries; the UPS fails and the generator succeeds, so the supply is interrupted until the generator takes the load; both fail and the supply is lost. A third heading for graceful shutdown would divide each red end state again into an orderly and a disorderly shutdown.

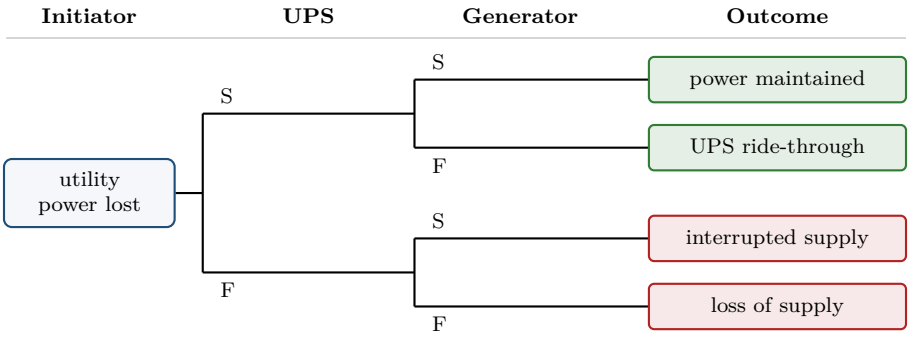


Figure 15.2: Conceptual event tree for the loss of utility power at a data center, with two headings and four sequences. The end-state labels need precise success criteria: UPS ride-through lasts only as long as the stored energy, and it is not sustained service unless the supply returns in time.

The tree is organized by the success or failure of functions, not of components. The second end state shows why the success criteria matter: "UPS ride-through" is green only for as long as the battery energy lasts, and whether it counts as a success depends on whether the utility supply or some other source returns before the batteries are exhausted. The complete tree of Section 15.3 adds that recovery as a third heading.

A second detail is that the probability of a later function can depend on the outcome of an earlier one. Generator success after the UPS has failed can have a different probability of preserving uninterrupted IT power than generator success after the UPS has operated, because the electrical conditions the generator meets, and the meaning of its success, differ. A rigorous event tree therefore uses conditional probabilities at every branch.

15.2.3 The general sequence probability

Suppose the initiating event I is followed by headings $1, 2, \ldots, n$, and let B_j denote the specified outcome, success or failure, at heading j along one particular sequence S. The sequence is the intersection $S = I \cap B_1 \cap B_2 \cap \cdots \cap B_n$, and its probability conditional on the initiating event is

$$\Pr(S \mid I) = \Pr(B_1, \ldots, B_n \mid I) = \prod_{j=1}^{n} \Pr(B_j \mid I, B_1, \ldots, B_{j-1}), \tag{15.6}$$

where the $j = 1$ factor is $\Pr(B_1 \mid I)$. This is the chain rule of probability. It follows from the definition of conditional probability, $\Pr(A \mid C) = \Pr(A \cap C)/\Pr(C)$, applied to each factor:

$$\prod_{j=1}^{n} \Pr(B_j \mid I, B_1, \ldots, B_{j-1}) = \prod_{j=1}^{n} \frac{\Pr(I \cap B_1 \cap \cdots \cap B_j)}{\Pr(I \cap B_1 \cap \cdots \cap B_{j-1})} = \frac{\Pr(I \cap B_1 \cap \cdots \cap B_n)}{\Pr(I)},$$

because the product telescopes: the numerator of each factor cancels the denominator of the next, leaving the last numerator over the first denominator, which is $\Pr(B_1, \ldots, B_n \mid I)$. The

derivation uses no independence assumption; each factor is conditioned on the initiating event and on the full history of the sequence up to that heading. For three headings,

$$\Pr(B_1, B_2, B_3 \mid I) = \Pr(B_1 \mid I) \times \Pr(B_2 \mid I, B_1) \times \Pr(B_3 \mid I, B_1, B_2).$$

In a PRA the initiating event is described by a frequency, not a probability; conditioning on I then means conditioning on "an initiating event of this type has just occurred", and the branch probabilities are probabilities per occurrence.

Write $S_1, S_2, \ldots$ for the sequences of a tree. The *sequence frequency* is the expected number of occurrences of sequence S_k per year,

$$f_{S_k} = f_{\mathrm{IE}} \Pr(S_k \mid I), \tag{15.7}$$

in yr^{-1}. The expected-count argument of Section 15.2.1 proves it under two conditions, which make the initiating-event rate and the conditional sequence model compatible: every occurrence of I is followed by the same conditional branch probabilities, and the sequence followed after one occurrence is independent of the other occurrences and of their number. The first fails if the conditional probabilities depend on the season, the load or the weather at the time of the event; the remedy is to divide the initiating event into families by operating condition, each with its own frequency and its own branch probabilities. The second fails if one event leaves equipment damaged and unrepaired when the next one arrives. (The sequence S_k is an event; it is unrelated to the unit status $S_i(t)$ of the notation table.)

15.2.4 The independent-branch approximation

A common simplification replaces every conditional branch probability by one conditioned on the initiating event alone,

$$\Pr(B_j \mid I, B_1, \ldots, B_{j-1}) = \Pr(B_j \mid I), \qquad \text{so that} \qquad \Pr(S \mid I) = \prod_{j=1}^{n} \Pr(B_j \mid I). \tag{15.8}$$

The assumption is that the outcomes of the headings are conditionally independent given the initiating event: knowing that the UPS failed says nothing new about whether the generator will succeed. The product then needs only one number per heading instead of one per heading and history.

The approximation fails when the headings share a support system, a control system, an environmental condition or an operator. A UPS and a generator transfer that rely on the same battery-backed control system, a generator and its fuel pumps that rely on the same switchboard, or two operator actions taken by the same crew under the same time pressure are dependent, and the product of marginal probabilities can then be wrong by a large factor in either direction. Chapter 13 showed why conditional dependence must be modeled explicitly; Section 15.5.2 returns to it for headings quantified by fault trees.

15.2.5 Completeness of the sequence set

A complete event tree partitions the outcomes that can follow the initiating event. Completeness has two requirements: at every branch point the branches are mutually exclusive and exhaust the outcomes of the heading given the history, so that for a binary heading

$$\Pr(\text{success} \mid \text{history}) + \Pr(\text{failure} \mid \text{history}) = 1; \tag{15.9}$$

and every branch is followed to an end state. A branch may skip later headings, as when a heading is not demanded on that path, and still end in an end state.

Principle 15.1. *In a complete event tree the conditional probabilities of the terminal sequences sum to one,*

$$\sum_k \Pr(S_k \mid I) = 1, \qquad \textit{and therefore} \qquad \sum_k f_{S_k} = f_{\mathrm{IE}}. \tag{15.10}$$

The proof is by induction on the number of headings a path can cross. A tree with no heading has one sequence, I itself, with conditional probability 1. Suppose the statement holds for every complete tree of depth at most $n-1$, and take a complete tree of depth n. Group its terminal sequences by the branch b they take at the first heading. Below branch b the tree is a complete tree of depth at most $n-1$, quantified conditionally on $I \cap b$, so by the induction hypothesis its sequence probabilities, conditional on $I \cap b$, sum to one. By the chain rule (15.6), each sequence below b has $\Pr(S_k \mid I) = \Pr(b \mid I)\ \Pr(S_k \mid I, b)$, so the sequences below b contribute $\Pr(b \mid I) \times 1$. Summing over the branches of the first heading gives $\sum_b \Pr(b \mid I) = 1$ by equation (15.9). The second statement follows by multiplying the first by f_{IE}.

The result is a validation check. If the sequence probabilities of a tree sum to 0.93 or 1.07, the tree is incomplete, internally inconsistent (two branches overlap, or a branch probability was taken from the wrong history), or incorrectly quantified. Event-tree software applies this normalization as one of its basic consistency checks, and the same check on the frequencies, which must add up to the initiating-event frequency, is applied to every quantified tree in this chapter.

15.3 Worked example: data center power continuity

This section builds and quantifies a complete event tree for the loss of utility supply at a data center, from its branch probabilities to sequence frequencies, consequence categories, expected annual interruption hours and expected annual financial loss. All numbers are illustrative.

15.3.1 The scenario and its branch probabilities

A data center loses its normal utility supply with frequency $f_{\mathrm{IE}} = 0.40\,\mathrm{yr}^{-1}$: on average, 0.4 utility-supply losses per year. The model has three headings, demanded in this order:

- *UPS*, event U: maintains IT electrical power immediately after the loss of utility supply;
- *generator*, event G: establishes sustained emergency power;
- *timely recovery*, event REC: if the generator fails, restores acceptable power before the stored energy of the UPS is exhausted.

Table 15.1 gives the conditional branch probabilities. Each failure branch has the complementary probability, by equation (15.5): $\Pr(\bar{U} \mid I) = 0.02$, $\Pr(\bar{G} \mid I, U) = 0.04$, $\Pr(\bar{G} \mid I, \bar{U}) = 0.20$ and $\Pr(\overline{\mathrm{REC}} \mid I, U, \bar{G}) = 0.30$.

Table 15.1: Conditional branch probabilities of the data center event tree (illustrative values).

Event-tree branch	Probability
UPS succeeds	$\Pr(U \mid I) = 0.98$
Generator succeeds, given the UPS succeeds	$\Pr(G \mid I, U) = 0.96$
Generator succeeds, given the UPS fails	$\Pr(G \mid I, \bar{U}) = 0.80$
Timely recovery, given the UPS succeeds and the generator fails	$\Pr(\mathrm{REC} \mid I, U, \bar{G}) = 0.70$

Generator success has two different probabilities, depending on the outcome of the UPS heading. The difference is deliberate. Success of the generator heading means providing the required service under the electrical conditions left by the preceding headings, not merely that

the engine starts. A generator may start properly and still fail to prevent an interruption if the UPS bridging has already been lost, and the success criterion of the heading after a UPS failure, "the generator picks up the load and holds it for the outage", is met with a different probability than after a UPS success.

15.3.2 The complete event tree

Figure 15.3 draws the complete tree. The initiating event is the box on the left. The three column headings are the UPS, the generator and timely recovery; each branch point is labeled with S or F and its conditional probability. To the right of the tree, each of the five terminal sequences carries its number k, its end state (green when IT service is maintained, red when it is interrupted), its conditional probability $\Pr(S_k \mid I)$ and its frequency f_{S_k}, computed in Section 15.3.3.

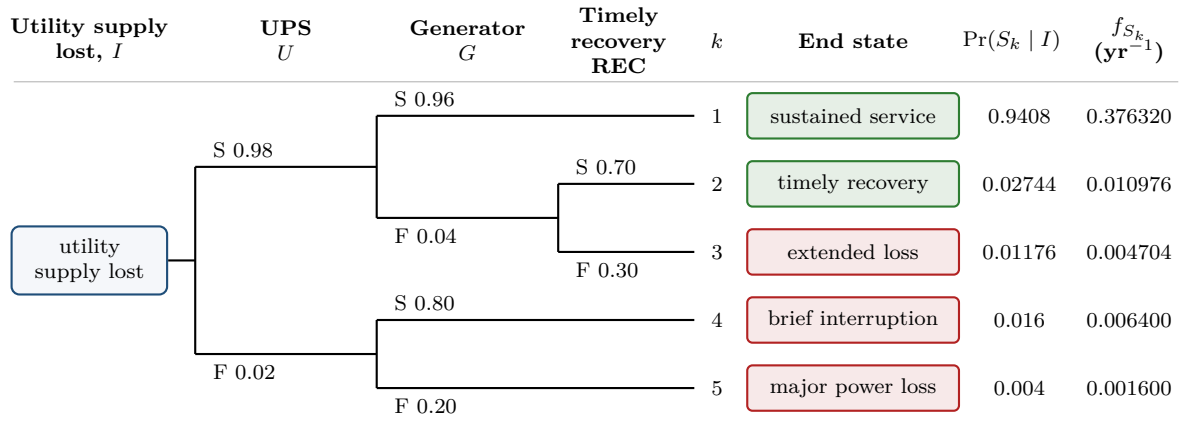


Figure 15.3: Complete event tree for the loss of utility supply at a data center, $f_{\mathrm{IE}} = 0.40\,\mathrm{yr}^{-1}$ (illustrative). Generator branches are conditional on the UPS outcome, and the recovery heading is demanded only when the UPS holds and the generator fails. The five end states are mutually exclusive; their probabilities sum to 1 and their frequencies to $0.40\,\mathrm{yr}^{-1}$.

The tree has five terminal sequences, not the eight that three binary headings would allow, because not every heading is demanded on every path. When the UPS and the generator both succeed, sustained emergency supply is already established and the recovery heading is not asked; the line of sequence 1 crosses the recovery column without a branch. When the UPS succeeds and the generator fails, recovery matters, and the path divides into sequences 2 and 3. When the UPS fails, the IT load is dropped at the instant the utility supply is lost, whatever happens afterward; the tree classifies a generator success on that path as a brief interruption (sequence 4) and a generator failure as a major power loss (sequence 5). The classification assumes that the generator starts slowly enough that the failed UPS bridging causes a brief interruption before the generator takes the load, and that recovery of the utility supply on that path is too slow to matter.

15.3.3 Sequence probabilities and annual frequencies

By the chain rule (15.6), each sequence probability is the product of the branch probabilities along its path:

$$\Pr(S_1 \mid I) = \Pr(U \mid I)\,\Pr(G \mid I, U) = 0.98 \times 0.96 = 0.9408,$$
$$\Pr(S_2 \mid I) = \Pr(U \mid I)\,\Pr(\bar{G} \mid I, U)\,\Pr(\mathrm{REC} \mid I, U, \bar{G}) = 0.98 \times 0.04 \times 0.70 = 0.02744,$$
$$\Pr(S_3 \mid I) = \Pr(U \mid I)\,\Pr(\bar{G} \mid I, U)\,\Pr(\overline{\mathrm{REC}} \mid I, U, \bar{G}) = 0.98 \times 0.04 \times 0.30 = 0.01176,$$
$$\Pr(S_4 \mid I) = \Pr(\bar{U} \mid I)\,\Pr(G \mid I, \bar{U}) = 0.02 \times 0.80 = 0.016,$$
$$\Pr(S_5 \mid I) = \Pr(\bar{U} \mid I)\,\Pr(\bar{G} \mid I, \bar{U}) = 0.02 \times 0.20 = 0.004.$$

The completeness check (15.10) holds: $0.9408 + 0.02744 + 0.01176 + 0.016 + 0.004 = 1$, and the tree is normalized.

Multiplying each conditional probability by the initiating-event frequency, equation (15.7) with $f_{\mathrm{IE}} = 0.40\,\mathrm{yr}^{-1}$, gives the annual sequence frequencies of Table 15.2. Its columns are the sequence number, the path through the headings, the end state, the conditional probability in percent and the expected frequency per year. The frequencies sum to $0.400000\,\mathrm{yr}^{-1}$, the initiating-event frequency, which is the second consistency check of equation (15.10). Sequence 1 carries 94 % of the initiating events; the four other sequences together have a frequency of $0.02368\,\mathrm{yr}^{-1}$, about once every 42 years on average.

Table 15.2: Sequence probabilities and annual frequencies of the data center event tree, $f_{S_k} = f_{\mathrm{IE}}\,\Pr(S_k \mid I)$ with $f_{\mathrm{IE}} = 0.40\,\mathrm{yr}^{-1}$ (illustrative).

k	Path	End state	$\Pr(S_k \mid I)$	f_{S_k} (yr^{-1})
1	UPS and generator succeed	sustained service	94.080 %	0.376320
2	UPS succeeds, generator fails, recovery succeeds	timely recovery	2.744 %	0.010976
3	UPS succeeds, generator fails, recovery fails	extended loss	1.176 %	0.004704
4	UPS fails, generator succeeds	brief interruption	1.600 %	0.006400
5	UPS and generator fail	major power loss	0.400 %	0.001600
	Total		100 %	0.400000

15.3.4 Consequence categories

The terminal sequences do not all produce the same consequence. For illustration, each sequence is assigned an IT service interruption duration D_k, in hours, and a financial consequence per occurrence C_k, in reais, listed in the first columns of Table 15.3: no interruption and no loss for sequences 1 and 2; 2.0 h and R$ 100 000 for the extended loss of sequence 3; 0.1 h and R$ 10 000 for the brief interruption of sequence 4; 6.0 h and R$ 500 000 for the major power loss of sequence 5. These are illustrative scenario consequences, not industry cost estimates. The remaining columns of the table are the contributions of each sequence to the expectations computed below.

Under these assumptions sequences 1 and 2 cause no IT interruption, and sequences 3, 4 and 5 do. Since the sequences are mutually exclusive, the conditional probability of an interruption, given a loss of utility supply, is the sum of their probabilities,

$$\Pr(\text{interruption} \mid I) = 0.01176 + 0.016 + 0.004 = 0.03176,$$

Table 15.3: Consequences of the data center sequences and their contributions to the expected interruption duration and the expected loss per initiating event and per year (illustrative).

k	End state	D_k (h)	C_k (R$)	$\Pr(S_k \mid I)\, D_k$ (h)	$\Pr(S_k \mid I)\, C_k$ (R$)	$f_{S_k} C_k$ (R$/yr)
1	sustained service	0	0	0	0	0
2	timely recovery	0	0	0	0	0
3	extended loss	2.0	100 000	0.02352	1176	470.40
4	brief interruption	0.1	10 000	0.00160	160	64.00
5	major power loss	6.0	500 000	0.02400	2000	800.00
	Total			0.04912	3336	1334.40

so about 3.18 % of initiating events lead to some interruption of IT service. The expected frequency of interruption sequences is

$$f_{\text{int}} = f_{\text{IE}} \Pr(\text{interruption} \mid I) = 0.40 \times 0.03176 = 0.012704\,\text{yr}^{-1}.$$

In reciprocal terms this is one interruption per $1/0.012704 = 78.7$ years on average, under a stationary model. It is not a guarantee of an interruption-free period of that length: by equation (15.3), the probability of at least one interruption in a given year is $1 - \mathrm{e}^{-0.012704} = 0.012624$.

15.3.5 Expected annual interruption hours and financial loss

Let D be the IT interruption duration that follows one initiating event; it equals D_k when sequence S_k occurs. By the law of total expectation over the mutually exclusive sequences, the expected interruption duration per initiating event is

$$\mathbb{E}[D \mid I] = \sum_k \Pr(S_k \mid I)\, D_k, \tag{15.11}$$

which for the data center is

$$\mathbb{E}[D \mid I] = 0.01176 \times 2 + 0.016 \times 0.1 + 0.004 \times 6 = 0.02352 + 0.00160 + 0.02400 = 0.04912\,\text{h}.$$

Over a year, the total interruption duration is the sum of the durations that follow each of the $N_I(1\,\text{yr})$ initiating events. When the outcomes are independent of one another and of their number, the expectation of such a random sum is the expected number of terms times the expected term (a standard result, Wald's identity), so the expected annual interruption duration is

$$\mathbb{E}[D_{\text{yr}}] = f_{\text{IE}}\, \mathbb{E}[D \mid I] = 0.40 \times 0.04912 = 0.019648\,\text{h per year},$$

about 1.18 minutes per year attributable to this initiating-event family. It is the same quantity as the expected downtime of equation (3.21), computed here from sequences instead of from an availability. It is an expectation over many possible years, not a typical outage duration: in most years nothing happens, and in the rare year with a major loss the interruption lasts six hours.

The expected financial loss per initiating event follows in the same way,

$$\mathbb{E}[C \mid I] = \sum_k \Pr(S_k \mid I)\, C_k = 0.01176 \times \text{R\$}\,100\,000 + 0.016 \times \text{R\$}\,10\,000 + 0.004 \times \text{R\$}\,500\,000,$$

that is, R$ 1176 + R$ 160 + R$ 2000 = R$ 3336, and the expected annual loss is

$$\mathcal{R} = \mathbb{E}[C_{\mathrm{yr}}] = f_{\mathrm{IE}}\,\mathbb{E}[C \mid I] = 0.40 \times \mathrm{R\$}\,3336 = \mathrm{R\$}\,1334.40 \text{ per year.} \tag{15.12}$$

This is the risk $\mathcal{R}$ of equation (3.19), expressed as an expected annual loss. It covers only the modeled utility-loss initiating events and the assumed consequences, and it does not represent the total operating risk of the data center.

The last column of Table 15.3 splits the annual loss by sequence. The major power loss, the rarest sequence of the three that cause interruptions with $f_{S_5} = 0.0016\,\mathrm{yr}^{-1}$, contributes R$ 800.00 of the R$ 1334.40, or 60 %, because its consequence is fifty times that of the brief interruption. The brief interruption is four times as frequent and contributes R$ 64.00.

Principle 15.2. *Frequency alone does not determine risk. A rare sequence with a severe consequence can contribute more expected loss than a more frequent sequence with a minor one.*

15.4 Worked example: transmission-line outage

The same construction applies to an electric network. This section quantifies an event tree for the outage of one transmission line and carries its sequence frequencies through to the expected unserved energy. The numbers are illustrative.

15.4.1 The power-system scenario

A transmission corridor serves a load area, and one of its lines experiences an outage, the initiating event, with frequency $f_{\mathrm{IE}} = 0.60\,\mathrm{yr}^{-1}$. The system has three successive responses. First, protection must correctly isolate the faulted line. Second, the remaining network capacity, with a redispatch of generation, must accommodate the outage, keeping the flows on the remaining lines within their ratings. Third, if the redispatch is insufficient, a remedial action scheme (RAS), an automatic scheme that trips generation or load or reconfigures the network when it detects a predefined condition, may prevent a more severe outcome. The conditional branch probabilities are

$$\begin{aligned} \Pr(\text{protection succeeds} \mid I) &= 0.99, \\ \Pr(\text{redispatch succeeds} \mid I, \text{protection succeeds}) &= 0.90, \\ \Pr(\text{RAS succeeds} \mid I, \text{protection succeeds}, \text{redispatch fails}) &= 0.75. \end{aligned}$$

For the running transmission case of this book, whose line ratings depend on the weather, the redispatch branch is where that dependence enters: whether the remaining lines can carry the redistributed flow depends on their ratings at the time of the outage. A more detailed tree would divide the initiating event into weather families, each with its own frequency and its own redispatch probability, as Section 15.2.3 prescribes when the branch probabilities depend on the operating condition.

15.4.2 The four sequences

Figure 15.4 draws the tree, in the same layout as Figure 15.3: the initiating event on the left; the headings protection, redispatch and RAS; the branch probabilities at each branch point; and, on the right, the sequence number, the end state, the conditional probability and the frequency of each of the four terminal sequences. When protection fails, the redispatch and RAS headings are not evaluated, and the line of sequence 4 crosses both columns without a branch; when redispatch succeeds, the RAS is not required, and the line of sequence 1 crosses the RAS

column. The tree assumes that a protection failure leads to a distinct severe disturbance. In a real network backup protection might still clear the fault, and further branches would then be appropriate.

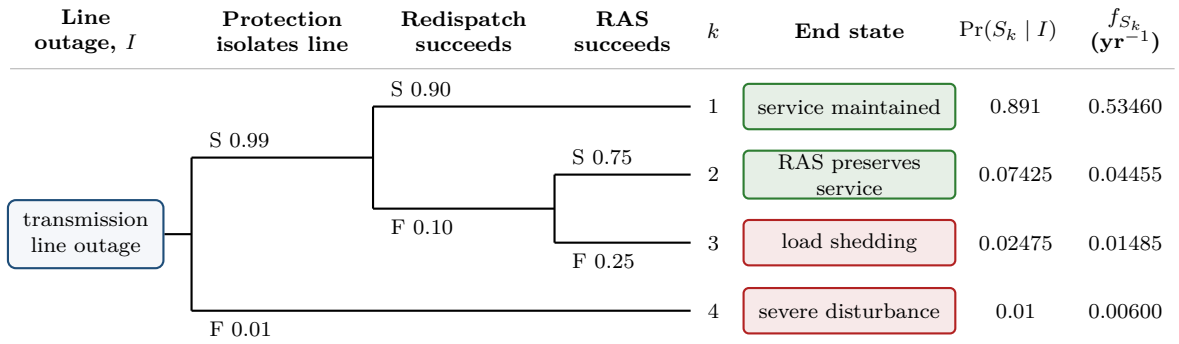


Figure 15.4: Event tree for the outage of one transmission line, $f_{\mathrm{IE}} = 0.60\,\mathrm{yr}^{-1}$ (illustrative). Redispatch and the remedial action scheme (RAS) are evaluated only after successful protection; a protection failure leads directly to a severe disturbance. The sequence probabilities sum to 1 and the frequencies to $0.60\,\mathrm{yr}^{-1}$.

The sequence probabilities follow from the chain rule:

$$\begin{aligned}
\Pr(S_1 \mid I) &= 0.99 \times 0.90 = 0.891,\\
\Pr(S_2 \mid I) &= 0.99 \times 0.10 \times 0.75 = 0.07425,\\
\Pr(S_3 \mid I) &= 0.99 \times 0.10 \times 0.25 = 0.02475,\\
\Pr(S_4 \mid I) &= 0.01.
\end{aligned}$$

They sum to $0.891 + 0.07425 + 0.02475 + 0.01 = 1$. Table 15.4 lists, for each sequence, the outcome at each heading ("not required" or "not evaluated" where the path skips a heading), the end state, the conditional probability, the frequency $f_{S_k} = 0.60 \times \Pr(S_k \mid I)$ and the unserved energy of one occurrence, derived in Section 15.4.3. The frequencies sum to $0.60000\,\mathrm{yr}^{-1}$. Load is shed in about 2.5 % of line outages, and a severe disturbance follows 1 % of them.

Table 15.4: Sequences of the transmission-line event tree, with their conditional probabilities, annual frequencies for $f_{\mathrm{IE}} = 0.60\,\mathrm{yr}^{-1}$, and unserved energy per occurrence (illustrative).

k	Protection	Redispatch	RAS	End state	$\Pr(S_k \mid I)$	f_{S_k} (yr^{-1})	E_k^{unserved} (MWh)
1	success	success	not required	service maintained	89.100 %	0.53460	0
2	success	failure	success	emergency control preserves service	7.425 %	0.04455	0
3	success	failure	failure	load shedding	2.475 %	0.01485	80
4	failure	not evaluated	not evaluated	severe protection-related disturbance	1.000 %	0.00600	400
				Total	100 %	0.60000	

15.4.3 From sequence frequency to expected unserved energy

Power-system reliability usually needs more than the frequency of each sequence: it needs the amount of load affected and the duration of the interruption. Let $L_k^{\text{unserved}}(u)$ be the unserved electrical load, in megawatts, at time u after the start of an occurrence of sequence S_k, and D_k the duration of the interruption. The unserved energy of one occurrence is

$$E_k^{\text{unserved}} = \int_0^{D_k} L_k^{\text{unserved}}(u)\,\mathrm{d}u, \tag{15.13}$$

in megawatt-hours, and if the unserved load is approximately constant, $E_k^{\text{unserved}} \approx L_k^{\text{unserved}}\, D_k$. Suppose sequence 3 sheds 40 MW for two hours, sequence 4 causes a severe disturbance with 100 MW unserved for four hours, and sequences 1 and 2 leave no load unserved. Then

$$E_3^{\text{unserved}} = 40\,\text{MW} \times 2\,\text{h} = 80\,\text{MWh}, \qquad E_4^{\text{unserved}} = 100\,\text{MW} \times 4\,\text{h} = 400\,\text{MWh}.$$

The expected unserved energy was defined in Section 3.6.3, equation (3.24), as the expected integral over the study period of the shortfall $(L - C_{\text{available}})_+$. Over a year, that integral is the sum of the unserved energies of the sequences that occur, and the expected-count argument of Section 15.2.1 gives its expectation as the sum of sequence frequencies times unserved energies. The contribution of this initiating-event family to the expected unserved energy is therefore

$$\text{EUE} = \sum_k f_{S_k}\, E_k^{\text{unserved}}, \tag{15.14}$$

in MWh per year, and for the transmission tree

$$\text{EUE} = 0.01485\,\text{yr}^{-1} \times 80\,\text{MWh} + 0.006\,\text{yr}^{-1} \times 400\,\text{MWh} = 1.188 + 2.400 = 3.588\,\text{MWh per year.}$$

Per initiating event the expected unserved energy is $0.02475 \times 80 + 0.01 \times 400 = 5.98$ MWh, and $0.60 \times 5.98 = 3.588$, as it must be. The severe disturbance, sequence 4, occurs less often than load shedding but contributes two thirds of the expected unserved energy. In the same way, the expected number of hours with load not served is $0.01485 \times 2 + 0.006 \times 4 = 0.0537$ h per year, the contribution of this family to the loss-of-load expectation of equation (3.23) if every hour with shed load counts as a shortfall hour.

The calculation is a simple probabilistic risk calculation in adequacy units, not a complete generation or transmission adequacy model: it covers one initiating-event family, fixed load levels and fixed durations. It shows how an event tree connects initiating events to the system response, the response to consequences, and the consequences to an expected annual impact. A later part of the book develops the loss-of-load probability, the loss-of-load expectation, the expected unserved energy and the detailed power-system reliability models in full.

15.5 Linking fault trees, conditional dependence and recovery timing

The trees of the two worked examples took their branch probabilities as given. This section shows where they come from: from fault trees for the headings, from conditional models that respect the history of each path, and from time-to-recovery distributions compared with the time available.

15.5.1 Branch probabilities from fault trees

A PRA must justify each branch probability. The data center tree used $\Pr(\bar{G} \mid I, U) = 0.04$; a fault tree for the failure of the generator function explains the number. Suppose the generator

fails to supply emergency power if the engine fails to start, the transfer switch fails, or the fuel supply is unavailable. With the basic events

$$X_1 = \{\text{engine fails to start}\}, \quad X_2 = \{\text{transfer switch fails}\}, \quad X_3 = \{\text{fuel supply unavailable}\},$$

the failure of the generator heading is the top event of an OR gate, $\bar{G} = X_1 \cup X_2 \cup X_3$. Its complement is the intersection of the complements, $G = \bar{X}_1 \cap \bar{X}_2 \cap \bar{X}_3$. If the basic events are independent under the demand and operating conditions of the sequence (here, given I and U), with probabilities q_1, q_2, q_3, then $\Pr(G \mid I, U) = (1 - q_1)(1 - q_2)(1 - q_3)$ and

$$\Pr(\bar{G} \mid I, U) = 1 - (1 - q_1)(1 - q_2)(1 - q_3). \tag{15.15}$$

With $q_1 = 0.015$, $q_2 = 0.010$ and $q_3 = 0.015$,

$$\Pr(\bar{G} \mid I, U) = 1 - 0.985 \times 0.990 \times 0.985 = 1 - 0.96052275 = 0.03947725,$$

about 3.95 %. Inclusion–exclusion gives the same value as $0.040 - 0.000525 + 0.00000225$, and the rare-event approximation, the sum $q_1 + q_2 + q_3 = 0.040$, is an upper bound that exceeds it by 0.0005. The fault-tree value is close to the 4 % assumed in the data center tree; used in its place, it lowers the expected annual loss of equation (15.12) from R$ 1334.40 to R$ 1328.25.

Figure 15.5 shows the division of labor. On the left is the fault tree of the generator heading: the top event $\bar{G}$, the OR gate and the three basic events with their probabilities. The arrow carries the top-event probability to the middle box, the branch failure probability. The second arrow carries it to the right, where it becomes the failure branch of the generator heading on the path that follows a UPS success. The fault tree explains the origin of the branch probability; the event tree explains how a generator failure affects the progression of the initiating event. Standard PRA procedures link fault trees to event-tree headings in exactly this way to quantify accident sequences.

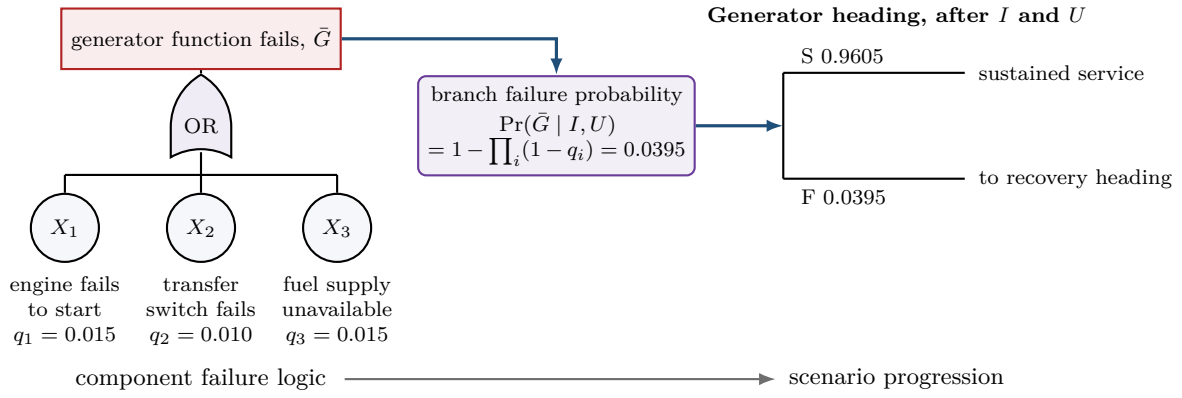


Figure 15.5: Linking a fault tree to an event-tree heading. The fault tree of the generator function (left) gives the top-event probability $\Pr(\bar{G} \mid I, U) = 0.0395$ from three independent basic events; that probability becomes the failure branch of the generator heading in the event tree (right). Illustrative values.

The fault tree of a heading must be built for the conditions of the sequence in which the heading is demanded. The probability that the fuel supply is unavailable, or that the transfer switch fails, can differ between a path on which the UPS holds the bus and a path on which it has dropped; a single fault tree evaluated once and copied into every branch discards that difference, the error that Section 15.5.3 examines.

15.5.2 Repeated basic events across headings

Suppose the same battery-backed control system is required both for UPS operation and for the generator transfer. A failure of that control system defeats both functions. Let X_{c} be the control-system failure, X_{u} a failure specific to the UPS and X_{g} a failure specific to the generator, with probabilities q_{c}, q_{u} and q_{g}, all independent. The fault trees of the two headings are

$$\bar{U} = X_{\mathrm{c}} \cup X_{\mathrm{u}}, \qquad \bar{G} = X_{\mathrm{c}} \cup X_{\mathrm{g}},$$

and they share the basic event X_{c}. Both functions fail when

$$\bar{U} \cap \bar{G} = (X_{\mathrm{c}} \cup X_{\mathrm{u}}) \cap (X_{\mathrm{c}} \cup X_{\mathrm{g}}) = X_{\mathrm{c}} \cup (X_{\mathrm{u}} \cap X_{\mathrm{g}}), \tag{15.16}$$

by the distributive law of Boolean algebra, $(A \cup B) \cap (A \cup C) = A \cup (B \cap C)$. The shared event appears once in the joint logic, as a minimal cut set of order one, $\{X_{\mathrm{c}}\}$, beside the cut set of order two $\{X_{\mathrm{u}}, X_{\mathrm{g}}\}$; this is the treatment of repeated basic events in Chapter 11. Since X_{c} and $\bar{X}_{\mathrm{c}} \cap X_{\mathrm{u}} \cap X_{\mathrm{g}}$ are mutually exclusive,

$$\Pr(\bar{U} \cap \bar{G}) = q_{\mathrm{c}} + (1 - q_{\mathrm{c}})\, q_{\mathrm{u}}\, q_{\mathrm{g}}.$$

The marginal probabilities computed from each fault tree separately cannot simply be multiplied. With illustrative values $q_{\mathrm{c}} = 0.005$, $q_{\mathrm{u}} = 0.015$ and $q_{\mathrm{g}} = 0.035$, the marginal failure probabilities are $\Pr(\bar{U}) = 1 - 0.995 \times 0.985 = 0.019925$ and $\Pr(\bar{G}) = 1 - 0.995 \times 0.965 = 0.039825$, close to the 0.02 and 0.04 of the data center tree. Their product is $0.019925 \times 0.039825 = 0.000794$; the correct joint probability is $0.005 + 0.995 \times 0.015 \times 0.035 = 0.005522$, seven times larger. The conditional probability that the generator fails given that the UPS has failed is $0.005522/0.019925 = 0.277$, against the marginal 0.040: a UPS failure is strong evidence that the shared control system has failed. The success branches are affected as well: the sequence "UPS fails, generator succeeds" is $\bar{U} \cap G = X_{\mathrm{u}} \cap \bar{X}_{\mathrm{c}} \cap \bar{X}_{\mathrm{g}}$, with probability $0.015 \times 0.995 \times 0.965 = 0.014403$, against $0.019925 \times 0.960175 = 0.019131$ from the product of marginals. An integrated fault-tree and event-tree model therefore evaluates each sequence as the Boolean combination of the heading fault trees along its path (failed headings as they are, successful headings complemented), reduces it to minimal cut sets, and only then quantifies it.

Common mistake

Multiplying heading probabilities computed from fault trees that share basic events. A sequence probability built from marginal function probabilities alone can be wrong by a large factor; the shared events must be kept in the joint Boolean logic of the sequence.

15.5.3 Conditional branch consistency

An event tree is not a product of fixed component reliabilities. The data center tree uses $\Pr(G \mid I, U) = 0.96$ but $\Pr(G \mid I, \bar{U}) = 0.80$: the two numbers describe the generator heading on two different histories. After a UPS failure the generator may still start, but power continuity has already been lost; a generator start can count as successful emergency operation without restoring uninterrupted IT service. The success criterion of each heading must therefore be stated for each history on which it is demanded.

If the probability of generator success following the initiating event, without regard to the UPS, is required, it is obtained by marginalizing over the UPS outcomes with the law of total probability,

$$\Pr(G \mid I) = \Pr(G \mid I, U)\, \Pr(U \mid I) + \Pr(G \mid I, \bar{U})\, \Pr(\bar{U} \mid I). \tag{15.17}$$

With the values of Table 15.1,

$$\Pr(G \mid I) = 0.96 \times 0.98 + 0.80 \times 0.02 = 0.9408 + 0.0160 = 0.9568,$$

which is neither 0.96 nor 0.80 but their average weighted by the probabilities of the two histories. Using this marginal probability in both branches would discard the dependence that the tree deliberately models. The effect can be computed. With 0.9568 on both generator branches, the sequence probabilities become 0.937664, 0.029635, 0.012701, 0.019136 and 0.000864; the probability of an interruption rises slightly, from 0.03176 to 0.032701, but the probability of a major power loss falls from 0.004 to 0.000864, and the expected annual loss falls from R$ 1334.40 to R$ 757.38, an understatement of 43 %.

Common mistake

Using the same success probability for a heading in every branch, without regard to the preceding sequence. The marginal probability of equation (15.17) is correct as a summary, but inserted into each branch it removes the dependence on the history and can misstate the frequency of the most severe end states.

15.5.4 Recovery as a time-dependent event

Recovery succeeds only if it happens in time. Consider a UPS that can carry the IT load for $T_{\mathrm{UPS}} = 15\,\mathrm{min}$ after the loss of utility supply, its battery autonomy. If the generator fails, power must be restored within those fifteen minutes to avoid an IT interruption. Let T_{R} be the time from the loss of supply to successful restoration of acceptable power. The probability of timely recovery is

$$\Pr(\mathrm{REC} \mid I, U, \bar{G}) = \Pr(T_{\mathrm{R}} \le T_{\mathrm{UPS}}) = M(T_{\mathrm{UPS}}), \tag{15.18}$$

where $M(t)$ is the cumulative distribution function of the restoration time, the maintainability function of equation (3.13). The branch probability is the maintainability evaluated at the time available.

If the restoration time is exponential with rate $\mu = 0.10\,\mathrm{min}^{-1}$ (a mean of ten minutes), then $M(t) = 1 - \mathrm{e}^{-\mu t}$, equation (3.14), and for the fifteen-minute autonomy

$$\Pr(\mathrm{REC}) = 1 - \mathrm{e}^{-0.10 \times 15} = 1 - \mathrm{e}^{-1.5} = 0.7769.$$

If the battery has only ten minutes of autonomy, the timely-recovery probability becomes $1 - \mathrm{e}^{-1} = 0.6321$: the shorter autonomy raises the probability of unsuccessful recovery from 0.2231 to 0.3679. The value 0.70 used in the data center tree corresponds, under this model, to an autonomy of $-\ln(0.30)/0.10 = 12.04\,\mathrm{min}$. Deriving the branch from a time-to-recovery distribution differs from assigning a generic recovery probability, because it ties the branch to the time actually available.

Table 15.5 carries the model through the data center tree for several autonomies, with all other branch probabilities and consequences as in Section 15.3. Its columns are the autonomy, the timely-recovery probability, the probability of sequence 3 (extended loss), the probability of an interruption per initiating event, the interruption frequency and the expected annual loss. For $T_{\mathrm{UPS}} = 10\,\mathrm{min}$, the row reads $\Pr(S_3 \mid I) = 0.98 \times 0.04 \times 0.367879 = 0.014421$, $\Pr(\mathrm{interruption} \mid I) = 0.014421 + 0.016 + 0.004 = 0.034421$ and $\mathcal{R} = 0.40 \times (0.014421 \times 100\,000 + 160 + 2000) = \mathrm{R\$}\,1440.83$ per year. Lengthening the autonomy from five to thirty minutes halves the interruption frequency and the expected loss, but the loss never falls below $0.40 \times (160 + 2000) = \mathrm{R\$}\,864$ per year, the part that comes from UPS failures, on which recovery has no effect.

The probability of timely recovery depends on more than the restoration time. It depends on the available UPS energy, the IT load, the battery condition, the generator start-up time, the automatic transfer delays, the operator intervention and the condition of the fuel system. Each of these changes either the time available or the time needed, and an event tree whose recovery branch is a fixed number hides all of them.

Table 15.5: Timely-recovery probability $M(T_{\mathrm{UPS}}) = 1 - \mathrm{e}^{-0.10\,T_{\mathrm{UPS}}}$ (T_{UPS} in minutes) and its effect on the data center tree, with the other branch probabilities and consequences of Section 15.3 and $f_{\mathrm{IE}} = 0.40\,\mathrm{yr}^{-1}$ (illustrative).

T_{UPS} (min)	Pr(REC)	$\Pr(S_3 \mid I)$	Pr(int. $\mid I$)	f_{int} (yr^{-1})	$\mathcal{R}$ (R\$/yr)
5	0.393469	0.023776	0.043776	0.017510	1815.04
10	0.632121	0.014421	0.034421	0.013768	1440.83
12.04	0.700000	0.011760	0.031760	0.012704	1334.40
15	0.776870	0.008747	0.028747	0.011499	1213.87
20	0.864665	0.005305	0.025305	0.010122	1076.21
30	0.950213	0.001952	0.021952	0.008781	942.07

15.5.5 A dynamic event tree

In a more detailed model the system state evolves continuously after the initiating event, and the time at which a heading is decided follows from that evolution. Let $E_{\mathrm{batt}}(t)$ be the usable energy remaining in the UPS batteries, in kilowatt-hours. While the UPS supplies the load, a simplified energy balance is

$$\frac{\mathrm{d}E_{\mathrm{batt}}}{\mathrm{d}t} = -P_{\mathrm{load}}(t), \tag{15.19}$$

with the IT load $P_{\mathrm{load}}(t)$ in kilowatts; a detailed model adds the conversion efficiency and the operating limits of the batteries. The UPS fails to provide the required service when $E_{\mathrm{batt}}(t) < E_{\min}$. For a constant load the depletion time, the time at which the UPS can no longer support the required load, is

$$T_{\mathrm{dep}} = \frac{E_{\mathrm{batt}}(0) - E_{\min}}{P_{\mathrm{load}}}. \tag{15.20}$$

A generator, or any other sustained source, picks up the load at an uncertain time T_{G}, and a simplified success event of the heading is $T_{\mathrm{G}} < T_{\mathrm{dep}}$. The event tree uses $\Pr(T_{\mathrm{G}} < T_{\mathrm{dep}})$ as its branch probability. If both times are uncertain, the branch probability is

$$\Pr(T_{\mathrm{G}} < T_{\mathrm{dep}}) = \int_0^{\infty} \Pr(T_{\mathrm{G}} < \tau \mid T_{\mathrm{dep}} = \tau)\, \mathrm{d}F_{\mathrm{dep}}(\tau), \tag{15.21}$$

where F_{dep} is the distribution function of the depletion time. If the two times are independent, the integrand reduces to the distribution function of T_{G} evaluated at τ; if they are dependent, for example because the hot weather that raises the cooling and IT load also delays the start or the loading of the generator, the full joint distribution must be used.

Figure 15.6 shows an illustrative case on a common time axis in minutes after the initiating event, for the path on which the UPS holds and the generator fails to start on demand. The upper panel plots equation (15.19): the usable energy above $E_{\min}$, 125 kWh at the start, falls linearly at a load of 500 kW (probability 0.7) or 750 kW (probability 0.3), reaching zero at $T_{\mathrm{dep}} = 125/500\,\mathrm{h} = 15\,\mathrm{min}$ or $125/750\,\mathrm{h} = 10\,\mathrm{min}$. The lower panel is the event tree. It branches first on the load level and then, on each path, at the depletion time of that path, marked by the dotted line from the upper panel: either a sustained source has picked up the load before depletion (restored) or it has not (depleted). The restoration time is the exponential model of Section 15.5.4, so the branch probabilities are $M(15\,\mathrm{min}) = 0.7769$ and $M(10\,\mathrm{min}) = 0.6321$. The end states on the right carry the products of the branch probabilities.

The figure shows the two features that distinguish a dynamic event tree from the static trees of the earlier sections. The time of each branch point is not fixed by the analyst; it is computed from the physical state, here the battery energy. And the same heading, restoration

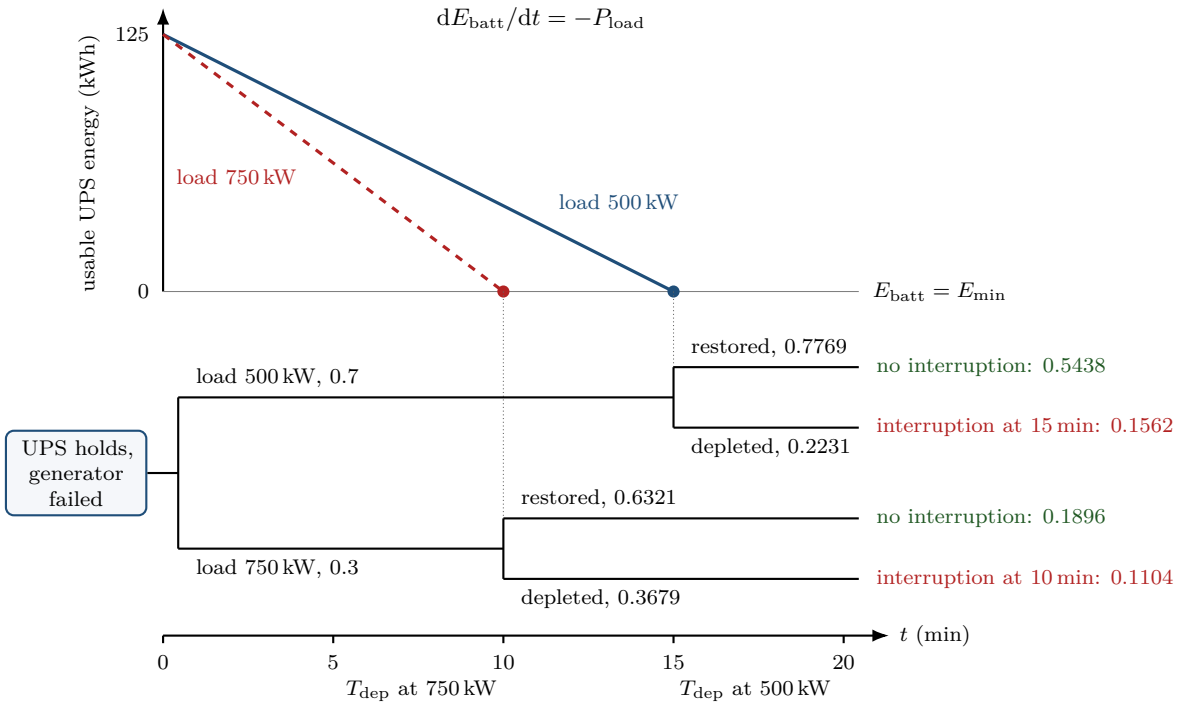


Figure 15.6: A dynamic event tree for the recovery heading of the data center (illustrative). Upper panel: usable UPS energy above E_{min}, $E_{batt}(t) - E_{min} = 125\,\text{kWh} - P_{load}\,t$, for loads of 500 kW and 750 kW. Lower panel: the tree branches on the load and then at each path's depletion time T_{dep}, with restoration before depletion probability $1 - e^{-0.10\,T_{dep}}$ (T_{dep} in minutes). End-state probabilities are conditional on the UPS holding and the generator failing.

before depletion, has a different probability on each path, because the time available differs. The timely-recovery branch probability, averaged over the load, is

$$\Pr(T_G < T_{dep}) = 0.7 \times 0.7769 + 0.3 \times 0.6321 = 0.5438 + 0.1896 = 0.7334,$$

equation (15.21) for a two-point load distribution independent of the restoration time. A static tree that used the fifteen-minute autonomy for every event would report 0.7769 and understate the probability of an interruption on this path, 0.2666, by 16 %. The dynamic event tree is the connection between event trees and the simulation of the physical system: the physics decides when the branches occur, and the event tree aggregates their probabilities.

15.6 Expected consequence, risk and the credibility of event-tree models

The worked examples ended in an expected annual loss and an expected unserved energy. This section states the general expression for risk from sequence frequencies, shows what an expected value leaves out, tabulates the sensitivity of the data center tree to its branch probabilities, and closes with the limits of the method.

15.6.1 From sequences to risk

Each sequence of an event tree has a conditional probability, an occurrence frequency and one or more consequences: a service interruption, unserved energy, equipment damage, a financial loss, a safety impact or an environmental impact. Let C_k be the consequence of one occurrence

of sequence S_k, measured in one stated unit. If it is a known number, the expected annual consequence is

$$\mathcal{R} = \mathbb{E}[C_{\mathrm{yr}}] = \sum_k f_{S_k}\, C_k, \tag{15.22}$$

which is equation (15.12) written sequence by sequence, and with $C_k = E_k^{\mathrm{unserved}}$ it is the expected unserved energy (15.14). The expression is exact, not an approximation: by linearity of expectation, the expected sum of the consequences of all occurrences is the sum over sequences of the expected number of occurrences times the consequence. It does not require the probability of a sequence in a year to be small, which is the advantage of working with frequencies rather than annual probabilities.

A facility is exposed to several initiating-event families i, each with its own tree, sequences $\mathcal{S}_i$ and frequencies $f_{S_{i,k}}$. The total expected annual consequence is

$$\mathcal{R} = \sum_i \sum_{k \in \mathcal{S}_i} f_{S_{i,k}}\, C_{i,k}. \tag{15.23}$$

The double sum assumes that the sequence occurrences and the consequence accounting are defined consistently across families, so that no loss is counted twice: a utility loss that also trips a chiller must be assigned to one family, and its cooling consequence included once. When the consequence of a sequence is itself uncertain, because the outage duration, the severity, the weather, the load or the operating conditions vary, $C_{i,k}$ is replaced by its conditional expectation,

$$\mathcal{R} = \sum_i \sum_{k \in \mathcal{S}_i} f_{S_{i,k}}\, \mathbb{E}[C \mid i, S_k], \tag{15.24}$$

which is the more realistic form and the one a PRA normally uses.

15.6.2 Risk beyond the expected loss

Two facilities can have the same expected annual loss and very different risk profiles. Suppose both have $\mathcal{R} = \mathrm{R\$}\,100\,000$ per year. Facility A has frequent small interruptions, an illustrative 20 per year at R\$ 5000 each; facility B has rare catastrophic ones, an illustrative 0.01 per year at R\$ 10 000 000 each. The expected losses are equal, $20 \times 5000 = 0.01 \times 10\,000\,000 = 100\,000$, but facility A loses about R\$ 100 000 every year, while facility B loses nothing in 99 years out of 100 and, with probability $1 - \mathrm{e}^{-0.01} = 0.00995$ in a given year, loses a sum that may exceed what the organization can absorb. For high-consequence systems the expected loss is therefore complemented by other measures: the probability of exceeding a consequence threshold, the distribution of annual losses, tail probabilities, the conditional value at risk (the expected loss given that the loss exceeds a stated quantile of its distribution), the frequencies of severe event categories, and the risk under extreme operating conditions.

The exceedance frequency $f(C > c)$ is the expected frequency of sequences whose consequence exceeds a threshold c. If the sequence consequences are deterministic,

$$f(C > c) = \sum_k f_{S_k}\, \mathbf{1}\{C_k > c\}, \tag{15.25}$$

and if they are uncertain, $f(C > c) = \sum_k f_{S_k}\, \Pr(C > c \mid S_k)$. Plotted against c, it is the frequency–consequence curve of risk analysis. The area under it is the expected annual consequence: writing $C_k = \int_0^\infty \mathbf{1}\{C_k > c\}\,\mathrm{d}c$ for a nonnegative consequence and exchanging sum and integral,

$$\int_0^\infty f(C > c)\,\mathrm{d}c = \sum_k f_{S_k}\, C_k = \mathcal{R}. \tag{15.26}$$

For the data center, $f(C > c) = 0.012704\,\mathrm{yr}^{-1}$ for $c <$ R\$ 10 000, $0.004704 + 0.0016 = 0.006304\,\mathrm{yr}^{-1}$ for R\$ 10 000 $\le c <$ R\$ 100 000, $0.0016\,\mathrm{yr}^{-1}$ for R\$ 100 000 $\le c <$ R\$ 500 000,

and zero beyond; the area is $0.012704 \times 10\,000 + 0.006304 \times 90\,000 + 0.0016 \times 400\,000 = 127.04 + 567.36 + 640.00 = 1334.40$, the expected annual loss of equation (15.12).

Figure 15.7 plots the three frequency–consequence curves. The horizontal axis is the consequence threshold c in reais and the vertical axis the exceedance frequency per year, both logarithmic. The solid blue step is facility A, the dashed red step facility B, and the dotted green staircase the data center tree of Section 15.3. Facilities A and B enclose the same area on linear axes, yet A's curve is high and ends at R$ 5000 while B's is low and extends to R$ 10 000 000. A requirement such as "losses above R$ 1 000 000 less often than once in a thousand years" is met by facility A and the data center and violated by facility B, whatever the expected losses say. These measures identify rare severe outcomes that need not dominate the average.

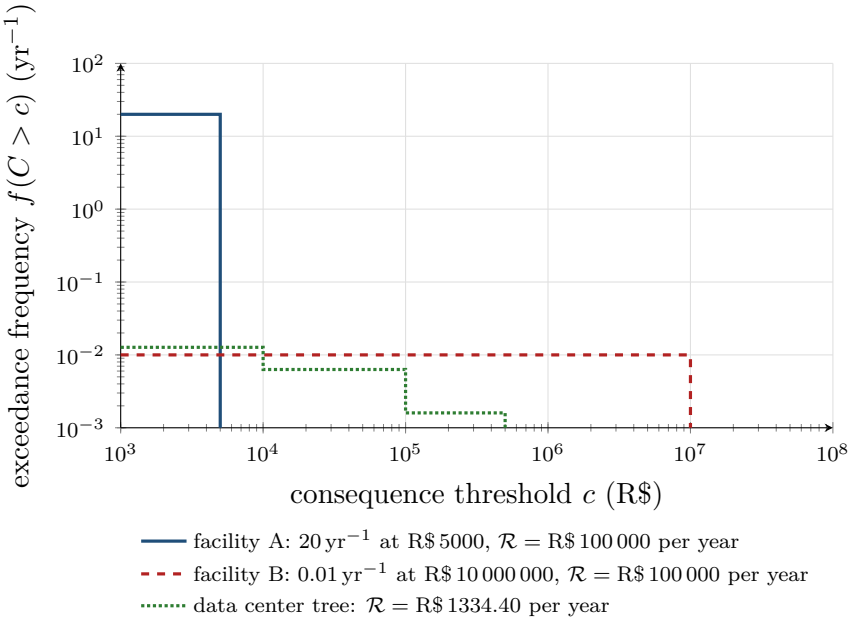


Figure 15.7: Frequency–consequence curves, $f(C > c) = \sum_k f_{S_k}\, \mathbf{1}\{C_k > c\}$ (equation (15.25)), for two illustrative facilities with equal expected annual loss and for the data center event tree of Table 15.3. A drop to the bottom edge means the exceedance frequency is zero beyond that threshold.

15.6.3 Sensitivity of the data center tree: a worked table

The data center tree can be recomputed for any set of branch probabilities. The model is that of Section 15.3: with $u = \Pr(U \mid I)$, $g_1 = \Pr(G \mid I, U)$, $g_0 = \Pr(G \mid I, \bar{U})$ and $r = \Pr(\text{REC} \mid I, U, \bar{G})$,

$$\Pr(S_1 \mid I) = u g_1, \qquad \Pr(S_2 \mid I) = u(1-g_1)r, \qquad \Pr(S_3 \mid I) = u(1-g_1)(1-r),$$
$$\Pr(S_4 \mid I) = (1-u)g_0, \qquad \Pr(S_5 \mid I) = (1-u)(1-g_0),$$

the interruption probability is the sum of the last three, and the expected annual interruption duration and loss follow from equations (15.11) and (15.22) with the consequences of Table 15.3. Table 15.6 evaluates it for the base case and for three single improvements: the UPS success probability raised from 0.98 to 0.999; the generator success after a UPS success raised from 0.96 to 0.99; and the timely recovery raised from 0.70 to 0.90. The rows above the rule are the inputs, the rows below the outputs.

The better-UPS column shows the arithmetic. With $u = 0.999$, $\Pr(S_3 \mid I) = 0.999 \times 0.04 \times 0.30 = 0.011988$, $\Pr(S_4 \mid I) = 0.001 \times 0.80 = 0.0008$ and $\Pr(S_5 \mid I) = 0.001 \times 0.20 = 0.0002$, so

Table 15.6: The data center event tree for four sets of branch probabilities, $f_{\mathrm{IE}} = 0.40\,\mathrm{yr}^{-1}$ and the consequences of Table 15.3 (illustrative). Changed inputs are in bold.

Quantity	Base	Better UPS	Better generator	Better recovery
f_{IE} (yr^{-1})	0.40	0.40	0.40	0.40
UPS success, $\Pr(U \mid I)$	0.98	**0.999**	0.98	0.98
Generator success after UPS success, $\Pr(G \mid I, U)$	0.96	0.96	**0.99**	0.96
Generator success after UPS failure, $\Pr(G \mid I, \bar{U})$	0.80	0.80	0.80	0.80
Timely recovery, $\Pr(\mathrm{REC} \mid I, U, \bar{G})$	0.70	0.70	0.70	**0.90**
$\Pr(S_1 \mid I)$, sustained service	0.940800	0.959040	0.970200	0.940800
$\Pr(S_2 \mid I)$, timely recovery	0.027440	0.027972	0.006860	0.035280
$\Pr(S_3 \mid I)$, extended loss	0.011760	0.011988	0.002940	0.003920
$\Pr(S_4 \mid I)$, brief interruption	0.016000	0.000800	0.016000	0.016000
$\Pr(S_5 \mid I)$, major power loss	0.004000	0.000200	0.004000	0.004000
$\Pr(\mathrm{interruption} \mid I)$	0.031760	0.012988	0.022940	0.023920
Interruption frequency f_{int} (yr^{-1})	0.012704	0.005195	0.009176	0.009568
Expected interruption hours, $\mathbb{E}[D_{\mathrm{yr}}]$ (h/yr)	0.019648	0.010102	0.012592	0.013376
Expected annual loss $\mathcal{R}$ (R$/yr)	1334.40	522.72	981.60	1020.80

the interruption probability is $0.011988 + 0.0008 + 0.0002 = 0.012988$, the interruption frequency $0.40 \times 0.012988 = 0.005195\,\mathrm{yr}^{-1}$, and the expected annual loss $0.40 \times (0.011988 \times 100\,000 + 0.0008 \times 10\,000 + 0.0002 \times 500\,000) = 0.40 \times 1306.80 = \mathrm{R\$}\,522.72$.

Three trends follow. First, an improvement to one function changes several sequence frequencies at once: raising the UPS success moves probability out of sequences 4 and 5 and into sequences 1, 2 and 3, so the extended-loss sequence becomes slightly more probable even though the system improved. Second, the three improvements have different values. The better UPS removes 60.8 % of the expected loss, because it acts on the major power loss that dominates the loss; the better generator removes 26.4 % and the better recovery 23.5 %, because they act only on the extended-loss path. Ranked by interruption frequency instead, the order of the last two is the same but the margin is small. Third, every output except the conditional probabilities is proportional to f_{IE}: doubling the frequency of utility losses doubles the interruption frequency, the interruption hours and the expected loss. The most valuable improvement depends on both the probability of the protection failure and the consequences of the sequences that it feeds.

15.6.4 What event trees do not capture

An event tree is an organizing structure; it does not by itself solve the engineering problems underneath its branches. Table 15.7 lists the issues that most often undermine an event-tree model, each with the reason it matters. The first is the most basic: an initiating event that is not modeled contributes no calculated risk, however frequent it is. The others concern the branches, the order of the headings, the timing, the consequences and the operating conditions, each of which the earlier subsections illustrated. Standard PRA practice separates the corresponding tasks: identification of initiating events, scenario modeling, quantification and the treatment of uncertainty.

Table 15.7: Issues that event trees do not capture automatically.

Issue	Why it matters
Missing initiating events	Unmodeled disturbances contribute no calculated risk.
Incorrect branch probabilities	Sequence frequencies become misleading.
Shared failure causes	Branches may be dependent (Section 15.5.2).
Incorrect event ordering	The modeled sequence may not represent the actual system behavior.
Recovery-time uncertainty	A protection function may succeed too late (Section 15.5.4).
Uncertain physical consequences	Identical equipment states may produce different outcomes.
Changes in operating conditions	The same initiating event may have different consequences under different loads or environmental conditions.

15.6.5 Event trees and dynamic simulation

A conventional event tree is effective when the initiating events are well defined, the protection functions can be represented as discrete outcomes, the number of significant branches is manageable, the branch probabilities can be estimated with reasonable confidence, and the terminal outcomes have identifiable consequences. A dynamic reliability model or a simulation of the physical system may be preferable when control actions depend continuously on the system state, transitions occur at uncertain times, thermal or electrical conditions evolve during the event, failures alter the subsequent failure hazards, or several operator and automatic actions interact.

The two can be combined. A dynamic simulation can determine whether a specific RAS operation prevents voltage collapse for a given contingency and operating state; a probabilistic event tree can then aggregate those outcomes across scenarios, operating states and protection failures, which is the structure of Figure 15.6 applied to a network. The objective is not to use the most elaborate model possible, but to represent the mechanisms that materially affect the decision.

A later chapter on integrated probabilistic risk assessment combines initiating events, event trees, linked fault trees, dependent failures and consequence models in one traceable study, with uncertainty propagation and sensitivity analysis.

15.7 What to retain

Definition 15.1 (Event tree and accident-sequence frequency)**.** An event tree partitions the progressions that can follow an initiating event into mutually exclusive accident sequences. The conditional probability of each sequence is the product of its branch probabilities, each conditioned on the initiating event and the preceding history; its frequency is the initiating-event frequency times that probability; and the expected annual consequence is the sum over sequences of frequency times consequence, which a frequency–consequence curve complements for rare severe outcomes.

Event trees model how disturbances progress; conditional probabilities are essential because a protection function depends on what preceded it; a sequence frequency is an expected number of occurrences per unit time, not a probability; fault trees supply the branch probabilities and event trees combine them into outcomes; consequences, not frequencies alone, determine risk; and timing, loading, thermal response, electrical dynamics and control actions may require

analysis beyond a static tree. Table 15.8 lists the chapter's terms, their meaning and units, and the subsection that defines each.

Table 15.8: The terms of this chapter and where each is defined.

Term	Meaning	Expression	Units	Defined in
Initiating event	disturbance that begins a sequence	I	–	15.1.2
Initiating-event frequency	expected occurrences per year	$f_{\mathrm{IE}} = \mathbb{E}[N_I(H)]/H$	yr^{-1}	15.1.2
Initiating-event probability	at least one in $[0, H]$	$1 - \mathrm{e}^{-f_{\mathrm{IE}} H}$	–	15.1.2
Heading, branch	function asked; its success or failure	$\Pr(\bar{G} \mid I) = 1 - p_G$	–	15.2.1
Accident sequence, end state	path to an outcome; its category	$S_k = I \cap B_1 \cap \cdots \cap B_n$	–	15.2.3
Sequence probability	product of conditional branch probabilities	equation (15.6)	–	15.2.3
Sequence frequency	expected occurrences per year	$f_{S_k} = f_{\mathrm{IE}} \Pr(S_k \mid I)$	yr^{-1}	15.2.3
Independent-branch approximation	headings independent given I	$\prod_j \Pr(B_j \mid I)$	–	15.2.4
Completeness	sequences partition the outcomes	$\sum_k \Pr(S_k \mid I) = 1$	–	15.2.5
Fault-tree link	heading failure as a top event	$1 - \prod_i (1 - q_i)$ for an OR gate	–	15.5.1
Repeated basic event	cause shared by two headings	$X_{\mathrm{c}} \cup (X_{\mathrm{u}} \cap X_{\mathrm{g}})$	–	15.5.2
Timely recovery	restoration within the time available	$M(T_{\mathrm{UPS}})$	–	15.5.4
Dynamic event tree	branch times from the physical state	$\Pr(T_{\mathrm{G}} < T_{\mathrm{dep}})$	–	15.5.5
Unserved energy per sequence	energy not supplied in one occurrence	$\int_0^{D_k} L_k^{\mathrm{unserved}} \,\mathrm{d}u$	MWh	15.4.3
Expected unserved energy	per initiating-event family	$\sum_k f_{S_k} E_k^{\mathrm{unserved}}$	MWh/yr	15.4.3
Expected annual consequence	risk from sequences	$\mathcal{R} = \sum_k f_{S_k} C_k$	R$/yr	15.6.1
Exceedance frequency	frequency of $C > c$	$\sum_k f_{S_k} \mathbf{1}\{C_k > c\}$	yr^{-1}	15.6.2

Further reading. NASA, *Probabilistic Risk Assessment Procedures*, for initiating events, scenario modeling, fault-tree links and frequency quantification; NUREG/CR-6823, *Handbook of Parameter Estimation for Probabilistic Risk Assessment*, for estimating initiating-event frequencies, equipment failure probabilities and recovery probabilities; Rausand, Barros and Høyland, *System Reliability Theory*, third edition, for event trees and reliability modeling; Kumamoto and Henley, *Probabilistic Risk Assessment and Management for Engineers and Scientists*, second edition, for accident sequences and risk quantification.

15.8 Exercises

Exercise 15.1. (a) State what distinguishes an initiating-event frequency from an initiating-event probability. (b) Utility-supply losses occur as a homogeneous Poisson process with $f_{\mathrm{IE}} = 0.20\,\mathrm{yr}^{-1}$. Compute the probability of at least one loss in one year and in five years, the expected number of losses in five years, and the probability that five years pass without one. (c) A sequence has an expected frequency of $0.02\,\mathrm{yr}^{-1}$. State whether this guarantees one occurrence every 50 years, and compute the probability of at least one occurrence in 50 years. (Answers: (a) a frequency has inverse-time units and is an expected number of occurrences per unit time; a probability is dimensionless and refers to a stated interval. (b) 0.1813, 0.6321, 1 and 0.3679. (c) No; it is an average occurrence frequency, not a deterministic recurrence interval; the probability is $1 - \mathrm{e}^{-1} = 0.6321$.)

Exercise 15.2. (a) State how a sequence probability is calculated when the branch outcomes are dependent, and what the probabilities of all complete, mutually exclusive terminal sequences must sum to. (b) In the transmission tree of Section 15.4, a protection upgrade raises $\Pr(\text{protection succeeds} \mid I)$ from 0.99 to 0.995, with the other branch probabilities, $f_{\mathrm{IE}} = 0.60\,\mathrm{yr}^{-1}$ and the unserved energies unchanged. Compute the four sequence probabilities and frequencies and check both normalizations. (c) Compute the expected unserved energy and its reduction from the 3.588 MWh per year of the original tree. (Answers: (a) by multiplying the branch probabilities, each conditioned on the initiating event and all preceding outcomes of the sequence; to one, since the sequences partition the outcomes given the initiating event, and their frequencies sum to f_{IE}. (b) 0.8955, 0.074625, 0.024875 and 0.005, sum 1; frequencies 0.5373, 0.044775, 0.014925 and $0.003\,\mathrm{yr}^{-1}$, sum $0.60\,\mathrm{yr}^{-1}$. (c) $0.014925 \times 80 + 0.003 \times 400 = 1.194 + 1.200 = 2.394$ MWh per year, a reduction of 33 %.)

Exercise 15.3. (a) State the role of a fault tree linked to an event-tree branch. (b) A generator heading fails if any of three independent basic events occurs, with $q_1 = 0.020$, $q_2 = 0.010$ and $q_3 = 0.005$. Compute the branch failure probability exactly and with the rare-event approximation. (c) The UPS and generator headings share a control-system failure: $\bar{U} = X_{\mathrm{c}} \cup X_{\mathrm{u}}$ and $\bar{G} = X_{\mathrm{c}} \cup X_{\mathrm{g}}$, with independent $q_{\mathrm{c}} = 0.01$, $q_{\mathrm{u}} = 0.01$ and $q_{\mathrm{g}} = 0.03$. Compute $\Pr(\bar{U})$, $\Pr(\bar{G})$, the correct $\Pr(\bar{U} \cap \bar{G})$, the product of the marginals, and $\Pr(\bar{G} \mid \bar{U})$. (Answers: (a) it computes the failure probability of the heading's function from its underlying basic events, which the event tree then uses as a branch probability. (b) $1 - 0.98 \times 0.99 \times 0.995 = 0.034651$ exactly, 0.035 approximately. (c) 0.0199, 0.0397, $0.01 + 0.99 \times 0.01 \times 0.03 = 0.010297$, against a product of 0.000790, thirteen times smaller; $\Pr(\bar{G} \mid \bar{U}) = 0.010297/0.0199 = 0.517$.)

Exercise 15.4. The UPS of the data center has only ten minutes of autonomy, and the restoration time is exponential with rate $0.10\,\mathrm{min}^{-1}$. All other branch probabilities and consequences are those of Section 15.3. (a) Compute the timely-recovery probability. (b) Compute $\Pr(S_2 \mid I)$, $\Pr(S_3 \mid I)$, the probability of an interruption per initiating event and the interruption frequency. (c) Compute the expected annual interruption duration and the expected annual loss. (d) Compute the exceedance frequency of losses above R$ 50 000 and compare it with the original tree. (Answers: (a) $1 - \mathrm{e}^{-1} = 0.632121$. (b) 0.024779, 0.014421, 0.034421 and $0.013768\,\mathrm{yr}^{-1}$. (c) $\mathbb{E}[D \mid I] = 0.028842 + 0.0016 + 0.024 = 0.054442$ h, so 0.021777 h (about 1.31 minutes) per year; $\mathcal{R} = 0.40 \times \mathrm{R\$}\,3602.09 = \mathrm{R\$}\,1440.83$ per year. (d) Sequences 3 and 5 exceed the threshold: $0.40 \times (0.014421 + 0.004) = 0.007368\,\mathrm{yr}^{-1}$, against $0.006304\,\mathrm{yr}^{-1}$ in the original tree.)

Chapter 16

Integrated Probabilistic Risk Assessment

16.1 From component reliability to an integrated risk assessment

The preceding chapters developed the building blocks of reliability engineering one at a time: lifetime models and the probability of failure of one component (Chapter 6); estimation of that probability from operating records and from prior knowledge (Chapters 7 and 8); reliability block diagrams and redundancy (Chapter 9); fault trees, minimal cut sets and their quantification (Chapters 10 and 11); importance measures (Chapter 12); common-cause and dependent failures (Chapters 13 and 14); and event trees with their accident sequences (Chapter 15). Each one answers part of a question that none of them answers alone. This chapter, the last of the part on dependencies and probabilistic risk assessment, joins them into one model. Its question is

how component reliability, shared failure mechanisms, initiating events, system responses and consequences are combined into one defensible assessment of engineering risk.

The running example is the power continuity of a data center, the system whose event tree Chapter 15 built. Carried through every step, it yields a model that answers seven questions in order: which disturbances can initiate an event; which protection functions may fail; how shared causes change the probabilities; which operational consequences can follow; what the expected annual risk is; which engineering improvements are worth considering; and how credible the result is, given that its inputs are uncertain estimates. The same structure applies to transmission networks and cooling plants, and the chapter ends by mapping it onto both.

Jargon: integrated PRA, risk triplet, linked fault tree, shared basic event, risk-informed decision, uncertainty propagation, traceability

A *probabilistic risk assessment* (PRA) is *integrated* when one model carries the analysis from the initiating events of a defined system, through the protection functions that respond to them and the fault trees that explain why those functions fail, to the end states, their consequences, the uncertainty of the result and the decisions it supports. The *risk triplet* is the standard statement of what such a model produces: the set of scenarios that can occur, the frequency of each, and the consequence of each. A *linked fault tree* is a fault tree whose top event is the failure of an event-tree heading, so that the heading's branch probability is computed from basic events rather than assigned. A *shared basic event* is a basic event that appears in the fault trees of more than one heading, for example the failure of a support system that two functions both need; it makes the headings dependent. A *risk-informed decision* uses the computed risk, with its uncertainty, as one input to an engineering choice, beside cost, feasibility and other requirements. *Uncertainty propagation* carries the uncertainty of the input parameters through the model to the uncertainty of its outputs. *Traceability* is the property that every number in the model can be followed back to the evidence and the assumption it rests on.

Key idea

Risk is the frequency-weighted sum of consequences over the mutually exclusive scenarios that follow each initiating event; an integrated PRA computes every term of that sum from one model, keeps the shared causes inside the joint logic of each scenario, and carries the uncertainty of every input through to the decision.

This section defines the chain of steps that makes a PRA integrated and explains why a fault-tree probability alone is not a risk.

16.1.1 The complete risk chain

A PRA starts from a defined system and a defined undesirable outcome. It then connects initiating events to sequences of equipment behavior, protective actions and consequences. Figure 16.1 draws the chain as six stages. Each box is one stage of the analysis; the text beneath a box states the question that stage answers and, in the last line, the quantity it hands to the next stage. The arrows give the order of the computation. The brace over the second and third boxes marks the scenario model, the fault trees and the event tree, which together turn an initiating event into a set of sequences with probabilities. The dashed arrow along the bottom returns from the decision stage to the fault trees: a proposed design change modifies basic-event probabilities or the logic, and the chain is evaluated again for the modified system.

The quantity at the end of the chain is usually an expected consequence rate. Let the initiating events be grouped into families $I_1, I_2, \ldots$, each with its frequency $f_{\mathrm{IE},i}$ in events per year and its event tree, whose terminal sequences $S_{i,k}$, $k \in \mathcal{S}_i$, partition what can follow one occurrence of I_i. Let C be the consequence of one occurrence, a random quantity whose conditional expectation given the family and the sequence is $\mathbb{E}[C \mid I_i, S_{i,k}]$. The expected consequence per occurrence of I_i follows from the law of total expectation over the mutually exclusive sequences, $\sum_k \Pr(S_{i,k} \mid I_i)\, \mathbb{E}[C \mid I_i, S_{i,k}]$; the expected consequence per year of the family is that amount times the expected number of occurrences per year, by the random-sum argument of Section 15.3.5; and the families add. Hence

$$\mathcal{R} = \sum_i f_{\mathrm{IE},i} \sum_{k \in \mathcal{S}_i} \Pr(S_{i,k} \mid I_i)\, \mathbb{E}[C \mid I_i, S_{i,k}], \tag{16.1}$$

which is equation (15.24) with the sequence frequency written out as $f_{S_{i,k}} = f_{\mathrm{IE},i} \Pr(S_{i,k} \mid I_i)$,

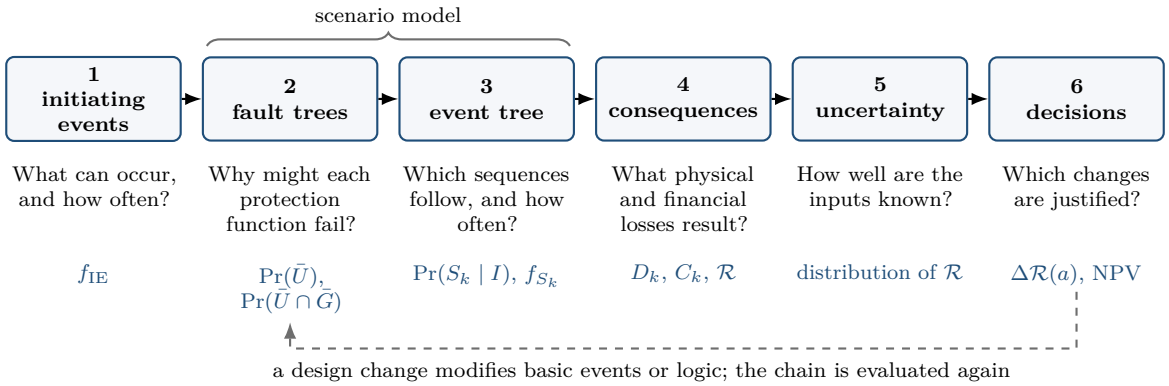


Figure 16.1: The risk chain of an integrated probabilistic risk assessment. Each stage answers one question and passes one quantity (blue) to the next; the dashed loop is the reevaluation of the chain for a proposed design change.

equation (15.7). If C is measured in reais, $\mathcal{R}$ is an expected loss in reais per year; if C is the unserved energy in megawatt-hours, $\mathcal{R}$ is the expected unserved energy in megawatt-hours per year. The three factors have different natures and different units: the frequency is a rate in yr^{-1}, the sequence probability is a dimensionless probability per initiating event, and the consequence is a physical or monetary amount per occurrence. Keeping them apart is what allows each to be estimated from its own evidence.

In the language of the risk triplet, the sequences $S_{i,k}$ are the scenarios, $f_{S_{i,k}}$ their frequencies and $\mathbb{E}[C \mid I_i, S_{i,k}]$ their consequences. The expected value $\mathcal{R}$ is one summary of the triplet; the frequency–consequence curve of Section 15.6.2 is another, and the decision may need both.

16.1.2 A PRA is not a fault-tree calculation

A fault tree can establish that the probability that both cooling-water pumps are unavailable is 0.5 %. What that probability means for the operation is a separate question. Another cooling subsystem may compensate; thermal storage may carry the load for twenty minutes; the resulting temperature rise may be small; or the failure may cause the immediate loss of a critical process. The same component-failure event can produce very different consequences in different system states, and a top-event probability does not distinguish among them.

A complete PRA therefore connects component failures to *functional outcomes*, the success or failure of the required function of the system, and not only to component states. The event tree supplies the functional outcomes, the fault trees supply the probabilities that the protection functions fail, and the consequence model supplies the physical and financial effects of each outcome. Removing any one of the three leaves a number whose meaning for the operation is undetermined.

Common mistake

Reporting a top-event probability as the risk of a system. A fault-tree result is a probability per demand or a snapshot unavailability of one function; it becomes a risk only when it is placed in a scenario that starts from an initiating event with a frequency and ends in a consequence.

16.2 The system, its protection functions and their dependencies

This section defines the data center power-continuity system and the scope of the assessment, assigns the initiating-event frequency, and builds the fault trees of the protection functions, which share one basic event. It then computes the marginal and joint failure probabilities of the functions and shows what an independence assumption would give instead.

16.2.1 The data center power-continuity system

The simplified electrical architecture of the data center has five elements: the utility supply; an uninterruptible power supply (UPS), which carries the IT load from its batteries without interruption when the utility supply is lost; an emergency standby generator; the automatic transfer and control equipment that switches the load between sources; and a recovery procedure, by which operators restore an acceptable supply if the generator fails. The assessment quantifies the risk that arises from the loss of utility supply.

The model is deliberately limited to one initiating-event family. Other sources of risk, among them internal switchgear faults, cooling failures, cyber incidents and failures of the downstream power distribution, are outside its scope. Table 16.1 states the top-level definition of the system in seven attributes: the required function, the initiating event, the assessment horizon, the unit of frequency, the consequences measured, the boundary of the model and what it excludes. The last two rows matter as much as the first: a clear boundary prevents a partial PRA result from being read as the total risk of the facility.

Table 16.1: Top-level definition of the data center power-continuity assessment.

Attribute	Definition
Required function	Maintain or recover IT electrical service
Initiating event	Loss of normal utility supply
Assessment horizon	One year
Frequency unit	Initiating events per year
Consequences	IT interruption duration and financial loss
Boundary	UPS, emergency generation, transfer and control function, and recovery
Excluded	Other initiating events and unrelated infrastructure failures

16.2.2 The initiating-event frequency

Suppose the operating history of the site and comparable sites gives

$$f_{\mathrm{IE}} = 0.40\,\mathrm{yr}^{-1}, \tag{16.2}$$

an expected 0.40 losses of utility supply per year, the value used in Section 15.3 and the estimate that four losses in ten years of records would give. If the losses occur approximately as a homogeneous Poisson process, equation (15.3) gives the probability of at least one loss in a year,

$$\Pr\bigl(N_I(1\,\mathrm{yr}) \geq 1\bigr) = 1 - \mathrm{e}^{-0.40} = 1 - 0.670320 = 0.329680,$$

about 32.97 %. The expected annual number of losses remains 0.40. The two numbers differ because some years have more than one loss: $\Pr(N_I \geq 2) = 1 - \mathrm{e}^{-0.40}(1 + 0.40) = 1 - 0.938448 = 0.061552$, so about 6.2 % of years have two or more.

The sequence frequencies are computed from f_{IE}, not from the annual probability 0.3297. The reason is the linearity of expectation used in equation (16.1): the expected number of sequences of a given kind in a year is the expected number of initiating events times the conditional probability of the sequence, whatever the number of events in that year. Multiplying the annual probability by the conditional probability would count a year with two losses as one and understate every sequence frequency by $1 - 0.3297/0.40 = 17.6\,\%$.

16.2.3 Protection functions and their basic events

Three protection functions respond to the loss of utility supply; they are the headings of the event tree of Section 16.3, and their events are written as in Chapter 15.

- *UPS ride-through*, event U: provides uninterrupted IT power after the initiating event; its failure is $\bar{U}$.
- *Emergency generator service*, event G: establishes an acceptable emergency power supply; its failure is $\bar{G}$.
- *Timely recovery*, event REC: restores power before the UPS energy is depleted, when the generator is unavailable; its failure is $\overline{\mathrm{REC}}$.

The failures of the first two functions are explained by three basic events, each with a probability per initiating event:

- X_{u}, a failure specific to the UPS, with $q_{\mathrm{u}} = 0.02$;
- X_{g}, a failure specific to the generator, with $q_{\mathrm{g}} = 0.04$;
- X_{c}, a failure of a shared electrical and control system that both functions need, with $q_{\mathrm{c}} = 0.01$.

The three basic events are assumed mutually independent. The function failures they produce are not independent, because X_{c} disables both. For the recovery heading, the probability of success is assigned directly,

$$\Pr(\mathrm{REC} \mid I, U, \bar{G}) = 0.70, \qquad \text{so that} \qquad q_{\mathrm{rec}} = \Pr(\overline{\mathrm{REC}} \mid I, U, \bar{G}) = 0.30,$$

the value of Section 15.3, which corresponds under the exponential restoration model of Section 15.5.4 to a battery autonomy of about twelve minutes. All four probabilities are illustrative.

In the terminology of Section 13.3.2, the shared event is a dependency on a support system, modeled explicitly as its own basic event, rather than a common-cause failure of similar components quantified by a beta or alpha factor. The structure is that of the explicit common-cause event model of equation (13.13): each function fails from its own mechanism or from the shared one. The shared event is called the common cause of the two function failures in what follows, in that sense.

16.2.4 The linked fault trees

The failure of each protection function is the top event of a fault tree with an OR gate:

$$\bar{U} = X_{\mathrm{c}} \cup X_{\mathrm{u}}, \qquad \bar{G} = X_{\mathrm{c}} \cup X_{\mathrm{g}}. \tag{16.3}$$

Figure 16.2 draws the two trees. The boxes at the top are the two top events, the UPS function failure on the left and the generator function failure on the right; each sits on an OR gate; the circles are the basic events, with their probabilities beneath them. The red circle in the middle is the shared event X_{c}, which feeds both OR gates. The annotations beside the top events give their marginal probabilities, and the box between them gives the probability that both fail, derived in Section 16.2.5.

The sharing of X_{c} is the critical feature of the model. If the marginal probability of a UPS failure were multiplied by the marginal probability of a generator failure, the shared cause

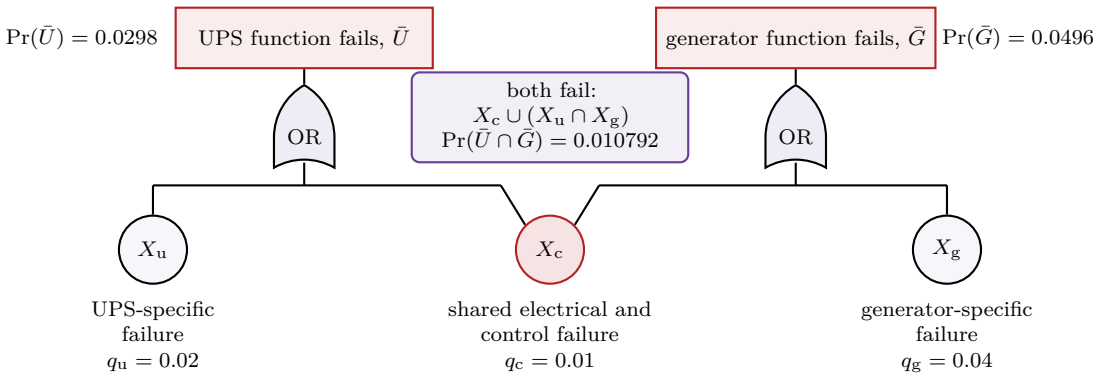


Figure 16.2: The linked fault trees of the UPS and generator functions (illustrative probabilities per initiating event). The shared event X_c appears in both trees; the joint failure probability 0.010792 is 7.3 times the product of the marginals, $0.0298 \times 0.0496 = 0.00147808$.

would be counted as if it acted twice, independently, and the probability that both functions fail would be badly underestimated. The fault trees are linked to the event tree in the sense of Section 15.5.1: their top events are the failure branches of the UPS and generator headings, and every sequence is evaluated from the joint logic of the basic events along its path, as Section 15.5.2 prescribed.

16.2.5 Marginal and joint function probabilities

The UPS function succeeds only if neither X_c nor X_u occurs, $U = \bar{X}_c \cap \bar{X}_u$, and by independence $\Pr(U) = (1 - q_c)(1 - q_u)$. The marginal failure probabilities of the two functions are therefore

$$\Pr(\bar{U}) = 1 - (1 - q_c)(1 - q_u), \qquad \Pr(\bar{G}) = 1 - (1 - q_c)(1 - q_g), \tag{16.4}$$

which give

$$\Pr(\bar{U}) = 1 - 0.99 \times 0.98 = 1 - 0.9702 = 0.0298, \qquad \Pr(\bar{G}) = 1 - 0.99 \times 0.96 = 1 - 0.9504 = 0.0496.$$

The UPS function fails on 2.98 % of initiating events and the generator function on 4.96 %; all probabilities in this section are conditional on the initiating event, and the condition is left implicit where no confusion arises.

Both functions fail when the shared event occurs, or when both specific mechanisms occur. By the distributive law, as in equation (15.16),

$$\bar{U} \cap \bar{G} = (X_c \cup X_u) \cap (X_c \cup X_g) = X_c \cup (X_u \cap X_g).$$

The two events of the union overlap where X_c and both specific failures occur together. Writing the union as the disjoint union of X_c and $\bar{X}_c \cap X_u \cap X_g$ removes the overlap, and independence gives

$$\Pr(\bar{U} \cap \bar{G}) = q_c + (1 - q_c)\, q_u\, q_g = 0.01 + 0.99 \times 0.02 \times 0.04 = 0.010792, \tag{16.5}$$

or 1.0792 %. Almost all of it, 0.01 of 0.010792, comes from the shared event. The same value follows from inclusion–exclusion as a check: $\Pr(\bar{U} \cup \bar{G}) = 1 - (1 - q_c)(1 - q_u)(1 - q_g) = 1 - 0.931392 = 0.068608$, and $\Pr(\bar{U} \cap \bar{G}) = \Pr(\bar{U}) + \Pr(\bar{G}) - \Pr(\bar{U} \cup \bar{G}) = 0.0298 + 0.0496 - 0.068608 = 0.010792$.

The dependence is clearest as a conditional probability. Given that the UPS function has failed, the probability that the generator function also fails is

$$\Pr(\bar{G} \mid \bar{U}) = \frac{\Pr(\bar{U} \cap \bar{G})}{\Pr(\bar{U})} = \frac{0.010792}{0.0298} = 0.362148,$$

against the marginal 0.0496. Bayes' rule explains the number: a UPS failure has two possible causes, and the probability that it was the shared one is $\Pr(X_\mathrm{c} \mid \bar{U}) = 0.01/0.0298 = 0.335570$; if it was, the generator fails with certainty, and if it was the UPS-specific mechanism, the generator fails with probability $q_\mathrm{g} = 0.04$. Hence $\Pr(\bar{G} \mid \bar{U}) = 0.335570 + 0.664430 \times 0.04 = 0.362148$. Given a UPS success, by contrast, X_c is excluded and $\Pr(\bar{G} \mid U) = q_\mathrm{g} = 0.04$.

16.2.6 What the independence assumption would give

Multiplying the marginal probabilities, as an analyst would if the two fault trees were quantified separately and treated as independent, gives

$$\Pr(\bar{U})\ \Pr(\bar{G}) = 0.0298 \times 0.0496 = 0.00147808,$$

or 0.1478 %. The correct joint probability (16.5) is $0.010792/0.00147808 = 7.30$ times larger. The error is not a rounding matter: the product counts the shared event as two independent events with probabilities near 0.01 each, whose coincidence is of order 10^{-4}, when in fact one occurrence of it fails both functions.

The error propagates to every quantity downstream. Section 16.3.6 shows that an event tree built from the marginal probabilities, with the generator heading given the same probability $\Pr(G) = 0.9504$ on both UPS branches, understates the expected annual loss by 63 %, while it slightly overstates the frequency of interruptions; the two errors go in opposite directions because the shared cause moves probability from the brief interruption into the major one. A common cause must therefore be preserved from the fault trees through the event tree to the consequences.

> **Principle 16.1.** *Dependence must be preserved throughout the integrated model. Every sequence is evaluated from the joint Boolean logic of the basic events along its path, with each shared basic event counted once; heading probabilities computed separately are never multiplied.*

16.3 The integrated event tree and its consequences

This section defines the five mutually exclusive sequences that follow the loss of utility supply, computes the probability of each from the basic events and checks that they sum to one, converts them into annual frequencies, attaches consequences, and computes the expected annual interruption frequency, interruption duration and financial loss.

16.3.1 Mutually exclusive sequences

After the initiating event, the outcomes are classified into five sequences, listed in the first columns of Table 16.2: the outcome of the UPS, generator and recovery headings on each, and the operational outcome. Recovery is credited only when the UPS succeeds and the generator fails; when both the UPS and the generator succeed it is not needed. If the UPS fails, the IT load is dropped at the instant of the supply loss, so a generator success restores service but does not prevent a brief interruption, and recovery is not credited on that path. The event tree is a classification of physical outcomes, not a rearrangement of component states, and in a real study each of these outcome assumptions would require engineering verification, for example of the generator start time against the time the IT equipment can tolerate without supply.

Each sequence is an event in the basic events and the recovery outcome:

$$
\begin{aligned}
S_1 &= U \cap G = \bar{X}_\mathrm{c} \cap \bar{X}_\mathrm{u} \cap \bar{X}_\mathrm{g},\\
S_2 &= U \cap \bar{G} \cap \mathrm{REC} = \bar{X}_\mathrm{c} \cap \bar{X}_\mathrm{u} \cap X_\mathrm{g} \cap \mathrm{REC},\\
S_3 &= U \cap \bar{G} \cap \overline{\mathrm{REC}} = \bar{X}_\mathrm{c} \cap \bar{X}_\mathrm{u} \cap X_\mathrm{g} \cap \overline{\mathrm{REC}},\\
S_4 &= \bar{U} \cap G = \bar{X}_\mathrm{c} \cap X_\mathrm{u} \cap \bar{X}_\mathrm{g},\\
S_5 &= \bar{U} \cap \bar{G} = X_\mathrm{c} \cup (X_\mathrm{u} \cap X_\mathrm{g}).
\end{aligned}
$$

The UPS success in S_1 to S_3 excludes X_c and X_u, so a generator failure on those paths can only come from X_g; the generator success in S_4 excludes X_c and X_g, so the UPS failure there comes from X_u; and S_5 is the joint failure of equation (16.5). The recovery outcome is assumed independent of the basic events, given the path on which it is demanded.

Figure 16.3 draws the tree in the layout of Figure 15.3: the initiating event on the left; the UPS, generator and recovery headings across the top; at each branch point the outcome (S or F) and its conditional probability; and on the right the sequence number, the end state (green when IT service is maintained, red when it is interrupted), the conditional sequence probability and the annual frequency computed in the next two subsections. The structure is that of Chapter 15; the branch probabilities differ. There, the UPS success probability 0.98 and the generator success probabilities 0.96 and 0.80 on the two UPS branches were assigned. Here they are derived from the linked fault trees: $\Pr(U \mid I) = 0.9702$, $\Pr(G \mid I, U) = 1 - q_\mathrm{g} = 0.96$, and $\Pr(G \mid I, \bar{U}) = 1 - 0.362148 = 0.637852$. The generator heading after a UPS failure is now much weaker than after a UPS success, because a UPS failure is evidence that the shared system has failed.

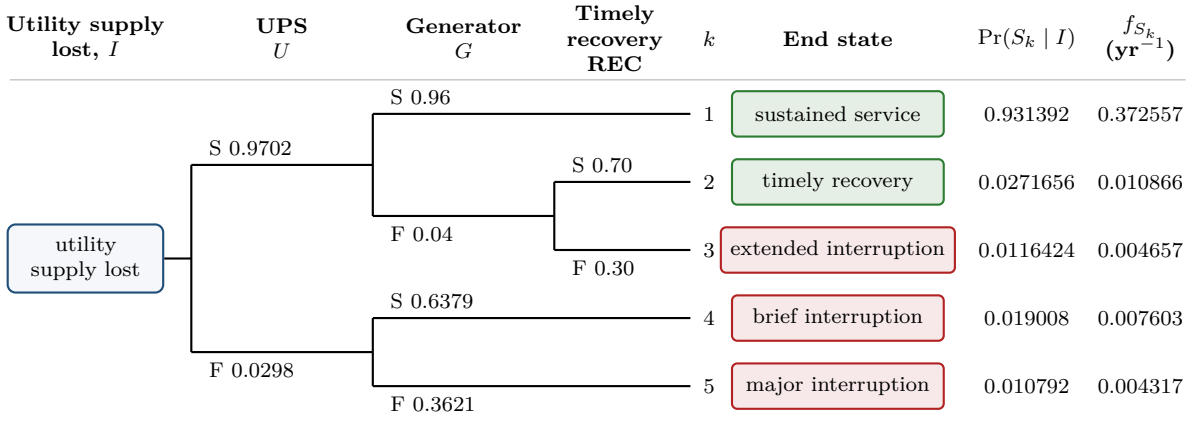


Figure 16.3: The integrated event tree for the loss of utility supply, $f_\mathrm{IE} = 0.40\,\mathrm{yr}^{-1}$ (illustrative). The UPS and generator branch probabilities are derived from the linked fault trees of Figure 16.2; the generator branch after a UPS failure, 0.6379 against 0.96, carries the shared cause. The sequence probabilities sum to 1 and the frequencies to $0.40\,\mathrm{yr}^{-1}$.

16.3.2 Sequence probabilities and the normalization check

Because X_{c}, X_{u} and X_{g} are independent, the probability of each sequence follows from its Boolean expression without counting the shared event twice:

$$\begin{aligned}
&\Pr(S_1 \mid I) = (1-q_{\mathrm{c}})(1-q_{\mathrm{u}})(1-q_{\mathrm{g}}) = 0.99 \times 0.98 \times 0.96 = 0.931392,\\
&\Pr(S_2 \mid I) = (1-q_{\mathrm{c}})(1-q_{\mathrm{u}})\,q_{\mathrm{g}}\,(1-q_{\mathrm{rec}}) = 0.99 \times 0.98 \times 0.04 \times 0.70 = 0.0271656,\\
&\Pr(S_3 \mid I) = (1-q_{\mathrm{c}})(1-q_{\mathrm{u}})\,q_{\mathrm{g}}\,q_{\mathrm{rec}} = 0.99 \times 0.98 \times 0.04 \times 0.30 = 0.0116424,\\
&\Pr(S_4 \mid I) = (1-q_{\mathrm{c}})\,q_{\mathrm{u}}\,(1-q_{\mathrm{g}}) = 0.99 \times 0.02 \times 0.96 = 0.019008,\\
&\Pr(S_5 \mid I) = q_{\mathrm{c}} + (1-q_{\mathrm{c}})\,q_{\mathrm{u}}\,q_{\mathrm{g}} = 0.010792.
\end{aligned} \tag{16.6}$$

The same numbers follow from the chain rule (15.6) along the branches of Figure 16.3: for instance $\Pr(S_4 \mid I) = \Pr(\bar{U} \mid I)\,\Pr(G \mid I, \bar{U}) = 0.0298 \times 0.637852 = 0.019008$.

The completeness check (15.10) holds numerically,

$$\sum_{k=1}^{5} \Pr(S_k \mid I) = 0.931392 + 0.0271656 + 0.0116424 + 0.019008 + 0.010792 = 1,$$

and it holds for every value of the parameters, which is the stronger check. The first three sequences add to $(1-q_{\mathrm{c}})(1-q_{\mathrm{u}})[(1-q_{\mathrm{g}}) + q_{\mathrm{g}}(1-q_{\mathrm{rec}}) + q_{\mathrm{g}}q_{\mathrm{rec}}] = (1-q_{\mathrm{c}})(1-q_{\mathrm{u}})$, the probability of a UPS success; the last two add to $(1-q_{\mathrm{c}})q_{\mathrm{u}}(1-q_{\mathrm{g}}) + q_{\mathrm{c}} + (1-q_{\mathrm{c}})q_{\mathrm{u}}q_{\mathrm{g}} = q_{\mathrm{c}} + (1-q_{\mathrm{c}})q_{\mathrm{u}}$, the probability of a UPS failure; and the two totals add to $(1-q_{\mathrm{c}}) + q_{\mathrm{c}} = 1$. The terminal sequences therefore form a complete partition of the outcomes, which is the first calculation to validate in any integrated event-tree model.

16.3.3 Annual sequence frequencies

Each conditional probability becomes an expected annual frequency by equation (15.7), $f_{S_k} = f_{\mathrm{IE}}\,\Pr(S_k \mid I)$ with $f_{\mathrm{IE}} = 0.40\,\mathrm{yr}^{-1}$. Table 16.2 lists, for each sequence, the outcome at each heading, the operational outcome, the conditional probability in percent and the expected frequency per year. The frequencies sum to $0.400000\,\mathrm{yr}^{-1}$, the initiating-event frequency, which is the second consistency check of equation (15.10).

Table 16.2: The five sequences of the integrated event tree, their conditional probabilities and their annual frequencies for $f_{\mathrm{IE}} = 0.40\,\mathrm{yr}^{-1}$ (illustrative).

k	UPS	Generator	Recovery	Operational outcome	$\Pr(S_k \mid I)$	f_{S_k} (yr^{-1})
1	succeeds	succeeds	not needed	sustained service	93.1392 %	0.372557
2	succeeds	fails	succeeds	service maintained through recovery	2.7166 %	0.010866
3	succeeds	fails	fails	extended IT interruption	1.1642 %	0.004657
4	fails	succeeds	not credited	brief IT interruption	1.9008 %	0.007603
5	fails	fails	not credited	major IT interruption	1.0792 %	0.004317
				Total	100 %	0.400000

The left panel of Figure 16.4 draws the frequencies as bars, one per sequence, with the frequency per year on the vertical axis and its value printed above each bar. Sequence 1

dominates by frequency: it carries 93 % of the initiating events, and the four other bars are barely visible beside it. Frequency alone, however, says little about the economic or physical risk until consequences are attached, which the right panel does in Section 16.3.6.

These frequencies are expected numbers of occurrences per year. They are not probabilities of at least one occurrence in a calendar year; converting one into the other needs a model of the occurrence times. Under the Poisson model of Section 16.2.2, each sequence type forms a thinned Poisson process with rate f_{S_k}, so the probability of at least one major interruption in a year is $1 - \mathrm{e}^{-0.0043168} = 0.0043075$, close to the frequency only because the frequency is small.

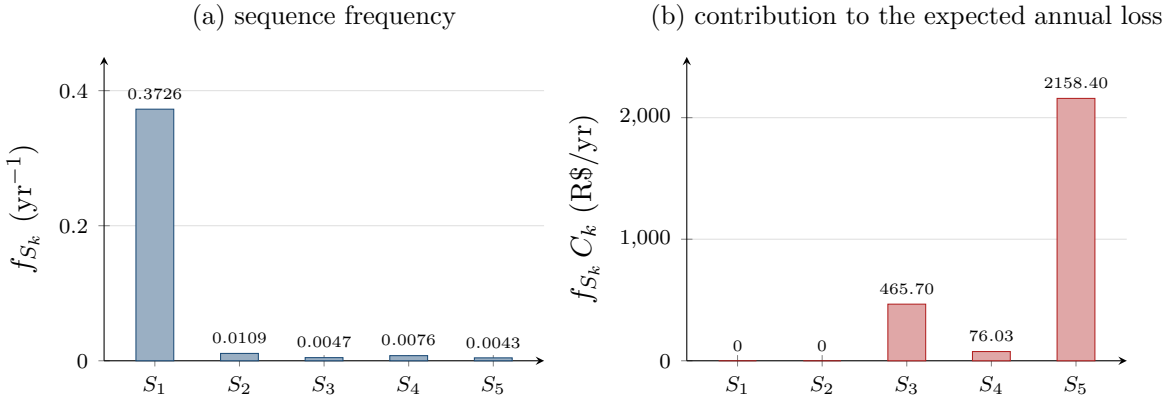


Figure 16.4: (a) Expected annual frequency of each sequence, $f_{S_k} = f_{\mathrm{IE}} \Pr(S_k \mid I)$, from Table 16.2. (b) Contribution of each sequence to the expected annual loss, $f_{S_k} C_k$, from Table 16.3. The most frequent sequence contributes nothing; the rarest of the interruption sequences contributes 80 %. Illustrative values.

16.3.4 Consequence severity

Each sequence is assigned an IT interruption duration D_k, in hours, and a financial loss per occurrence C_k, in reais, listed in the first columns of Table 16.3: no interruption and no loss for sequences 1 and 2; 2 h and R$ 100 000 for the extended interruption of sequence 3; 0.1 h and R$ 10 000 for the brief interruption of sequence 4; and 6 h and R$ 500 000 for the major interruption of sequence 5. They are the consequences of Table 15.3, kept so that the two chapters can be compared, and they are illustrative. The remaining columns of the table are the contributions of each sequence to the expected annual duration and loss computed below.

A real PRA derives these values from physical and operational analysis rather than assigning averages. For a data center the consequence of an interruption depends on the workload at the time, the contractual obligations to customers, the requirements for restarting the IT systems, any equipment or data damage, and the duration of the outage; each sequence's consequence is then a conditional expectation $\mathbb{E}[C \mid I, S_k]$ over those conditions, as equation (16.1) allows.

16.3.5 Expected interruption frequency and duration

Sequences 3, 4 and 5 interrupt the IT service. They are mutually exclusive, so the conditional probability of an interruption, given a loss of utility supply, is the sum of their probabilities,

$$\Pr(\text{interruption} \mid I) = 0.0116424 + 0.019008 + 0.010792 = 0.0414424,$$

and the expected annual frequency of interruptions is

$$f_{\text{int}} = f_{\mathrm{IE}} \Pr(\text{interruption} \mid I) = 0.40 \times 0.0414424 = 0.016577\,\text{yr}^{-1}.$$

Table 16.3: Consequences of the five sequences and their contributions to the expected annual interruption duration and the expected annual loss (illustrative).

k	Operational outcome	D_k (h)	C_k (R$)	f_{S_k} (yr^{-1})	$f_{S_k} D_k$ (h/yr)	$f_{S_k} C_k$ (R$/yr)
1	sustained service	0	0	0.3725568	0	0
2	timely recovery	0	0	0.0108662	0	0
3	extended interruption	2.0	100 000	0.0046570	0.0093139	465.70
4	brief interruption	0.1	10 000	0.0076032	0.0007603	76.03
5	major interruption	6.0	500 000	0.0043168	0.0259008	2158.40
	Total			0.4000000	0.0359750	2700.13

In reciprocal terms this is one interruption per $1/0.016577 = 60.3$ years on average. It is not a prediction of the time until the next interruption, and it covers only the modeled initiating-event family. The probability of at least one interruption in a given year is $1 - \mathrm{e}^{-0.016577} = 0.016440$.

The expected annual interruption duration is the frequency-weighted sum of the durations, by the argument of equation (15.11) and the random-sum identity that follows it:

$$\mathbb{E}[D_{\mathrm{yr}}] = \sum_k f_{S_k} D_k = 0.00465696 \times 2 + 0.0076032 \times 0.1 + 0.0043168 \times 6 = 0.035975 \,\mathrm{h} \text{ per year}, \tag{16.7}$$

the column total of Table 16.3: about 2.16 minutes of expected IT interruption per year attributable to this initiating-event family, or 0.0899 h per initiating event. A small expected annual duration does not imply that any individual outage is short: the major-interruption sequence lasts six hours, and in most years nothing happens at all.

16.3.6 Expected annual financial loss and the dominant sequence

The contribution of sequence k to the expected annual loss is $f_{S_k} C_k$, and the total is equation (15.22),

$$\mathcal{R} = \sum_k f_{S_k} C_k = 465.70 + 76.03 + 2158.40 = \mathrm{R\$}\, 2700.13 \text{ per year}, \tag{16.8}$$

the last column of Table 16.3, or R$ 6750.32 per initiating event. The right panel of Figure 16.4 draws the contributions in the same layout as the frequencies in the left panel. The picture inverts: sequence 1, which dominates the frequencies, contributes nothing, and sequence 5, with a conditional probability of only 1.0792 %, contributes $2158.40/2700.13 = 0.799$, about 80 % of the expected annual loss.

This is the important result of the integrated analysis, and Table 16.4 shows where it comes from. The table evaluates the same tree with the same consequences under three models of the UPS and generator headings: the linked fault trees of this chapter, in which the shared event is kept in the joint logic; the same fault trees quantified separately and treated as independent, with $\Pr(U) = 0.9702$ and $\Pr(G) = 0.9504$ on every branch; and the assigned branch probabilities of Chapter 15. Its rows are the five sequence probabilities and three annual outputs. The independence model puts 0.00148 on the major interruption instead of 0.010792 and computes an expected loss of R$ 986.37 per year, 63 % below the correct R$ 2700.13; at the same time it overstates the interruption frequency by 7 %, because the probability it removes from sequence 5 reappears in the brief interruption of sequence 4. The tree of Chapter 15, with its assigned conditional probability 0.80 for the generator after a UPS failure, captured part of the dependence and lies between the two.

Table 16.4: The data center event tree under three models of the UPS and generator headings, with $f_{\mathrm{IE}} = 0.40\,\mathrm{yr}^{-1}$ and the consequences of Table 16.3 (illustrative).

Quantity	Linked fault trees, shared event	Separate fault trees, independence	Assigned branches, Chapter 15
$\Pr(S_1 \mid I)$, sustained service	0.931392	0.922078	0.940800
$\Pr(S_2 \mid I)$, timely recovery	0.027166	0.033685	0.027440
$\Pr(S_3 \mid I)$, extended interruption	0.011642	0.014437	0.011760
$\Pr(S_4 \mid I)$, brief interruption	0.019008	0.028322	0.016000
$\Pr(S_5 \mid I)$, major interruption	0.010792	0.001478	0.004000
Interruption frequency f_{int} (yr^{-1})	0.016577	0.017695	0.012704
Expected interruption $\mathbb{E}[D_{\mathrm{yr}}]$ (min/yr)	2.159	0.974	1.179
Expected annual loss $\mathcal{R}$ (R$/yr)	2700.13	986.37	1334.40

Principle 16.2. *Expected risk depends on the severity of consequences as well as on frequency. A rare sequence with a severe consequence, here the joint failure produced by a shared cause, can dominate the expected loss, and a model that misses the dependence misses the dominant term.*

16.4 Risk importance and engineering decisions

With the consequences attached, the integrated model expresses the expected annual loss as an explicit function of the basic-event probabilities. This section writes that function, ranks the failure mechanisms by their effect on it, evaluates four engineering improvements by recomputing the whole tree, and weighs the risk reduction of one against its investment.

16.4.1 The risk function

Substituting the sequence probabilities (16.6) into $\mathcal{R} = f_{\mathrm{IE}} \sum_k \Pr(S_k \mid I)\, C_k$, with the consequences of Table 16.3, gives the expected annual loss as an analytic expression in the five parameters:

$$\begin{aligned}\mathcal{R}(f_{\mathrm{IE}}, q_{\mathrm{u}}, q_{\mathrm{g}}, q_{\mathrm{c}}, q_{\mathrm{rec}}) = f_{\mathrm{IE}}\big[\,&100\,000\,(1-q_{\mathrm{c}})(1-q_{\mathrm{u}})\,q_{\mathrm{g}}\,q_{\mathrm{rec}} + 10\,000\,(1-q_{\mathrm{c}})\,q_{\mathrm{u}}\,(1-q_{\mathrm{g}}) \\ &+ 500\,000\,\big(q_{\mathrm{c}} + (1-q_{\mathrm{c}})\,q_{\mathrm{u}}\,q_{\mathrm{g}}\big)\big],\end{aligned} \tag{16.9}$$

in reais per year. At the baseline values it returns $0.40 \times (1164.24 + 190.08 + 5396) = 0.40 \times 6750.32 = \mathrm{R\$}\,2700.13$. The expression connects the component probabilities directly to the expected consequence; it is the complete model of this simplified PRA, and every result of the rest of the chapter is obtained by evaluating or differentiating it.

The function is *multilinear*: it is a polynomial in which each parameter appears at most to the first power in every term. Two properties follow. With all parameters but one held fixed, $\mathcal{R}$ is a straight line in that parameter, so a change of one parameter has an exactly linear effect. And $\mathcal{R}$ is proportional to f_{IE}: doubling the frequency of utility losses doubles the expected loss.

16.4.2 Risk sensitivity and risk importance

The *risk sensitivity* to a basic-event probability q_i is the partial derivative $\partial\mathcal{R}/\partial q_i$, in reais per year per unit of probability. It is the analogue, for the expected loss, of the Birnbaum importance of Section 12.2.1, which is the derivative of the top-event probability; the risk sensitivity weights every sequence by its consequence instead of counting only whether the top event occurs. Differentiating equation (16.9) term by term,

$$\begin{aligned}
\frac{\partial\mathcal{R}}{\partial q_{\mathrm{c}}} &= f_{\mathrm{IE}}\big[-100\,000\,(1-q_{\mathrm{u}})q_{\mathrm{g}}q_{\mathrm{rec}} - 10\,000\,q_{\mathrm{u}}(1-q_{\mathrm{g}}) + 500\,000\,(1-q_{\mathrm{u}}q_{\mathrm{g}})\big],\\
\frac{\partial\mathcal{R}}{\partial q_{\mathrm{g}}} &= f_{\mathrm{IE}}(1-q_{\mathrm{c}})\big[100\,000\,(1-q_{\mathrm{u}})q_{\mathrm{rec}} - 10\,000\,q_{\mathrm{u}} + 500\,000\,q_{\mathrm{u}}\big],\\
\frac{\partial\mathcal{R}}{\partial q_{\mathrm{u}}} &= f_{\mathrm{IE}}(1-q_{\mathrm{c}})\big[-100\,000\,q_{\mathrm{g}}q_{\mathrm{rec}} + 10\,000\,(1-q_{\mathrm{g}}) + 500\,000\,q_{\mathrm{g}}\big],\\
\frac{\partial\mathcal{R}}{\partial q_{\mathrm{rec}}} &= f_{\mathrm{IE}}\,100\,000\,(1-q_{\mathrm{c}})(1-q_{\mathrm{u}})\,q_{\mathrm{g}}.
\end{aligned} \tag{16.10}$$

At the baseline, for the shared event, $\partial\mathcal{R}/\partial q_{\mathrm{c}} = 0.40\times(-1176-192+499\,600) = 0.40\times 498\,232 = 199\,292.80$ reais per year per unit probability; the others are $\partial\mathcal{R}/\partial q_{\mathrm{g}} = 0.40\times 0.99\times 39\,200 = 15\,523.20$, $\partial\mathcal{R}/\partial q_{\mathrm{u}} = 0.40\times 0.99\times 28\,400 = 11\,246.40$ and $\partial\mathcal{R}/\partial q_{\mathrm{rec}} = 0.40\times 100\,000\times 0.9702\times 0.04 = 1552.32$. The negative terms have a physical meaning: a shared-cause failure removes the possibility of the milder sequences 3 and 4, so part of what it adds to sequence 5 it takes from them.

Because $\mathcal{R}$ is linear in each parameter separately, a change Δq_i of one parameter changes the expected loss by exactly $\Delta q_i\,\partial\mathcal{R}/\partial q_i$, not approximately: lowering q_{c} by 0.01 lowers $\mathcal{R}$ by R\$ 1992.93 per year, provided the other parameters and the consequences are unchanged. The linearity also gives the importance measures of Chapter 12 computed on the risk. Writing $\mathcal{R}(q_i=0)$ and $\mathcal{R}(q_i=1)$ for the risk with the event made impossible or certain, linearity gives $\mathcal{R} = \mathcal{R}(q_i=0) + q_i\,\partial\mathcal{R}/\partial q_i$ and $\partial\mathcal{R}/\partial q_i = \mathcal{R}(q_i=1) - \mathcal{R}(q_i=0)$, so that

$$I_i^{\mathrm{FV}} = \frac{\mathcal{R}-\mathcal{R}(q_i=0)}{\mathcal{R}} = \frac{q_i}{\mathcal{R}}\frac{\partial\mathcal{R}}{\partial q_i}, \qquad \mathrm{RAW}_i = \frac{\mathcal{R}(q_i=1)}{\mathcal{R}}, \qquad \mathrm{RRW}_i = \frac{\mathcal{R}}{\mathcal{R}(q_i=0)}, \tag{16.11}$$

the definitions of equations (12.18), (12.21) and (12.22) with the risk in place of the top-event probability. The Fussell–Vesely importance on the risk is the fraction of the expected loss removed by eliminating the event, and it equals the elasticity of the risk, the relative change of $\mathcal{R}$ per relative change of q_i.

Table 16.5 evaluates all of them. Its columns are the baseline probability, the risk sensitivity, the risk with the event impossible and certain, and the three ratios of equation (16.11). The shared event leads on every measure: eliminating it removes 73.8 % of the expected loss, and making it certain multiplies the loss by 74, since every initiating event then ends in a major interruption, $\mathcal{R} = 0.40\times 500\,000 =$ R\$ 200 000 per year. Below it the measures disagree. The recovery failure has the smallest sensitivity and the smallest RAW, yet its Fussell–Vesely importance, 17.2 %, is twice that of the UPS-specific failure, because its baseline probability, 0.30, is fifteen times larger. Sensitivity asks what a unit change of probability does; Fussell–Vesely asks what the present probability contributes; RAW asks what a failed or removed item would cost. These derivatives and ratios cannot be compared as costs of intervention without the feasible parameter changes and the implementation expenses, which the next two subsections add.

16.4.3 Four engineering improvements

Four improvements are under consideration, each halving one probability:

Table 16.5: Risk sensitivity and importance measures of the four probabilities on the expected annual loss, from equations (16.10) and (16.11) at the baseline $\mathcal{R}$ = R$ 2700.13 per year (illustrative).

Basic event	q_i	$\partial\mathcal{R}/\partial q_i$ (R$/yr)	$\mathcal{R}(q_i=0)$ (R$/yr)	$\mathcal{R}(q_i=1)$ (R$/yr)	I_i^{FV}	RAW_i	RRW_i
Shared cause, q_{c}	0.01	199 292.80	707.20	200 000.00	0.738	74.07	3.818
Generator-specific, q_{g}	0.04	15 523.20	2079.20	17 602.40	0.230	6.519	1.299
UPS-specific, q_{u}	0.02	11 246.40	2475.20	13 721.60	0.083	5.082	1.091
Recovery failure, q_{rec}	0.30	1552.32	2234.43	3786.75	0.172	1.402	1.208

- option A: the shared-cause probability q_{c} from 1 % to 0.5 %;
- option B: the generator-specific probability q_{g} from 4 % to 2 %;
- option C: the UPS-specific probability q_{u} from 2 % to 1 %;
- option D: the recovery failure probability q_{rec} from 30 % to 15 %.

For each option every sequence probability of equation (16.6) is recomputed, and with it the annual outputs, as the decision formula (12.30) requires. For option D, for example, only sequences 2 and 3 change: $\Pr(S_3 \mid I) = 0.9702 \times 0.04 \times 0.15 = 0.0058212$ and $\Pr(S_2 \mid I) = 0.9702 \times 0.04 \times 0.85 = 0.0329868$, so $\mathcal{R} = 0.40 \times (582.12 + 190.08 + 5396) = 0.40 \times 6168.20 =$ R$ 2467.28 per year, a reduction of R$ 232.85, which is 0.15×1552.32, as linearity requires.

Table 16.6 gives the four results. Its columns are the expected annual loss, the annual risk reduction from the baseline and its fraction of the baseline, the interruption frequency, the expected interruption minutes per year, and the present value of ten years of risk reduction derived in Section 16.4.4. Figure 16.5 draws the annual risk reduction of each option as a horizontal bar, with the value printed at the end of each bar.

Table 16.6: Four improvements, each halving one probability, evaluated by recomputing the integrated event tree; $f_{\mathrm{IE}} = 0.40\,\mathrm{yr}^{-1}$ (illustrative).

Probability halved	$\mathcal{R}$ (R$/yr)	$\Delta\mathcal{R}$ (R$/yr)	Relative	f_{int} (yr^{-1})	$\mathbb{E}[D_{\mathrm{yr}}]$ (min/yr)	PV of $\Delta\mathcal{R}$, 10 yr (R$)
Baseline	2700.13	–	–	0.016577	2.159	–
A: shared cause q_{c}	1703.66	996.46	36.9 %	0.014640	1.442	7694
B: generator-specific q_{g}	2389.66	310.46	11.5 %	0.014248	1.823	2397
C: UPS-specific q_{u}	2587.66	112.46	4.2 %	0.012664	2.084	868
D: recovery failure q_{rec}	2467.28	232.85	8.6 %	0.014248	1.879	1798

Under these assumptions, halving the shared-cause probability gives the largest reduction of the expected loss, R$ 996.46 per year, more than the other three options together (R$ 655.78). It acts on the major-interruption sequence that dominates the loss. The ranking depends on the consequence measure, however. By interruption frequency the best option is C, the UPS improvement, which reduces the frequent brief interruptions of sequence 4 and leaves the fewest interruptions, $0.012664\,\mathrm{yr}^{-1}$; option A is the worst by that measure. By expected interruption minutes option A is again first. A decision that counts interruptions, as a service-level agreement with a penalty per event might, and one that counts lost revenue would choose differently.

The options are not additive. Applying A and B together gives $\mathcal{R}$ = R$ 1391.63 per year, a reduction of R$ 1308.50, not the sum $996.46 + 310.46 =$ R$ 1306.93: the risk is linear in each parameter but contains products of parameters, and the cross term $\Delta q_{\mathrm{c}}\,\Delta q_{\mathrm{g}}\,\partial^2\mathcal{R}/\partial q_{\mathrm{c}}\partial q_{\mathrm{g}}$ adds R$ 1.57. Here the interaction is small; with larger changes or stronger coupling it need not be, and combinations must be evaluated on the model.

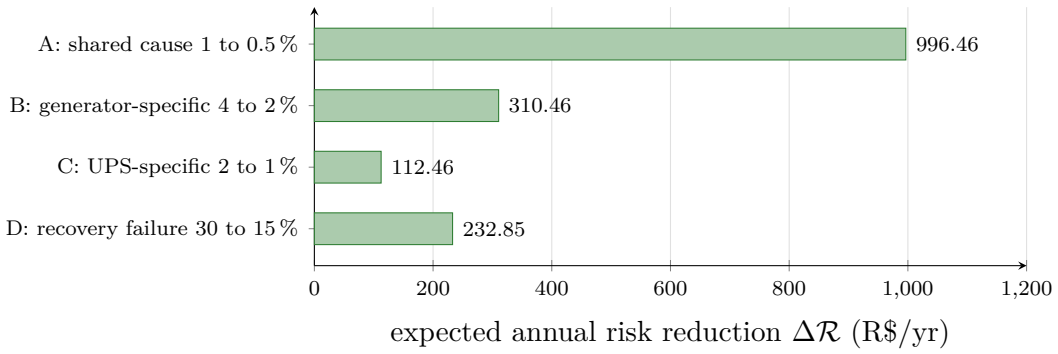


Figure 16.5: Expected annual risk reduction of the four improvements of Table 16.6, each halving one probability of the integrated model; $\Delta\mathcal{R} = \Delta q_i \, \partial\mathcal{R}/\partial q_i$ exactly, because the risk is linear in each probability. Illustrative values.

16.4.4 Risk reduction against investment value

A reduction of the expected loss is not yet an investment decision. The decision needs the cost of the improvement and the period over which its benefit lasts, the terms of the net benefit (12.31). Suppose option A requires an up-front investment of R$ 5000, its annual risk reduction $\Delta\mathcal{R}_A$ = R$ 996.46 persists for ten years, money is discounted at an annual rate $\rho = 5\,\%$, and maintenance and replacement costs are ignored. The present value of the risk reduction is the discounted sum

$$\mathrm{PV}_A = \sum_{y=1}^{10} \frac{\Delta\mathcal{R}_A}{(1+\rho)^y}. \tag{16.12}$$

With $x = 1/(1+\rho)$, the sum is the geometric series $\sum_{y=1}^{n} x^y = x\,(1-x^n)/(1-x)$, and since $x/(1-x) = 1/\rho$ it equals $\Delta\mathcal{R}_A$ times the annuity factor

$$\sum_{y=1}^{n} (1+\rho)^{-y} = \frac{1-(1+\rho)^{-n}}{\rho} = \frac{1-1.05^{-10}}{0.05} = \frac{1-0.613913}{0.05} = 7.7217. \tag{16.13}$$

Therefore $\mathrm{PV}_A = 996.46 \times 7.7217 =$ R$ 7694, and the net present value of the investment is

$$\mathrm{NPV}_A = \mathrm{PV}_A - C_{\mathrm{impl}}(A) = 7694 - 5000 = +\text{R\$ } 2694.$$

This is the net benefit of equation (12.31) with the benefits of later years discounted to the present. Under these assumptions option A has a positive expected net value.

The last column of Table 16.6 gives the same present value for every option. It is the largest investment that each improvement could justify on expected loss alone over ten years: R$ 7694 for A, R$ 2397 for B, R$ 1798 for D and R$ 868 for C. An option whose cost exceeds this value has a negative expected net value, whatever its rank by importance.

A positive net present value is not sufficient evidence to approve a real design change. The benefit rests on uncertain failure probabilities, on assumed consequences, on the effectiveness of the mitigation (whether the modification actually halves q_c), and on the assumption that the modification introduces no new dependence or failure mode. The first of these is the subject of the next section.

Principle 16.3. *Engineering improvements are evaluated as interventions on the full system model: each one is applied to the basic events or the logic, the whole event tree is recomputed, and the resulting risk reduction is weighed against its cost over the period of its benefit. Importance and sensitivity rankings identify where to look, not what to buy.*

16.5 Uncertainty propagation through the integrated model

The calculations so far treated every input as a known number. This section treats the inputs as estimates, propagates their uncertainty to the expected annual loss by a first-order expansion and by Monte Carlo sampling, separates the two levels of randomness in the result, and ends with a worked table of the expected loss under several strengths of the dependence and several improvements.

16.5.1 Probabilities as estimates

The probabilities q_{u}, q_{g}, q_{c} and q_{rec} are not known constants. They are estimated from limited observations, engineering studies, tests, operating history or expert judgment, by the methods of Chapters 7 and 8. The shared-cause probability q_{c} is usually the least certain, because shared-cause events are rare and a site may never have observed one. The initiating-event frequency and the consequence values are estimates as well.

Collect the uncertain inputs in a parameter vector

$$\boldsymbol{\theta} = (q_{\mathrm{u}}, q_{\mathrm{g}}, q_{\mathrm{c}}, q_{\mathrm{rec}}, f_{\mathrm{IE}}, \mathbf{c}),$$

where $\mathbf{c}$ holds the parameters of the consequence model. The expected annual loss is a function of it, $\mathcal{R} = \mathcal{R}(\boldsymbol{\theta})$, given by equation (16.9), and the uncertainty of $\boldsymbol{\theta}$ propagates through that function into uncertainty about $\mathcal{R}$. In what follows the consequences are held fixed, so that the propagation concerns the five probabilistic parameters.

16.5.2 First-order uncertainty propagation

Let $\bar{\boldsymbol{\theta}} = \mathbb{E}[\boldsymbol{\theta}]$ be the vector of parameter means and $\Sigma_{\boldsymbol{\theta}}$ the covariance matrix of the parameters. For small parameter uncertainty, a first-order Taylor expansion of the risk about the means gives

$$\mathcal{R}(\boldsymbol{\theta}) \approx \mathcal{R}(\bar{\boldsymbol{\theta}}) + \nabla\mathcal{R}^{\mathsf{T}} \, (\boldsymbol{\theta} - \bar{\boldsymbol{\theta}}),$$

where $\nabla\mathcal{R}$ is the gradient of sensitivities $\partial\mathcal{R}/\partial\theta_i$ evaluated at the means. The right-hand side is a constant plus a linear function of the random vector $\boldsymbol{\theta} - \bar{\boldsymbol{\theta}}$, whose mean is zero and whose covariance is $\Sigma_{\boldsymbol{\theta}}$. The variance of a linear function $\mathbf{a}^{\mathsf{T}}\mathbf{x}$ is $\mathbf{a}^{\mathsf{T}} \operatorname{Var}(\mathbf{x})\, \mathbf{a}$, so

$$\operatorname{Var}(\mathcal{R}) \approx \nabla\mathcal{R}^{\mathsf{T}} \, \Sigma_{\boldsymbol{\theta}} \, \nabla\mathcal{R}, \tag{16.14}$$

which is equation (12.32) with the risk in place of the top-event probability. If the parameters are approximately independent, $\Sigma_{\boldsymbol{\theta}}$ is diagonal and

$$\operatorname{Var}(\mathcal{R}) \approx \sum_{i} \left(\frac{\partial \mathcal{R}}{\partial \theta_i} \right)^2 \operatorname{Var}(\theta_i). \tag{16.15}$$

Each term is the contribution of one parameter: the square of the sensitivity of the risk to it times the variance of its estimate. The sensitivity analysis of Section 16.4.2 thus becomes an uncertainty analysis once the variances are known.

To quantify it, the parameters are given illustrative distributions whose means are the baseline values. A per-demand failure probability is given a beta distribution $\mathrm{Beta}(a, b)$, with density proportional to $q^{a-1}(1-q)^{b-1}$ on $0 < q < 1$, mean $a/(a+b)$ and variance $ab/[(a+b)^2(a+b+1)]$. It is the standard conjugate distribution for a probability estimated from a failures in $a+b$ demands, so its parameters can be read as pseudo-counts. The initiating-event frequency is given the gamma distribution of Chapter 8, $\mathrm{Gamma}(4, 10\,\mathrm{yr})$, with mean $4/10 = 0.40\,\mathrm{yr}^{-1}$ and standard deviation $\sqrt{4}/10 = 0.20\,\mathrm{yr}^{-1}$, the posterior that four utility losses in ten years of records would give under a vague prior. Table 16.7 lists

the five distributions with their means and standard deviations, the sensitivity of the risk to each parameter, and its contribution to the first-order variance. For the shared event the contribution is $(199\,292.80 \times 0.009900)^2 = 1973.1^2 = 3.893 \times 10^6$; for the frequency it is $(6750.32 \times 0.20)^2 = 1350.1^2 = 1.823 \times 10^6$, the sensitivity to f_{IE} being $\mathcal{R}/f_{\mathrm{IE}} = 6750.32$ because $\mathcal{R}$ is proportional to it.

Table 16.7: Illustrative parameter distributions and the first-order variance of the expected annual loss, equation (16.15), with independent parameters and fixed consequences.

Parameter	Distribution	Mean	Standard deviation	$\partial\mathcal{R}/\partial\theta_i$	Contribution (R\$/yr)^2	Share
f_{IE} (yr^{-1})	Gamma(4, 10 yr)	0.40	0.2000	6750.32	1.823×10^6	31.0 %
q_{u}	Beta(2, 98)	0.02	0.0139	11 246.40	2.454×10^4	0.4 %
q_{g}	Beta(4, 96)	0.04	0.0195	15 523.20	9.162×10^4	1.6 %
q_{c}	Beta(1, 99)	0.01	0.0099	199 292.80	3.893×10^6	66.2 %
q_{rec}	Beta(3, 7)	0.30	0.1382	1552.32	4.600×10^4	0.8 %
Total					5.878×10^6	100 %

The first-order variance is 5.878×10^6, a standard deviation of R\$ 2424 per year against a point estimate of R\$ 2700: the expected loss is known only to within a factor of about two. The shared cause carries two thirds of the variance and the initiating-event frequency almost a third; the recovery failure, whose probability has the largest standard deviation of all, contributes under one percent, because the risk is insensitive to it. Two different notions must therefore be kept apart. A parameter may have high risk importance because the risk function is sensitive to it; another may dominate the uncertainty of the risk because its estimate is especially imprecise. Here the shared cause is both, which makes better evidence about it the most valuable information the assessment could acquire; that question, the value of information, returns in a later chapter.

The multilinear form of the risk function gives two further results. First, when the parameters are independent the mean of the risk is the risk at the mean parameters,

$$\mathbb{E}[\mathcal{R}(\boldsymbol{\theta})] = \mathcal{R}(\bar{\boldsymbol{\theta}}), \tag{16.16}$$

because each term of equation (16.9) is a product of distinct independent parameters, each to the first power, and the expectation of such a product is the product of the expectations. The point estimate R\$ 2700.13 is therefore also the mean of the uncertain risk; that would not hold for a model with squared parameters, such as a 2-out-of-3 group of identical units with a shared estimated probability. Second, the first-order variance is not exact, because the products of parameters contribute terms that the linearization drops: for two independent factors, $\mathrm{Var}(XY) = \bar{X}^2\,\mathrm{Var}(Y) + \bar{Y}^2\,\mathrm{Var}(X) + \mathrm{Var}(X)\,\mathrm{Var}(Y)$, and the expansion keeps only the first two. The exact variance, computed from the moments of the five distributions, is 6.909×10^6, a standard deviation of R\$ 2628; the first-order value understates it by 8 %, nearly all of the gap coming from the product $f_{\mathrm{IE}}\,q_{\mathrm{c}}$ of the two most uncertain inputs, whose omitted term is $(500\,000 \times 0.20 \times 0.0099)^2 = 0.98 \times 10^6$.

16.5.3 Monte Carlo uncertainty propagation

A more general method samples the parameters from their distributions and evaluates the model for each sample. With the distributions of Table 16.7, the procedure is:

1. draw a parameter vector $\boldsymbol{\theta}^{(j)}$, each component from its distribution;
2. compute the linked fault-tree probabilities $\Pr(\bar{U})$, $\Pr(\bar{G})$ and $\Pr(\bar{U} \cap \bar{G})$;
3. compute the five sequence probabilities (16.6);

4. compute $\mathcal{R}^{(j)} = \mathcal{R}(\boldsymbol{\theta}^{(j)})$ from equation (16.9);
5. repeat for $j = 1, \ldots, J$.

The sample $\mathcal{R}^{(1)}, \ldots, \mathcal{R}^{(J)}$ approximates the distribution of the expected annual loss induced by the parameter uncertainty. The method needs no linearization, handles any distribution and any dependence among the parameters, and its error falls as $1/\sqrt{J}$. The shared event must be sampled once per draw and used in both fault trees, exactly as it is counted once in the joint logic.

With $J = 200\,000$ draws, the sample mean is R\$ 2700 (standard error R\$ 6), in agreement with equation (16.16), and the sample standard deviation is R\$ 2629, in agreement with the exact value and 8 % above the first-order one. The distribution is strongly skewed: the median is R\$ 1895, below the mean; the 5th and 95th percentiles are R\$ 400 and R\$ 7716 per year, a factor of nineteen apart. Figure 16.6(a) draws the histogram. The horizontal axis is the expected annual loss $\mathcal{R}$ in reais per year, in bins of R\$ 500; the vertical axis is the fraction of the draws in each bin; the dashed vertical line marks the mean R\$ 2700 and the dotted line the median. The long right tail comes from draws in which both f_{IE} and q_{c} are high; 0.4 % of the draws lie beyond the plotted range of R\$ 16 000.

The sample supports statements such as

$$\Pr(\mathcal{R} > \text{R\$ } 5000 \text{ per year} \mid \mathcal{D}) = 0.134,$$

where $\mathcal{D}$ denotes the evidence on which the parameter distributions rest: given that evidence, the expected annual loss exceeds R\$ 5000 with probability 13.4 %, and it exceeds the point estimate with probability 34.7 %. Such a statement is a Bayesian (or otherwise probabilistic) statement of uncertainty about a model quantity, and it is only as good as the justification of the distributions; those of Table 16.7 are illustrative choices, not calibrated to evidence.

16.5.4 Aleatory variability and epistemic uncertainty

A PRA has two sources of variation, and they must not be confused. *Aleatory variability* is the randomness of future events under a model with given parameters: whether a utility loss occurs this year, whether the generator starts on a particular demand, which sequence follows. *Epistemic uncertainty* is the incomplete knowledge of the parameters and of the model itself: how large the generator failure probability really is, or whether one aggregate shared event represents the coupling correctly. The first is reduced by nothing but a change of the system; the second is reduced by evidence.

The distribution of Figure 16.6(a) is epistemic. It is a distribution over the *expected* annual loss $\mathcal{R}$, a property of the system, and it is not the distribution of the loss C_{yr} actually realized in a year. The realized loss also depends on the number and the outcomes of the initiating events in that year. With the baseline parameters, the interruption sequences 3, 4 and 5 form independent Poisson processes with the rates of Table 16.2, so the probability of a year without any loss is $\mathrm{e}^{-f_{\mathrm{int}}} = \mathrm{e}^{-0.016577} = 0.98356$. In 98.4 % of years the realized loss is zero; in the others it is R\$ 10 000, R\$ 100 000, R\$ 500 000 or a sum of these. The expected value R\$ 2700 is never the realized loss of any year.

Figure 16.6(b) draws this aleatory distribution as an exceedance curve. The horizontal axis is a loss threshold c in reais, logarithmic; the vertical axis is the probability that the realized annual loss exceeds c, logarithmic. The solid blue step is computed with the baseline parameters: 0.01644 for any loss, 0.00896 above R\$ 10 000, 0.00435 above R\$ 100 000 and 6.2×10^{-5} above R\$ 500 000, the last being the probability of a major interruption together with another interruption in the same year. The dashed red step averages the same probability over the epistemic distribution of the parameters, from the first 20 000 Monte Carlo draws; it is the predictive exceedance probability, which combines both levels. The two curves nearly coincide for common outcomes and separate in the tail: above R\$ 500 000 the predictive probability is

8.8×10^{-5}, 42 % higher, because draws with a high shared-cause probability make a major interruption together with a second interruption in the same year more likely. Parameter uncertainty matters most for the rare severe outcomes, which are also the ones that decide whether a loss can be absorbed.

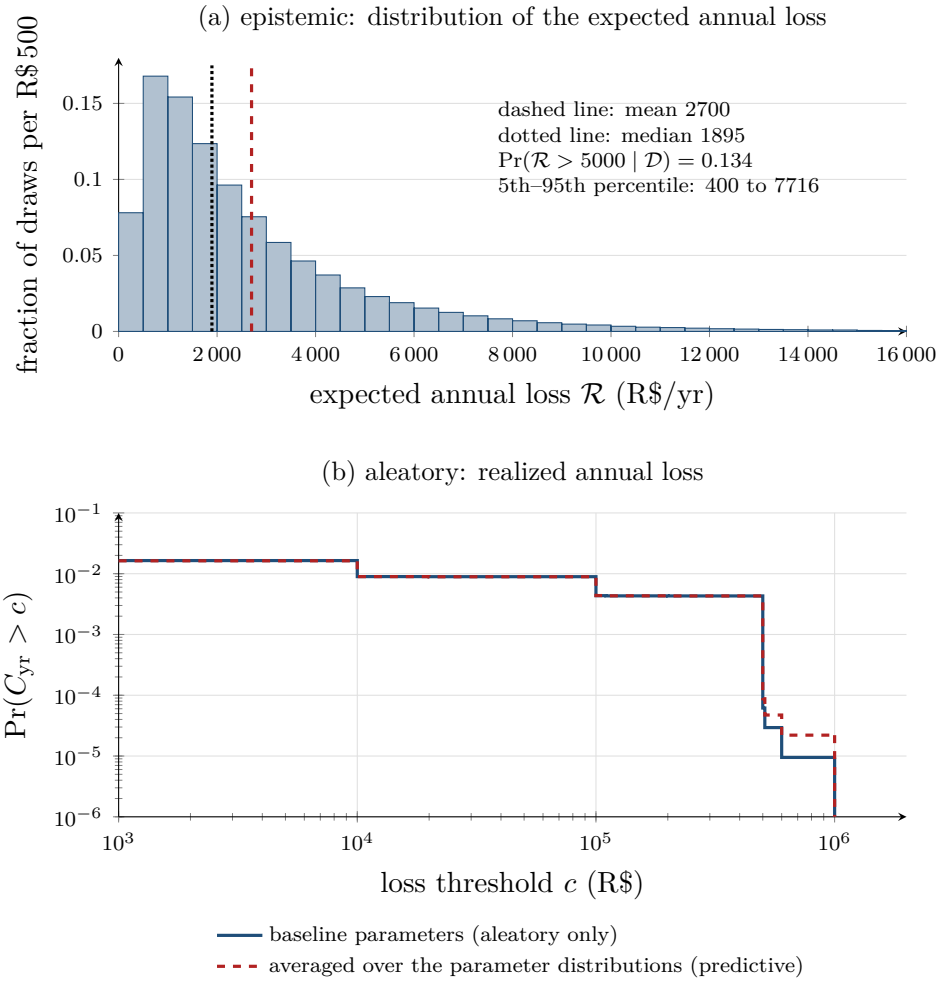


Figure 16.6: The two levels of randomness. (a) Histogram of $J = 200\,000$ Monte Carlo draws of the expected annual loss $\mathcal{R}(\boldsymbol{\theta})$, equation (16.9), with the parameter distributions of Table 16.7. (b) Probability that the loss realized in one year exceeds c, for the baseline parameters and averaged over the parameter distributions; a drop to the bottom edge means the probability falls below 10^{-6}. Illustrative distributions.

A defensible uncertainty analysis keeps the two levels separate and says which one each reported number describes: the point estimate and the histogram describe the expected loss and what is known about it; the exceedance curve describes what can happen in a year. The mean of the predictive distribution of the realized loss is again R$ 2700, by equation (16.16) and the linearity of expectation, so the two levels agree on the average and differ in everything else.

16.5.5 Annual loss under dependence strength and improvements: a worked table

Because the shared-cause probability is both the most important and the least certain parameter, the decision must hold across its plausible range. The integrated model of equations (16.6) and (16.9), with the shared event kept in the joint probabilities, $f_{\mathrm{IE}} = 0.40\,\mathrm{yr}^{-1}$ and the fixed consequences 0, 0, R$ 100 000, R$ 10 000 and R$ 500 000 for sequences 1 to 5, is evaluated in

Table 16.8 for six values of q_{c}, from no shared cause to 5 %, and for five improvement cases: none; option B, C or D of Section 16.4.3; and the three together. Its columns are the shared-cause probability, the major-interruption probability, the interruption frequency and the expected interruption minutes per year without improvement, and then the expected annual loss for each improvement case. The row $q_{\mathrm{c}} = 0.01$ is the baseline, and the row $q_{\mathrm{c}} = 0.005$ without improvement is option A.

Table 16.8: Expected annual loss of the integrated model for several strengths of the shared cause and several improvements; $q_{\mathrm{u}} = 0.02$, $q_{\mathrm{g}} = 0.04$, $q_{\mathrm{rec}} = 0.30$ unless improved, $f_{\mathrm{IE}} = 0.40\,\mathrm{yr}^{-1}$ (illustrative). B halves q_{g}, C halves q_{u}, D halves q_{rec}.

	No improvement			Expected annual loss $\mathcal{R}$ (R$/yr)				
q_{c}	$\Pr(S_5 \mid I)$	f_{int} (yr^{-1})	$\mathbb{E}[D_{\mathrm{yr}}]$ (min/yr)	none	B	C	D	B, C and D
0	0.000800	0.012704	0.726	707.20	393.60	593.60	472.00	198.00
0.0025	0.003298	0.013672	1.084	1205.43	892.62	1092.12	970.82	697.51
0.005	0.005796	0.014640	1.442	1703.66	1391.63	1590.63	1469.64	1197.01
0.01	0.010792	0.016577	2.159	2700.13	2389.66	2587.66	2467.28	2196.02
0.02	0.020784	0.020450	3.591	4693.06	4385.73	4581.73	4462.56	4194.04
0.05	0.050760	0.032069	7.889	10 671.84	10 373.92	10 563.92	10 448.40	10 188.10

One row of the arithmetic, $q_{\mathrm{c}} = 0.05$ without improvement: $\Pr(S_3 \mid I) = 0.95 \times 0.98 \times 0.04 \times 0.30 = 0.011172$, $\Pr(S_4 \mid I) = 0.95 \times 0.02 \times 0.96 = 0.01824$ and $\Pr(S_5 \mid I) = 0.05 + 0.95 \times 0.02 \times 0.04 = 0.05076$, so $f_{\mathrm{int}} = 0.40 \times 0.080172 = 0.032069\,\mathrm{yr}^{-1}$ and $\mathcal{R} = 0.40 \times (1117.20 + 182.40 + 25\,380) = 0.40 \times 26\,679.60 = \mathrm{R\$}\,10\,671.84$ per year.

Three trends follow. First, the expected loss grows almost linearly with the shared-cause probability, by about R$ 2000 per year for each percentage point, from R$ 707 with no shared cause to R$ 10 672 at 5 %, a factor of fifteen; over the same range the interruption frequency grows only by a factor of 2.5, because the shared cause converts interruptions that would have been brief into major ones more than it creates new ones. The share of the major interruption in the loss rises from 23 % at $q_{\mathrm{c}} = 0$ to 95 % at 5 %. Second, improvements B, C and D remove nearly the same absolute amount at every q_{c}, about R$ 300, R$ 110 and R$ 230 per year, because they act on the independent mechanisms; their relative value therefore collapses as the dependence strengthens. Halving the generator-specific failure removes 44 % of the loss when there is no shared cause and 3 % when $q_{\mathrm{c}} = 0.05$. Third, at the baseline the three independent improvements together reach R$ 2196 per year, while halving the shared cause alone reaches R$ 1704: no combination of improvements to the independent mechanisms substitutes for reducing the dependence. Every entry is proportional to f_{IE}; at $0.20\,\mathrm{yr}^{-1}$ or $0.80\,\mathrm{yr}^{-1}$ the baseline loss is R$ 1350.06 or R$ 5400.26 per year, and the rankings do not change.

Principle 16.4. *Uncertainty is propagated to the final decision quantities. A precise risk estimate is credible only when its inputs are defensible and the decision it supports survives the plausible range of the most uncertain of them, which in a model with a shared cause is usually the shared-cause probability.*

16.6 Model credibility and application to other engineered systems

A complete calculation is not necessarily a credible PRA. This section states what makes the assessment defensible, lists the independent checks that catch implementation errors, and applies

the same structure to the risk of unserved energy in a power network and to the probability that a cooling system lets a temperature limit be exceeded.

16.6.1 What makes a PRA defensible

A defensible assessment connects each numerical input to appropriate evidence and documents the assumptions used to build the model, which is the traceability defined at the start of the chapter. Table 16.9 is a practical evidence table for the data center model. Each row is one model element; the second column is the evidence it requires, and the third the main concern about the credibility of that evidence. The concerns differ by element: for the initiating-event frequency, whether the event definitions and operating conditions of the records match the model; for the generator, whether the recorded tests exercised the whole function (start, transfer and load acceptance) or only the engine start; for the shared cause, whether the coupling mechanism has been correctly identified at all.

Table 16.9: Evidence table for the data center PRA: the evidence each model element requires and the main concern about its credibility.

Model element	Evidence required	Main credibility concern
Initiating-event frequency	Historical outage records, comparable site data	Are the event definitions and operating conditions consistent?
UPS failure probability	Test records, demand failures, equipment data	Are the observed failures representative of actual demands?
Generator failure probability	Start tests, operating records, manufacturer evidence	Does the function definition include transfer and load acceptance?
Shared-cause probability	Common-cause failure databases, failure investigations, engineering analysis	Is the coupling mechanism correctly identified?
Recovery probability	Restoration-time records, procedures, simulation	Are the timing and operating conditions realistic?
Consequence estimate	IT impact, outage simulations, contractual costs	Does the model capture severity and duration?

The present model has explicit limitations, which accompany any risk result reported from it:

- the initiating-event family is incomplete relative to the total risk of the facility;
- the basic-event probabilities are hypothetical;
- the shared-cause mechanism is represented by one aggregate event;
- time-varying load and battery energy are not modeled (the dynamic event tree of Section 15.5.5 would add them);
- the consequence values are illustrative;
- repair and restoration enter only through the recovery branch;
- the parameter and structural uncertainties have not been calibrated from evidence.

16.6.2 Independent model checks

Several checks identify mathematical and implementation errors. Each is applied here to the data center model.

1. All terminal-sequence probabilities sum to one: shown numerically and algebraically in Section 16.3.2.

2. The terminal-sequence frequencies sum to the initiating-event frequency: $0.400000\,\mathrm{yr}^{-1}$ in Table 16.2.
3. Shared basic events are not counted as independent duplicates: $\Pr(S_5 \mid I) = 0.010792$, not the product 0.00147808.
4. Units are consistent: frequencies in yr^{-1}, sequence probabilities dimensionless, consequences in reais per occurrence, risk in reais per year.
5. Consequence values are nonnegative and tied to defined outcomes: each C_k belongs to one end state of Table 16.3.
6. Alternative Boolean formulations give the same probabilities: the joint failure from the disjoint decomposition and from inclusion–exclusion both give 0.010792 (Section 16.2.5); the chain rule along the tree and the Boolean expressions both give the sequence probabilities.
7. Limiting cases behave sensibly: with $q_\mathrm{u} = q_\mathrm{g} = q_\mathrm{c} = 0$ every interruption sequence vanishes and $\mathcal{R} = 0$; with $q_\mathrm{c} = 1$ every initiating event ends in a major interruption and $\mathcal{R} = f_\mathrm{IE} \times \mathrm{R\$}\,500\,000$.

These checks do not prove that the model is adequate, that its structure, boundary and success criteria represent the real system; they reduce the risk of mathematical and implementation errors, which are common in large models and invisible in their final number.

16.6.3 Power systems: the risk of unserved energy

The integrated method extends directly to power networks. Consider a set of transmission-line initiating events $I_1, I_2, \ldots, I_m$, each with frequency $f_{\mathrm{IE},i}$. After an initiating outage the network may remain secure, require redispatch of generation, require a remedial action scheme, experience overloads, shed load, or suffer further cascading outages. Each sequence may require a power-flow or a dynamic-security calculation to determine its consequence. With the unserved energy of equation (15.13) as the consequence, equation (16.1) becomes the expected unserved energy,

$$\mathrm{EUE} = \sum_{i=1}^{m} f_{\mathrm{IE},i} \sum_{k \in \mathcal{S}_i} \Pr(S_{i,k} \mid I_i)\, \mathbb{E}[E^{\mathrm{unserved}} \mid I_i, S_{i,k}], \tag{16.17}$$

in megawatt-hours per year. The same framework connects contingency probabilities to physical consequences, and it can incorporate uncertain demand, generation, weather, line ratings, protection reliability and control actions.

The weather dependence of the transmission case shows why the conditional expectation in equation (16.17) matters. The line-outage tree of Section 15.4 used one family with $f_\mathrm{IE} = 0.60\,\mathrm{yr}^{-1}$, redispatch success 0.90 and 80 MWh shed when the RAS fails, and gave EUE = 3.588 MWh per year. Suppose, illustratively, that the outages divide into a normal family with $f_{\mathrm{IE},1} = 0.45\,\mathrm{yr}^{-1}$, redispatch success 0.95 and 80 MWh shed, and a hot, low-wind family with $f_{\mathrm{IE},2} = 0.15\,\mathrm{yr}^{-1}$, in which the weather-dependent line ratings are low, redispatch succeeds with probability 0.75 only, and the load shed is 60 MW for two hours, 120 MWh. Protection (0.99), the RAS (0.75) and the 400 MWh of the protection-failure sequence are as before. The frequency-weighted redispatch success, $(0.45 \times 0.95 + 0.15 \times 0.75)/0.60 = 0.90$, is unchanged, but

$$\begin{aligned}\mathrm{EUE} &= 0.45\,[0.99 \times 0.05 \times 0.25 \times 80 + 0.01 \times 400] \\ &\quad + 0.15\,[0.99 \times 0.25 \times 0.25 \times 120 + 0.01 \times 400] \\ &= 0.45 \times 4.99 + 0.15 \times 11.425 = 2.2455 + 1.7138 = 3.959\,\mathrm{MWh\ per\ year},\end{aligned}$$

10 % more than the single-family result. The hot family has a quarter of the outages and contributes 43 % of the expected unserved energy, because the weather that makes redispatch fail also makes the shedding larger: the expectation of a product of two quantities that depend on the same condition is not the product of their expectations. Even a single family with

the averaged redispatch probability and the averaged shed energy, 90 MWh, would give only 3.737 MWh per year.

A purely static event tree may be insufficient for cascading outages or for time-varying physical constraints, in which the next failure depends on the flows and temperatures that the previous ones produced; those require dynamic or sequential models, of which the dynamic event tree of Section 15.5.5 is the simplest. Power-system adequacy measures are developed in full in a later part of the book.

16.6.4 Cooling systems: the probability of exceeding a thermal limit

For the data center cooling plant the undesirable outcome is not the unavailability of equipment but a violation of the thermal requirement: the inlet temperature $T_{\text{in}}(t)$ of the critical IT equipment must stay at or below $T_{\max} = 27^{\circ}\text{C}$, the criterion of Chapter 2. Over a period $[0, \tau]$ after a disturbance, the failure event is the performance-based event

$$F = \Big\{ \max_{0 \le t \le \tau} T_{\text{in}}(t) > T_{\max} \Big\}. \tag{16.18}$$

A pump or chiller failure can initiate the disturbance, but whether F occurs depends on the remaining cooling capacity, the IT load, the thermal storage and inertia of the room and the water loop, the airflow and coolant distribution, backup operation and the recovery time. The PRA connects the component-failure sequences to a physical thermal model; in its simplest lumped form, an energy balance on one thermal node,

$$C_{\text{th}} \frac{\mathrm{d}T_{\text{in}}}{\mathrm{d}t} = \dot{Q}_{\text{IT}}(t) - \dot{Q}_{\text{cool}}(t), \tag{16.19}$$

where C_{th} is the effective thermal capacitance of the room air, the equipment mass and the water loop, in kilojoules per kelvin, $\dot{Q}_{\text{IT}}$ is the heat generated by the IT equipment and $\dot{Q}_{\text{cool}}$ the heat removed by the cooling plant, both in kilowatts. (The dot marks a heat rate; $\dot{Q}$ is unrelated to the top-event probability Q_{top}.) A component failure changes $\dot{Q}_{\text{cool}}$, but whether the threshold is crossed depends on the whole trajectory.

When the heat rates are constant during the deficit, $\dot{Q}_{\text{IT}} > \dot{Q}_{\text{cool}}$, equation (16.19) integrates to a straight line, $T_{\text{in}}(t) = T_0 + (\dot{Q}_{\text{IT}} - \dot{Q}_{\text{cool}})\, t / C_{\text{th}}$, from the initial temperature T_0. Setting $T_{\text{in}} = T_{\max}$ gives the time available before the limit is reached,

$$t_{\text{lim}} = \frac{C_{\text{th}}\,(T_{\max} - T_0)}{\dot{Q}_{\text{IT}} - \dot{Q}_{\text{cool}}}, \tag{16.20}$$

and the limit is exceeded if and only if cooling is not restored before t_{lim}. Take the cooling plant of the book's running case, three chillers of about 430 kW each of which two are required, and illustrative values $C_{\text{th}} = 30\,\text{MJ/K}$, which is 500 kW min/K, and $T_0 = 22^{\circ}\text{C}$. If two chillers are lost, the remaining 430 kW faces an IT load of 800 kW in the afternoon or 600 kW at night. The deficits are 370 kW and 170 kW, the temperature rises at $370/500 = 0.74\,\text{K/min}$ or $170/500 = 0.34\,\text{K/min}$, and $t_{\text{lim}} = 500 \times 5/370 = 6.76$ min or $500 \times 5/170 = 14.71$ min.

Figure 16.7 draws three trajectories of equation (16.19). The horizontal axis is the time in minutes after the loss of the two chillers, the vertical axis the inlet temperature, and the red horizontal line the limit $T_{\max} = 27^{\circ}\text{C}$. Each trajectory rises during the deficit and falls after a chiller is restored, when the two available chillers, 860 kW, again exceed the load. At 800 kW with restoration after five minutes (solid blue), the temperature peaks at 25.7°C and the requirement is met. At the same load with restoration after ten minutes (dashed red), it crosses the limit at 6.76 minutes and peaks at 29.4°C. At 600 kW with restoration after ten minutes (dotted green), it peaks at 25.4°C. The same component-failure sequence, and even the same restoration time, produce a violation or not depending on the load. The loss of one chiller alone would leave 860 kW of capacity and no temperature rise at either load.

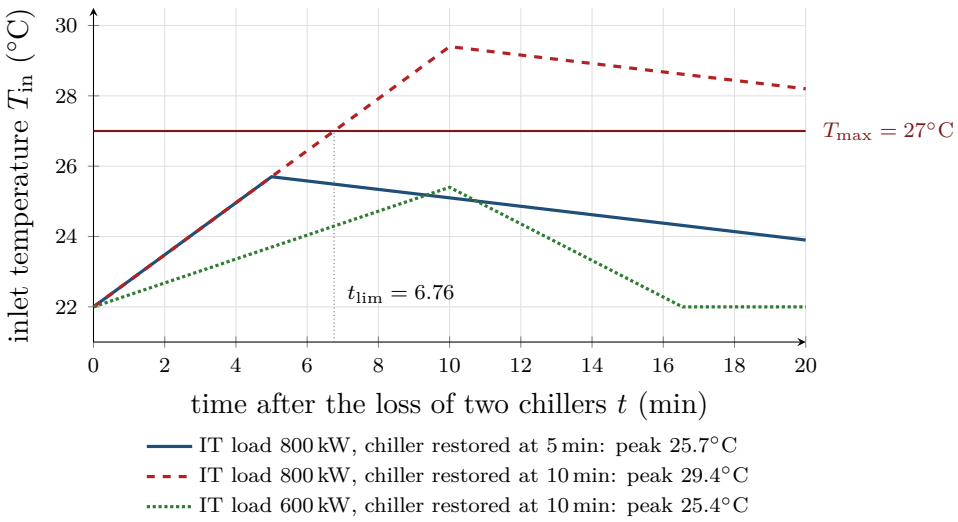


Figure 16.7: Inlet temperature after the loss of two of three chillers, $C_{\text{th}}\,\mathrm{d}T_{\text{in}}/\mathrm{d}t = \dot{Q}_{\text{IT}} - \dot{Q}_{\text{cool}}$ with $C_{\text{th}} = 500\,\text{kW min/K}$, $T_0 = 22$°C, $\dot{Q}_{\text{cool}} = 430\,\text{kW}$ during the deficit and 860 kW after restoration (illustrative). The limit is exceeded only when restoration comes after t_{lim}, which depends on the load.

The probability of the failure event follows by combining the thermal model with the probabilistic model of restoration and load. If the restoration time T_{R} is exponential with rate $\mu = 0.10\,\text{min}^{-1}$, as in Section 15.5.4, and independent of the load, which is 800 kW with probability 0.3 and 600 kW with probability 0.7, then

$$\begin{aligned}\Pr(F \mid \text{two chillers lost}) &= \sum_{\ell} \Pr(\ell)\,\Pr(T_{\text{R}} > t_{\text{lim}}(\ell)) = 0.3\,\mathrm{e}^{-0.676} + 0.7\,\mathrm{e}^{-1.471} \\ &= 0.3 \times 0.5088 + 0.7 \times 0.2298 = 0.3135,\end{aligned} \tag{16.21}$$

the sum running over the load levels ℓ. A model that counted the loss of two chillers as a system failure, the component-state criterion of a 2-out-of-3 block diagram, would assign this event the probability 1; the thermal criterion assigns it 0.31 under these assumptions, and identifies the afternoon load and slow restoration as the conditions that produce the violation. The value of integrating the probabilistic model with the physical state evolution is that the final risk measure corresponds to the actual engineering failure criterion rather than to a proxy for it.

The next chapter turns from nonrepairable mission models to repairable systems, whose failures, repairs and returns to service recur and are described by counting processes (Chapter 17).

16.7 What to retain

Definition 16.1 (Integrated probabilistic risk assessment)**.** An integrated PRA computes the risk triplet of a defined system (its scenarios, their frequencies and their consequences) from one model that links initiating events, the fault trees of the protection functions and the event tree of their responses to physical and financial consequences; it preserves shared causes in the joint logic of every sequence, evaluates improvements as interventions on the whole model, and propagates the uncertainty of every input to the decision quantities.

Six results carry forward. A PRA integrates initiating events, system functions, dependencies and consequences, and does not stop at a component failure probability. Dependence must be preserved throughout: a shared basic event made the joint failure of the UPS and the generator 7.3 times more probable than the product of the marginals and raised the expected loss 2.7-fold

relative to an independence model. Expected risk depends on severity as well as frequency, and a rare severe sequence carried 80 % of it. Improvements are evaluated on the full model, and their ranking depends on the consequence measure and on the cost. Uncertainty is propagated to the final quantities, with aleatory and epistemic randomness kept apart. And modeling the physical system makes the PRA measure the actual failure criterion, a meaningful violation of the required function rather than a component outage. Table 16.10 lists the chapter's terms, their meaning and units, and the subsection that defines each.

Table 16.10: The terms of this chapter and where each is defined.

Term	Meaning	Expression	Units	Defined in
Risk chain	initiating events to decisions	six stages, Figure 16.1	–	16.1.1
Expected consequence rate	risk over all families	equation (16.1)	R$/yr, MWh/yr	16.1.1
Linked fault tree	heading failure as a top event	$\bar{U} = X_{\mathrm{c}} \cup X_{\mathrm{u}}$	–	16.2.4
Shared basic event	one cause in two fault trees	X_{c}, probability q_{c}	–	16.2.3
Joint function failure	both functions fail	$q_{\mathrm{c}} + (1 - q_{\mathrm{c}}) q_{\mathrm{u}} q_{\mathrm{g}}$	–	16.2.5
Sequence probability	from the basic events	equation (16.6)	–	16.3.2
Sequence frequency	expected occurrences per year	$f_{S_k} = f_{\mathrm{IE}} \Pr(S_k \mid I)$	yr^{-1}	16.3.3
Interruption frequency	interruption sequences per year	$f_{\mathrm{IE}} \sum_{k=3}^{5} \Pr(S_k \mid I)$	yr^{-1}	16.3.5
Expected annual interruption	expected duration per year	$\sum_k f_{S_k} D_k$	h/yr	16.3.5
Expected annual loss	risk in money	$\sum_k f_{S_k} C_k$	R$/yr	16.3.6
Risk function	loss as a function of parameters	equation (16.9)	R$/yr	16.4.1
Risk sensitivity	derivative of the risk	$\partial \mathcal{R} / \partial q_i$	R$/yr	16.4.2
Risk importance	FV, RAW, RRW on the risk	equation (16.11)	–	16.4.2
Present value, NPV	discounted risk reduction, less cost	$\Delta \mathcal{R} \, [1 - (1 + \rho)^{-n}] / \rho$	R$	16.4.4
First-order propagation	linearized variance	$\nabla \mathcal{R}^{\mathsf{T}} \Sigma_{\boldsymbol{\theta}} \nabla \mathcal{R}$	$(\mathrm{R\$/yr})^2$	16.5.2
Monte Carlo propagation	sampled distribution of $\mathcal{R}$	$\mathcal{R}^{(j)} = \mathcal{R}(\boldsymbol{\theta}^{(j)})$	R$/yr	16.5.3
Aleatory, epistemic	random outcomes; incomplete knowledge	C_{yr} against $\mathcal{R}$	–	16.5.4
Thermal-limit failure	limit exceeded in the period	$\max_{t \le \tau} T_{\mathrm{in}}(t) > T_{\max}$	–	16.6.4

Further reading. Kumamoto and Henley, *Probabilistic Risk Assessment and Management for Engineers and Scientists*, second edition, for integrated PRA methods, fault trees, event

trees and consequence analysis; Rausand, Barros and Høyland, *System Reliability Theory*, third edition, for reliability modeling and probabilistic risk assessment; NASA, *Probabilistic Risk Assessment Procedures Guide*, for integrated scenario modeling, uncertainty and risk-informed decisions; the NRC's PRA guidance and parameter-estimation handbooks, for initiating-event data, system models and quantification practices.

16.8 Exercises

Exercise 16.1. (a) State why the marginal UPS and generator failure probabilities of the data center model cannot simply be multiplied. (b) The shared-cause probability is doubled to $q_c = 0.02$, with $q_u = 0.02$ and $q_g = 0.04$. Compute $\Pr(\bar{U})$, $\Pr(\bar{G})$, the correct joint failure probability, the product of the marginals and their ratio. (c) Compute $\Pr(\bar{G} \mid \bar{U})$ and $\Pr(G \mid \bar{U})$, the two generator branches after a UPS failure. (Answers: (a) they share a common-cause failure event, which makes the two function failures dependent. (b) $1 - 0.98 \times 0.98 = 0.0396$, $1 - 0.98 \times 0.96 = 0.0592$, $0.02 + 0.98 \times 0.0008 = 0.020784$, against a product of 0.00234432, 8.87 times smaller. (c) $0.020784/0.0396 = 0.524848$ and $0.98 \times 0.02 \times 0.96/0.0396 = 0.475152$.)

Exercise 16.2. (a) State how a conditional sequence probability is converted into an annual sequence frequency, and which outcome dominates the expected annual loss of the baseline model. (b) A second site has the same equipment but a weaker grid connection, $f_{IE} = 0.60\,\mathrm{yr}^{-1}$, and its shared-cause probability is $q_c = 0.02$; the other probabilities and the consequences are those of Table 16.3. Compute the five sequence probabilities, check that they sum to one, and compute the five frequencies. (c) Compute the interruption frequency, the expected interruption minutes per year, the expected annual loss and the share of sequence 5 in it. (Answers: (a) multiply by the initiating-event frequency; sequence 5, the major interruption, which is rare but has a much larger consequence per occurrence. (b) 0.921984, 0.0268912, 0.0115248, 0.018816 and 0.020784, sum 1; frequencies 0.553190, 0.016135, 0.006915, 0.011290 and $0.012470\,\mathrm{yr}^{-1}$, sum $0.60\,\mathrm{yr}^{-1}$. (c) $f_{int} = 0.60 \times 0.0511248 = 0.030675\,\mathrm{yr}^{-1}$; $0.60 \times (0.0230496 + 0.0018816 + 0.124704) = 0.089781\,\mathrm{h}$, or 5.39 minutes per year; $\mathcal{R} = 0.60 \times (1152.48 + 188.16 + 10\,392) =$ R$ 7039.58 per year, of which sequence 5 contributes R$ 6235.20, 88.6 %.)

Exercise 16.3. (a) State why a ranking by risk sensitivity is not automatically a ranking of investments. (b) Derive $\partial\mathcal{R}/\partial q_{rec}$ from equation (16.9) and evaluate it at the baseline. (c) Option B costs R$ 2000 and option D R$ 1500, both with a ten-year benefit, a 5 % discount rate and no other costs. Compute their net present values. (d) Options A and B are applied together. Compute the expected annual loss and compare its reduction with the sum of the separate reductions of Table 16.6. (Answers: (a) real projects have different costs, feasible improvements and side effects, such as new dependencies, so specific interventions must be evaluated on the full model. (b) Only the sequence 3 term contains q_{rec}: $f_{IE} \times 100\,000\,(1 - q_c)(1 - q_u)q_g = 0.40 \times 100\,000 \times 0.9702 \times 0.04 = 1552.32$ reais per year per unit probability. (c) $310.46 \times 7.7217 - 2000 = +R 397 and $232.85 \times 7.7217 - 1500 = +R 298. (d) $0.40 \times (0.995 \times 0.98 \times 0.02 \times 0.30 \times 100\,000 + 0.995 \times 0.02 \times 0.98 \times 10\,000 + (0.005 + 0.995 \times 0.0004) \times 500\,000) =$ R$ 1391.63, a reduction of R$ 1308.50 against R$ 1306.93 for the sum; the difference, R$ 1.57, is the cross term in $q_c q_g$.)

Exercise 16.4. (a) State the difference between epistemic and aleatory uncertainty, and which one the histogram of Figure 16.6(a) describes. (b) Suppose only f_{IE} (standard deviation $0.20\,\mathrm{yr}^{-1}$) and q_c (standard deviation 0.0099) are uncertain, independent, with the sensitivities of Table 16.7. Compute the first-order standard deviation of $\mathcal{R}$. (c) With the baseline parameters, compute the probability that a year passes without any IT interruption from this initiating-event family, and the probability of at least one interruption in ten years. (d) Explain why the realized annual loss is almost never close to R$ 2700. (Answers: (a) epistemic uncertainty

concerns incomplete knowledge of the parameters and the model, aleatory variability the random future outcomes under a given model; the histogram is epistemic, a distribution over the expected loss. (b) $\sqrt{1.823 \times 10^6 + 3.893 \times 10^6} = \text{R\$}\,2391$ per year. (c) $\mathrm{e}^{-0.016577} = 0.98356$ and $1 - \mathrm{e}^{-0.16577} = 0.15276$. (d) the realized loss is zero in 98.4 % of years and otherwise at least R\$ 10 000; the expected value averages these outcomes and is not itself a possible outcome.)

Part V

Repairable Systems and Availability

Chapter 17

Repairable Systems and Counting Processes

17.1 Repairable systems and recurrent failures

Chapters 5 to 8 studied the time to the first failure of an item: the lifetime T, its reliability function $R(t) = \Pr(T > t)$, its hazard function, and the estimation of their parameters from censored records. Chapters 9 to 16 combined such items into systems and accident scenarios, again up to the first loss of a required function. That framework is appropriate for nonrepairable items and for any item that is evaluated only until its first failure. Much industrial equipment is not used that way. A cooling-water pump fails after 2000 h, is repaired, operates for another 1500 h, fails again and returns to service. Transmission lines trip and are restored; data center chillers fail and are repaired; emergency generators experience repeated faults; inverters and battery subsystems pass through many repair cycles. For such equipment a single random lifetime does not describe the failure history, which is a sequence of recurrent events. The question of this chapter is

how equipment that fails, is repaired and fails again is modeled, and how it is decided from its failure history whether its reliability is improving or deteriorating over time.

The chapter answers it in five steps. It defines the failure counting process of a repairable system and the times between its failures, and it introduces recurrent failure data with the mean number of failures per system and the rate of occurrence of failures, which it separates from the hazard function of a lifetime. It derives the homogeneous Poisson process, the model with a constant rate of occurrence of failures, and the Poisson distribution of its failure counts. It generalizes to the nonhomogeneous Poisson process, whose rate changes with the age of the system, with the power-law process as the standard example of improving and deteriorating systems. It states what repair does to a system, as good as new, as bad as old or in between, and separates that assumption from the choice of a failure-process model and from availability. Finally, it estimates the models from failure timestamps, tests whether a constant rate is adequate, and applies the ideas to transmission-line outages, data center cooling and battery energy storage.

Jargon: repairable system, recurrent event, failure epoch, counting process, interfailure time, cumulative intensity, MCF, ROCOF, conditional intensity, HPP, NHPP, power-law process, perfect, minimal and imperfect repair, virtual age

A *repairable system* is an item that is restored to operation after a failure without being replaced as a whole, so that it can fail again; its failures are *recurrent events*. The *failure epochs* $T_1 < T_2 < \cdots$ are the times of its successive failures, measured from a common origin. The *counting process* $N(t)$ is the number of failures in $(0, t]$, and the *interfailure times* are $X_j = T_j - T_{j-1}$. The *cumulative intensity*, also called the *mean function*, is $W(t) = \mathbb{E}[N(t)]$, the expected number of failures by time t; in recurrent-data analysis it is called the *mean cumulative function* (MCF), the name also used for its nonparametric estimate. The *rate of occurrence of failures* (ROCOF) is $w(t) = \mathrm{d}W/\mathrm{d}t$, in failures per hour: the unconditional rate at which the system accumulates failures. The *conditional intensity* $\lambda^*(t \mid \mathcal{H}_t)$ is the rate of failure at t given the observed failure history $\mathcal{H}_t$. The *homogeneous Poisson process* (HPP) has a constant ROCOF; the *nonhomogeneous Poisson process* (NHPP) has a ROCOF that varies with time; the *power-law process* is the NHPP with $W(t) = (t/\eta)^b$. *Perfect repair* restores the system as good as new; *minimal repair* restores its function and leaves it as bad as old, at the condition it had just before the failure; *imperfect repair* leaves it in between. The *virtual age* V_n is the effective age of the system just after its n-th repair.

Key idea

The hazard function belongs to one lifetime and stops at the first failure; the rate of occurrence of failures $w(t)$ belongs to the failure process of a repairable system and runs through all its failures. Whether a system is improving or deteriorating is a statement about the slope of its expected cumulative number of failures, $W(t)$, and the model that fits it, the HPP or an NHPP, is a separate choice from the assumption about what repair does.

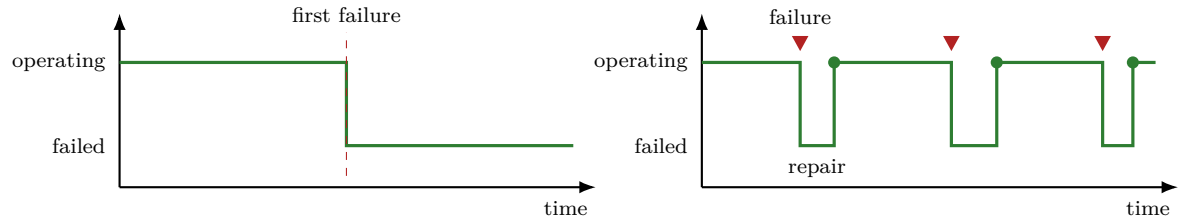
This section distinguishes repairable from nonrepairable items, defines the counting process that records the failures of a repairable system, and separates that process from the operating state of the system.

17.1.1 Nonrepairable and repairable items

For a nonrepairable item the life ends at the first failure. Its behavior is a single transition from the operating state to the failed state, and the natural questions concern that transition, such as $\Pr(T > 1000\,\mathrm{h})$, the probability of surviving the first 1000 h. A repairable item makes the same transition many times, each followed by a period of repair and a return to service. Figure 17.1 draws the two cases. In each panel the horizontal axis is time and the vertical axis is the state of the item, operating or failed. Panel (a) is a nonrepairable item: one transition, at the first failure, after which the item stays failed. Panel (b) is a repairable item: red triangles mark failures, each followed by a short down interval of repair, and green dots mark the restorations to service. The schedule is schematic and not to scale.

For the repairable item of panel (b), the probability of surviving the first 1000 h is still a meaningful quantity, but it no longer answers the questions that an operator or a maintenance planner asks. These are

- how many failures are expected in the next 1000 h;
- whether the rate at which failures occur is increasing;
- how long a repair takes;
- what fraction of the time the system is operating;



(a) nonrepairable item (b) repairable item

Figure 17.1: State of a nonrepairable and of a repairable item over time, schematic. (a) One transition from operating to failed; the life ends at the first failure. (b) Repeated failures (red triangles), repairs and restorations (green dots); the item passes through the operating and failed states many times.

- whether maintenance restores the system to an as-good-as-new condition;
- whether the system deteriorates as failures accumulate.

The first two and the last concern the sequence of failure times; the third and fourth concern the durations of the down periods; the fifth concerns what a repair does to the condition of the system. Each needs its own stochastic model, and the rest of the chapter builds the models for the failure sequence and for the effect of repair. Repair durations and the fraction of time operating, availability, are introduced in Section 17.5.6 and developed in the following chapters.

17.1.2 The failure counting process

Let $T_1, T_2, T_3, \ldots$ denote the times of the successive failures of a repairable system, measured from a common origin, normally the moment the system is first put into service. They are the *failure epochs*, random variables with $0 < T_1 < T_2 < \cdots$. The *failure counting process* is

$$N(t) = \sum_{j=1}^{\infty} \mathbf{1}\{T_j \le t\}, \tag{17.1}$$

where $\mathbf{1}\{\cdot\}$ is the indicator function, equal to one when the condition in braces holds and to zero otherwise. $N(t)$ is the number of failures observed in $(0, t]$. It is a random function of time with four properties: $N(0) = 0$; its values are nonnegative integers; it is nondecreasing; and it increases by exactly one at each failure epoch, because two failures are assumed never to occur at the same instant (a process with this property is called *orderly*). For a nonrepairable item the counting process makes a single jump, at $T_1 = T$, and stays at one.

As an example, suppose a cooling pump fails at $T_1 = 400\,\text{h}$, $T_2 = 1100\,\text{h}$ and $T_3 = 1800\,\text{h}$. Its counting process is

$$N(t) = \begin{cases} 0, & 0 \le t < 400, \\ 1, & 400 \le t < 1100, \\ 2, & 1100 \le t < 1800, \\ 3, & t \ge 1800, \end{cases}$$

with t in hours. Figure 17.2 draws this sample path. The horizontal axis of both panels is calendar time in hours. The lower panel is $N(t)$: a step function that is flat between failures and rises by one at each failure epoch, where the filled dot marks the value taken at the epoch itself and the open dot the value just before it. The braces above the steps are the interfailure times defined in Section 17.2.1. The upper panel is the operating status of the same pump, discussed next; in this illustration each repair takes 50 h.

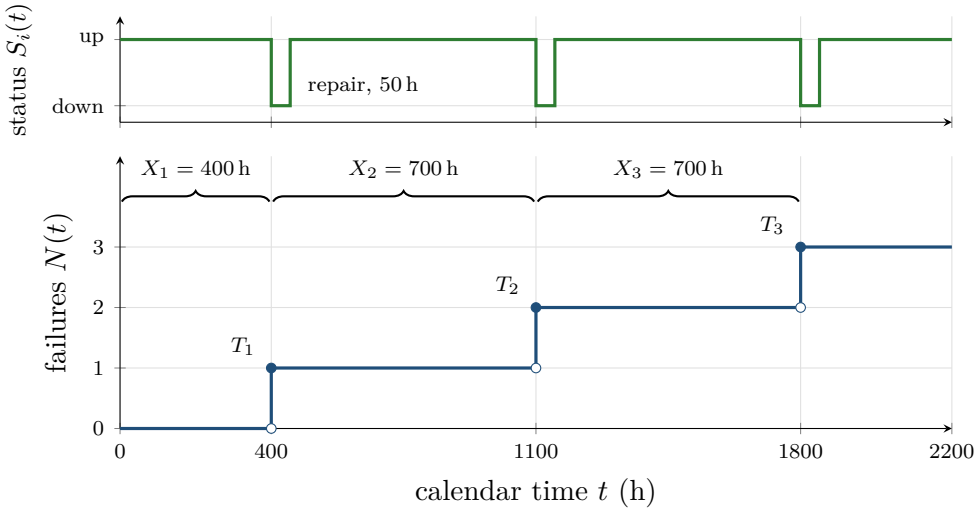


Figure 17.2: A sample path of a repairable pump, illustrative. Lower panel: the failure counting process $N(t)$ of equation (17.1), with failure epochs $T_1 = 400$, $T_2 = 1100$ and $T_3 = 1800$ h and the interfailure times X_j between them; filled dots are the values at the epochs, open dots the values just before. Upper panel: the operating status of the same pump, down for 50 h of repair after each failure. $N(t)$ records the failures but not the repairs.

17.1.3 Failure occurrence and operating state

The counting process records failures, not the duration of the repairs that follow them. In the lower panel of Figure 17.2, $N(t)$ is the same whether each repair takes 50 h or 500 h. Two processes must therefore be distinguished. The *failure occurrence process* is $N(t)$, the sequence of failure epochs. The *operating state process* is the status $S_i(t)$ of the system, unit i, equal to one while it performs its required function and to zero while it is down, drawn in the upper panel. $N(t)$ counts the transitions of S_i from one to zero; it says nothing about how long S_i stays at zero.

The two processes answer different questions and can disagree in either direction. A system may have accumulated five failures and be operating now; another may have had only one failure and still be unavailable because its repair has not been completed. The failure occurrence process answers the questions about the number and the trend of failures; the operating state process, which needs a model of repair durations as well, answers the questions about availability. This chapter models the first; Section 17.5.6 shows where the second begins.

17.2 Recurrent failure data and the rate of occurrence of failures

The data of a repairable system are its failure times, read on a clock that must be stated. This section defines the interfailure times and the two clocks on which they can be measured, presents a recurrent failure dataset for a fleet of pumps, defines the expected cumulative number of failures and its derivative, the rate of occurrence of failures, distinguishes that rate from the hazard function, and estimates the expected cumulative number of failures from the fleet's records.

17.2.1 Interfailure times

The successive *interfailure times*, or times between failures, are

$$X_1 = T_1, \qquad X_j = T_j - T_{j-1}, \quad j \geq 2, \tag{17.2}$$

so that each failure epoch is the sum of the interfailure times before it, $T_k = X_1 + X_2 + \cdots + X_k$. The first interfailure time is the time to first failure, $X_1 = T_1$, which is the lifetime T of the earlier chapters when the system starts new. For the pump of Figure 17.2, with $T_1 = 400$, $T_2 = 1100$ and $T_3 = 1800$ h, the interfailure times are

$$X_1 = 400\,\text{h}, \qquad X_2 = 1100 - 400 = 700\,\text{h}, \qquad X_3 = 1800 - 1100 = 700\,\text{h}.$$

A repairable system can be described either by its counting process $N(t)$ or by the sequence of its interfailure times; the two carry the same information, since $N(t) \geq k$ exactly when $T_k \leq t$. The models of this chapter differ in what they assume about the interfailure times: whether they are independent, whether they have the same distribution, and whether that distribution depends on the age of the system.

17.2.2 Calendar and operating clocks

The interfailure times are differences of failure epochs, and their meaning depends on the clock on which the epochs are read. Two clocks are common. The *calendar-time clock* runs continuously, during operation and during repair. The *operating-time clock* advances only while the equipment is operating and therefore exposed to the failure mechanisms of interest; it is the clock of an hour meter, and its readings are the *operating ages* of the equipment. Between two failures, the calendar interval includes the repair time that follows the first failure; the operating interval does not.

Suppose a pump fails at calendar hour 400, takes 50 h to repair, and fails again at calendar hour 1100, as in Figure 17.2. The calendar time between the two failure events is $1100 - 400 = 700$ h. If the pump operated continuously outside the repair period, its operating exposure between the failures is $700 - 50 = 650$ h. Using the wrong clock biases an estimated rate. One failure in 650 operating hours is a rate of $1/650 = 1.538 \times 10^{-3}$ failures per operating hour; the same failure counted against 700 calendar hours gives 1.429×10^{-3} per calendar hour, 7.1 % lower, because the repair period, during which the pump could not fail, has been counted as exposure. The distinction matters because failure occurrence models are usually specified in operating time, the time during which the failure mechanisms act, while availability models use calendar time, in which repair periods are part of the history. Every recurrent failure analysis must state its clock.

17.2.3 A recurrent failure dataset

As a running example, five similar cooling pumps are each monitored for 3000 operating hours, and every failure is recorded with the pump's cumulative operating age at that moment; downtime is excluded, so the operating-time clock is used throughout. After each failure the pump is repaired and returned to service. Table 17.1 lists the failure ages and the number of failures of each pump. The data are illustrative.

Figure 17.3 draws the records as observed histories. The horizontal axis is operating age; each row is one pump, observed from age zero to 3000 h. A red cross marks each failure, after which the pump was repaired and the history continues on the same line; the green bar at the right marks the end of observation, with the pump operating. The number at the right of each row is its count of failures. Unlike the lifetime records of Chapter 7, where each history ended at the first failure, every history here continues through its failures to the end of the observation window.

The simplest summary of the fleet is its overall recurrence rate. The total operating exposure is the sum of the observation times, and with the notation of Chapter 7, r for the number of observed failures and T_{tot} for the total exposure,

$$T_{\text{tot}} = 5 \times 3000\,\text{h} = 15\,000\,\text{h}, \qquad \frac{r}{T_{\text{tot}}} = \frac{16}{15\,000\,\text{h}} = 1.0667 \times 10^{-3}\,\text{h}^{-1},$$

Table 17.1: Recurrent failures of five cooling pumps, each observed to an operating age of 3000 h, illustrative. Ages are cumulative operating hours with downtime excluded.

Pump	Failure ages (h)	Number of failures
P1	400, 1100, 1800, 2600	4
P2	800, 1700, 2400	3
P3	600, 1500, 2800	3
P4	900, 2000	2
P5	300, 1000, 1900, 2700	4
Total		16

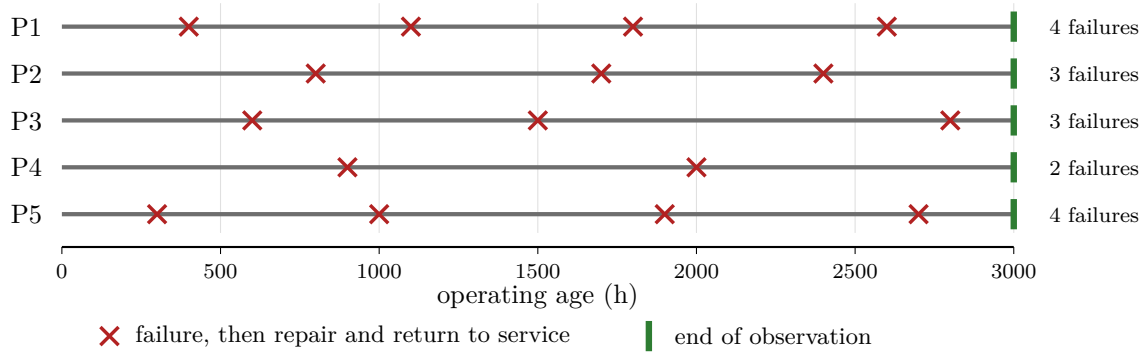


Figure 17.3: Recurrent failure histories of the five pumps of Table 17.1, illustrative, on the operating-time clock. Each pump is observed from age zero to 3000 h and continues through its failures; 16 failures in 15 000 pump-hours.

that is, 15 000 pump-hours and 1.0667×10^{-3} failures per pump-hour, about 1.067 failures per 1000 pump-operating-hours. This number is an average over the whole window. It cannot say whether the rate is constant as the pumps age, rising or falling, because it would be the same for any arrangement of the same 16 failures within the 3000 h. Answering that question requires a model of how the rate of failures depends on age, which the rest of this section defines.

17.2.4 Cumulative intensity and the rate of occurrence of failures

The expected number of failures of a system by time t is the *cumulative intensity*, or *mean function*,

$$W(t) = \mathbb{E}[N(t)]. \tag{17.3}$$

It is dimensionless, a count of failures, and as an expectation it need not be an integer. Its derivative, where it exists, is the *rate of occurrence of failures* (ROCOF),

$$w(t) = \frac{\mathrm{d}W(t)}{\mathrm{d}t}, \qquad W(t) = \int_0^t w(u)\,\mathrm{d}u, \tag{17.4}$$

where the integral form uses $W(0) = 0$. The ROCOF has units of failures per hour, h^{-1} when time is in hours. Its meaning follows from the definition of the derivative. Over a short interval,

$$\mathbb{E}[N(t + \Delta t) - N(t)] = W(t + \Delta t) - W(t) \approx w(t)\,\Delta t,$$

and for an orderly process, in which at most one failure occurs in a sufficiently short interval, the expected number of failures in $(t, t + \Delta t]$ is the probability of a failure there, up to terms of higher order in Δt. The ROCOF is therefore the unconditional probability of a failure in a

short interval at time t, per unit length of the interval: unconditional in the sense that it does not assume anything about whether, or how often, the system has failed before t.

As an illustration, suppose that a pump population has $W(2000) = 2.5$ and $W(3000) = 4.0$, with ages in operating hours. Then

$$W(3000) - W(2000) = 4.0 - 2.5 = 1.5$$

is the expected number of additional failures per pump between the ages of 2000 and 3000 h. A value $W(2000) = 2.5$ does not mean that each pump has failed exactly 2.5 times, which is impossible; it is an expectation across comparable system histories, some of which have two failures by that age, some three, and some other numbers. The renewal function $M(t)$ of Chapter 18 is the same quantity $W(t)$ for the particular process that arises when every repair is a replacement by a new unit.

17.2.5 ROCOF against the hazard function

The ROCOF and the hazard function are both rates in h^{-1}, both are often called a failure rate, and both can be written as a limit of a probability in a short interval divided by its length. They are different quantities, and confusing them is the most common error in the analysis of repairable systems. The hazard function of Chapter 5 is

$$\lambda(t) = \lim_{\Delta t \to 0^+} \frac{\Pr(t < T \le t + \Delta t \mid T > t)}{\Delta t},$$

the instantaneous rate of the first failure of an item conditional on its survival to age t. It is a property of the distribution of one lifetime, and it is defined only up to that lifetime's single failure. A repairable system needs a rate that continues through all its failures. Let $\mathcal{H}_t$ denote the observed failure history up to time t, the number and times of the failures before t. The *conditional intensity* of the counting process is

$$\lambda^*(t \mid \mathcal{H}_t) = \lim_{\Delta t \to 0^+} \frac{\Pr[N(t + \Delta t) - N(t) = 1 \mid \mathcal{H}_t]}{\Delta t}, \tag{17.5}$$

under the assumption that the probability of two or more failures in a short interval is negligible. The asterisk marks that the intensity may depend on the history: on the time since the last failure, on the number of failures so far, or on any other feature of $\mathcal{H}_t$. The ROCOF is its average over the possible histories,

$$w(t) = \mathbb{E}[\lambda^*(t \mid \mathcal{H}_t)],$$

under the usual regularity conditions, because the unconditional probability of a failure in $(t, t + \Delta t]$ is the expectation of the conditional one.

Table 17.2 sets the three quantities side by side. Each row gives the quantity, its notation, what it is conditioned on, and its interpretation. The first is a property of a lifetime; the second and third are properties of the failure process. They coincide in some special models and differ in general.

Two examples show when the three coincide and when they do not. In a process whose conditional intensity does not depend on the history, the intensity is a deterministic function of time and equals the ROCOF; the Poisson processes of Sections 17.3 and 17.4 are of this kind. For such a process the probability of no failure in $(0, t]$ is $\mathrm{e}^{-W(t)}$ (Section 17.4.2), so the time to the first failure has reliability function $\mathrm{e}^{-W(t)}$ and hazard $w(t)$: the hazard of the first lifetime and the ROCOF are the same function. They still mean different things, since the hazard stops at the first failure while the ROCOF describes the second, third and every later failure. In a process in which every repair replaces the failed unit by a new one, on the other hand, the

Table 17.2: Three rates that must not be treated as interchangeable. Units h^{-1} in every row.

Quantity	Notation	Conditioned on	Interpretation
Hazard function of a lifetime	$\lambda(t)$	no failure before t, $T > t$	instantaneous rate of the first failure of an item of age t
Conditional intensity	$\lambda^*(t \mid \mathcal{H}_t)$	the observed failure history	instantaneous rate of the next failure given the history
Rate of occurrence of failures (ROCOF)	$w(t) = \mathrm{d}W/\mathrm{d}t$	nothing; averaged over histories	slope of the expected cumulative number of failures

conditional intensity after the n-th failure is the hazard of a new unit evaluated at the time since that failure. If the units wear out, with an increasing hazard, the ROCOF of the process nevertheless settles to a constant equal to one over the mean time to failure, as Chapter 18 shows: each unit ages, but the system, continually renewed, does not.

Common mistake

Calling the ROCOF a hazard, or reading one for the other. The hazard $\lambda(t)$ is a property of one lifetime, conditional on survival to t; it says nothing about failures after the first. The ROCOF $w(t)$ is a property of the failure process of a repairable system, unconditional, and runs through all its failures. A system whose units each have an increasing hazard can have a constant ROCOF, when failed units are replaced, and an increasing ROCOF does not show that any unit has an increasing hazard. "The failure rate is increasing" must say which of the two is meant.

Principle 17.1. *The hazard function describes when an item fails for the first time; the rate of occurrence of failures describes how fast a repairable system accumulates failures over its life. A repairable system is deteriorating when its ROCOF increases and improving when its ROCOF decreases; the shape of a component's hazard does not by itself settle either.*

17.2.6 The empirical mean cumulative function

The cumulative intensity $W(t)$ can be estimated from recurrent failure data without assuming a model. When K comparable systems are all observed to the same age, as the five pumps are, the natural estimate is the average count per system,

$$\hat{W}(t) = \frac{1}{K} \sum_{i=1}^{K} N_i(t), \tag{17.6}$$

where $N_i(t)$ is the counting process of system i and the hat marks an estimate. This is the *empirical mean cumulative function*. Table 17.3 evaluates it for the pumps at four ages. The columns are the operating age, the failures of the whole fleet in the preceding 1000 h, the cumulative failures of the fleet, and the estimate. At 1000 h, for example, the fleet has six failures, at 400, 800, 600, 900, 300 and 1000 h, so $\hat{W}(1000) = 6/5 = 1.20$; the failure of P5 at exactly 1000 h is counted, since $N(t)$ counts failures in $(0, t]$.

Figure 17.4 draws the estimate at every age. The horizontal axis is operating age and the vertical axis the expected number of failures per pump. The black step function is $\hat{W}(t)$, which rises by $1/5 = 0.2$ at each of the 16 failure ages of the fleet; the three dots are the values of

Table 17.3: Empirical mean cumulative function of the five pumps, equation (17.6), at four operating ages.

Operating age (h)	Fleet failures in the preceding 1000 h	Cumulative fleet failures	$\hat{W}(t)$ per pump
0	–	0	0
1000	6	6	1.20
2000	6	12	2.40
3000	4	16	3.20

Table 17.3. The dashed blue line and the dash-dotted gold curve are the homogeneous and the power-law Poisson models fitted to the same data in Sections 17.3.4 and 17.6.2. The estimate increases with age, as it must, since it counts failures. The relevant question is whether its slope, the estimated ROCOF, increases or decreases. The fleet's counts in the three successive 1000 h intervals are 6, 6 and 4, which suggests a modest decline of the recurrence rate in the last interval. Five pumps and sixteen failures, however, provide limited statistical evidence, and random variation could produce such a pattern from a constant rate; the fleet should not be declared to be improving without further analysis, which Section 17.6.4 carries out.

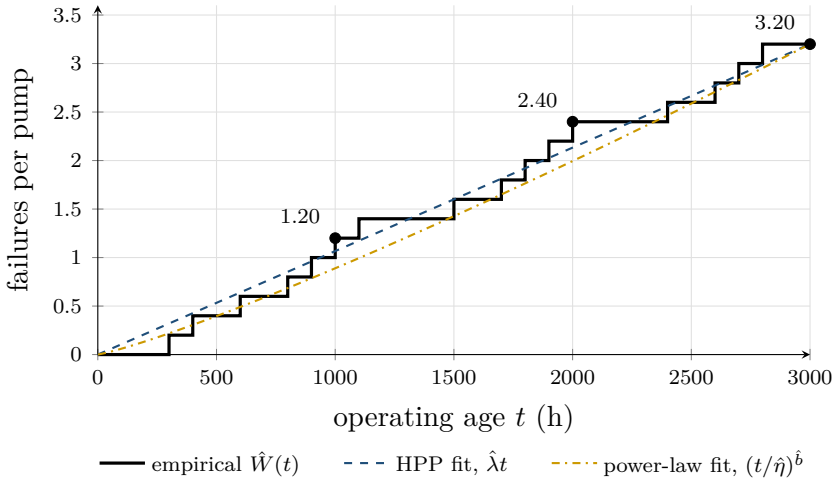


Figure 17.4: Empirical mean cumulative function of the five pumps, equation (17.6), with the HPP fit $\hat{\lambda}t$, $\hat{\lambda} = 1.0667 \times 10^{-3}\,\mathrm{h}^{-1}$ (Section 17.3.4), and the power-law fit $(t/\hat{\eta})^{\hat{b}}$, $\hat{b} = 1.165$, $\hat{\eta} = 1106\,\mathrm{h}$ (Section 17.6.2). Both fits pass through 3.2 at the end of the window, as their estimating equations require.

The estimate of equation (17.6) is simple because all pumps were observed to the same age. When some systems leave observation earlier, because they were installed later, retired or because the data were extracted on a fixed date, the changing number of systems under observation must be taken into account. For independent, comparable systems with noninformative end of observation, in the sense of Section 7.3.5, the nonparametric estimate is

$$\hat{W}(t) = \sum_{j:\, t_j \le t} \frac{d_j}{n_j}, \tag{17.7}$$

where the t_j are the distinct observed failure ages, d_j is the number of failures at age t_j across all systems, and n_j is the number of systems still under observation just before t_j. Each term is the average number of failures per observed system at t_j, and the sum accumulates them. When all K systems are observed throughout, $n_j = K$ and the sum reduces to equation (17.6).

The estimator resembles the Kaplan–Meier estimator of Chapter 7, with one essential difference. In a Kaplan–Meier calculation a unit leaves the risk set at its first failure, because its lifetime has ended. A repaired system does not leave the set of systems under observation merely because it has failed; it leaves only when its observation ends. The estimator also requires a consistent clock across systems and an explicit treatment of periods in which a system is under observation but not exposed, such as long outages or standby periods on the operating-time clock.

17.3 The homogeneous Poisson process

The simplest model of recurrent failures has a constant rate of occurrence of failures. This section defines it, derives the exponential distribution of its interfailure times and the Poisson distribution of its failure counts, fits it to the pump fleet, and uses the fit to predict failures over the next operating interval.

17.3.1 Definition

A *homogeneous Poisson process* (HPP) with rate $\lambda > 0$ is a counting process with three properties:

1. $N(0) = 0$;
2. *independent and stationary increments*: the numbers of failures in disjoint intervals are independent, and the distribution of the number in an interval depends only on the interval's length, not on its position;
3. in a short interval of length Δt, the probability of one failure is $\lambda\,\Delta t + o(\Delta t)$ and the probability of two or more is $o(\Delta t)$, where $o(\Delta t)$ denotes terms that vanish faster than Δt.

The rate λ is a constant rate of occurrence of failures, in h^{-1}. Since the expected number of failures in a short interval is $\lambda\,\Delta t$ at every time, the ROCOF and the cumulative intensity are

$$w(t) = \lambda, \qquad W(t) = \lambda t. \tag{17.8}$$

Because the increments are independent of the past, the conditional intensity does not depend on the history either, $\lambda^*(t \mid \mathcal{H}_t) = \lambda$: a system that has just failed and been repaired and a system that has run for a long time without failure have the same rate of failure. The HPP is the standard baseline model for repairable systems, against which trends are judged.

17.3.2 Exponential interfailure times

The interfailure times of an HPP follow from property 3. Let $P_0(x)$ be the probability of no failure in an interval $(t, t+x]$ of length x. For no failure to occur in $(t, t+x+\Delta x]$, none may occur in $(t, t+x]$ and none in the following Δx; by the independence of increments,

$$P_0(x + \Delta x) = P_0(x)\,[1 - \lambda\,\Delta x + o(\Delta x)].$$

Subtracting $P_0(x)$, dividing by Δx and letting $\Delta x \to 0$ gives the differential equation $\mathrm{d}P_0/\mathrm{d}x = -\lambda P_0(x)$, with $P_0(0) = 1$, whose solution is

$$\Pr[N(t+x) - N(t) = 0] = P_0(x) = \mathrm{e}^{-\lambda x}. \tag{17.9}$$

The waiting time X from any time to the next failure exceeds x exactly when no failure occurs in the next x hours, so $\Pr(X > x) = \mathrm{e}^{-\lambda x}$: the waiting time is exponentially distributed with

rate λ. Applied at a failure epoch T_j, after which the increments are independent of everything up to T_j, the same argument shows that each interfailure time X_{j+1} is exponential with rate λ and independent of the earlier ones. The interfailure times of an HPP are therefore independent and identically distributed exponential random variables, with mean

$$\mathbb{E}[X] = \frac{1}{\lambda}. \tag{17.10}$$

The converse also holds: a counting process whose interfailure times are independent exponential variables with rate λ is an HPP. The memoryless property of the exponential distribution (Chapter 6) is what makes the rate independent of the history.

The mean interfailure time $1/\lambda$ is measured on the clock of the process. On the operating-time clock it is the mean operating time between failures, the mean up time, which this book calls the MTTF of a repairable item (Section 3.4.1). Many sources call the same quantity the MTBF. In this book's convention the MTBF is the mean elapsed time from failure to failure, repair included, MTBF = MTTF + MTTR, and it equals $1/\lambda$ only when the HPP is written on the calendar clock or when repair times are negligible. A reported MTBF must therefore state its clock.

17.3.3 The Poisson count distribution

The distribution of the number of failures by time t follows from the exponential interfailure times in two steps.

Step 1: the distribution of the k-th failure epoch. The epoch $T_k = X_1 + \cdots + X_k$ is a sum of k independent exponential variables. Let f_k denote its density. For $k = 1$, $f_1(s) = \lambda \mathrm{e}^{-\lambda s}$. Suppose that $f_k(s) = \lambda^k s^{k-1} \mathrm{e}^{-\lambda s}/(k-1)!$. Since $T_{k+1} = T_k + X_{k+1}$ with X_{k+1} independent of T_k, the density of the sum is the convolution

$$f_{k+1}(s) = \int_0^s f_k(u)\, \lambda \mathrm{e}^{-\lambda(s-u)}\, \mathrm{d}u = \frac{\lambda^{k+1}\mathrm{e}^{-\lambda s}}{(k-1)!} \int_0^s u^{k-1}\, \mathrm{d}u = \frac{\lambda^{k+1} s^k \mathrm{e}^{-\lambda s}}{k!},$$

because the product of the two exponentials, $\mathrm{e}^{-\lambda u}\mathrm{e}^{-\lambda(s-u)} = \mathrm{e}^{-\lambda s}$, does not depend on u. By induction the formula holds for every $k \geq 1$; it is the Erlang density, a gamma density with integer shape k.

Step 2: the count. Exactly k failures occur by time t when the k-th failure has occurred and the $(k+1)$-th has not: $\{N(t) = k\} = \{T_k \leq t,\ X_{k+1} > t - T_k\}$. Conditioning on $T_k = s$ and using the independence of X_{k+1},

$$\Pr[N(t) = k] = \int_0^t f_k(s)\, \mathrm{e}^{-\lambda(t-s)}\, \mathrm{d}s = \frac{\lambda^k \mathrm{e}^{-\lambda t}}{(k-1)!} \int_0^t s^{k-1}\, \mathrm{d}s = \frac{(\lambda t)^k}{k!}\, \mathrm{e}^{-\lambda t},$$

and for $k = 0$ the probability is $\Pr(T_1 > t) = \mathrm{e}^{-\lambda t}$, the same formula. Hence

$$\Pr[N(t) = k] = \frac{(\lambda t)^k}{k!}\, \mathrm{e}^{-\lambda t}, \qquad k = 0, 1, 2, \ldots, \tag{17.11}$$

the Poisson distribution with mean λt, written $N(t) \sim \text{Poisson}(\lambda t)$. By stationarity of the increments the same holds for any interval of length t: $N(s+t) - N(s) \sim \text{Poisson}(\lambda t)$.

The mean and variance follow from the series of the exponential. The mean is

$$\mathbb{E}[N(t)] = \sum_{k=1}^{\infty} k\, \frac{(\lambda t)^k}{k!}\, \mathrm{e}^{-\lambda t} = \lambda t\, \mathrm{e}^{-\lambda t} \sum_{k=1}^{\infty} \frac{(\lambda t)^{k-1}}{(k-1)!} = \lambda t,$$

which is $W(t)$ of equation (17.8). In the same way $\mathbb{E}[N(t)(N(t)-1)] = (\lambda t)^2$, so that

$$\operatorname{Var}[N(t)] = \mathbb{E}[N(N-1)] + \mathbb{E}[N] - (\mathbb{E}[N])^2 = (\lambda t)^2 + \lambda t - (\lambda t)^2 = \lambda t. \tag{17.12}$$

For an HPP the mean and the variance of the number of failures in a fixed interval are equal. Failure counts that vary from interval to interval much more than their mean, which is common in fleets of unequal equipment or under changing conditions, are a first sign that a single HPP does not describe the data.

17.3.4 Rate estimate for the pump fleet

The likelihood of an HPP for a system observed on $(0, T_{\text{obs}}]$, with failures at $t_1 < \cdots < t_n$, is the joint density of the observed interfailure times multiplied by the probability that no failure occurred after the last one,

$$\prod_{j=1}^{n} \lambda \, \mathrm{e}^{-\lambda(t_j - t_{j-1})} \times \mathrm{e}^{-\lambda(T_{\text{obs}} - t_n)} = \lambda^n \, \mathrm{e}^{-\lambda T_{\text{obs}}},$$

with $t_0 = 0$; the exponents add to the whole window because the interfailure times and the final failure-free interval tile $(0, T_{\text{obs}}]$. The final interval is a right-censored interfailure time in the sense of Chapter 7. For K independent systems the likelihoods multiply, and with r the total number of failures and T_{tot} the total exposure,

$$L(\lambda) = \lambda^r \, \mathrm{e}^{-\lambda T_{\text{tot}}}, \qquad \ell(\lambda) = r \ln \lambda - \lambda \, T_{\text{tot}},$$

which is equation (7.7) of Chapter 7. Setting the derivative $r/\lambda - T_{\text{tot}}$ to zero gives the maximum likelihood estimate of equation (7.8), $\hat{\lambda} = r / T_{\text{tot}}$. For the pumps,

$$\hat{\lambda} = \frac{16}{15\,000\,\text{h}} = 1.0667 \times 10^{-3}\,\text{h}^{-1}, \qquad \frac{1}{\hat{\lambda}} = 937.5\,\text{h},$$

an estimated mean of 937.5 operating hours between failures. Its standard error, from equation (7.9), is $\hat{\lambda}/\sqrt{r} = 2.667 \times 10^{-4}\,\text{h}^{-1}$, a quarter of the estimate. Because the five pumps were observed over a window fixed in advance and failures arrive as a Poisson process in the accumulated exposure, the time-terminated chi-square limits of equation (7.10) apply here in the exact form derived in Section 7.5.3, conservative in coverage. With $\chi^2_{32,0.025} = 18.291$ and $\chi^2_{34,0.975} = 51.966$,

$$\lambda_L = \frac{18.291}{30\,000\,\text{h}} = 6.097 \times 10^{-4}\,\text{h}^{-1}, \qquad \lambda_U = \frac{51.966}{30\,000\,\text{h}} = 1.732 \times 10^{-3}\,\text{h}^{-1},$$

so the mean operating time between failures lies between $1/\lambda_U = 577\,\text{h}$ and $1/\lambda_L = 1640\,\text{h}$ with 95 % confidence. These numbers are a fitted average under the HPP assumption. They do not establish that the rate is constant; that is a separate question, answered in Section 17.6.4.

17.3.5 Predicting failures in the next 500 hours

Under the HPP, the number of failures of one pump in any further $\tau = 500$ operating hours is Poisson with mean

$$\Delta W = \mathbb{E}[N(t+500) - N(t)] = \hat{\lambda} \times 500 = 1.0667 \times 10^{-3} \times 500 = 0.5333,$$

whatever the pump's present age t. The probability of no failure is

$$\Pr(\Delta N = 0) = \mathrm{e}^{-0.5333} = 0.5866,$$

where $\Delta N = N(t+500) - N(t)$, so the probability of at least one failure is $1 - 0.5866 = 0.4134$. The probability of at least two failures uses the first two terms of equation (17.11),

$$\Pr(\Delta N \geq 2) = 1 - \mathrm{e}^{-\Delta W}(1 + \Delta W) = 1 - 0.5866 \times 1.5333 = 0.1005.$$

A counting-process model gives probabilities for any number of recurrent failures, not only for the next one. Table 17.4 collects the probabilities at the estimate and at the two confidence limits of Section 17.3.4. Each row is one value of the rate; the columns give the expected number of failures in 500 h and the probabilities of none, exactly one, and two or more. The spread between the rows is the epistemic uncertainty of the rate carried into the prediction: the probability of at least one failure in the next 500 h ranges from 0.263 to 0.579 around the point value 0.413.

Table 17.4: Predicted failures of one pump in the next 500 operating hours under the HPP, at the estimated rate and at its 95 % chi-square limits. Equation (17.11) with λt replaced by $\Delta W = 500\lambda$.

Rate λ ($10^{-3}\,\mathrm{h}^{-1}$)	ΔW	$\Pr(\Delta N = 0)$	$\Pr(\Delta N = 1)$	$\Pr(\Delta N \geq 2)$
lower limit, 0.6097	0.3048	0.7372	0.2247	0.0380
estimate, 1.0667	0.5333	0.5866	0.3129	0.1005
upper limit, 1.7322	0.8661	0.4206	0.3643	0.2151

17.4 The nonhomogeneous Poisson process

A constant rate of occurrence of failures excludes the two behaviors that matter most for maintenance: deterioration and improvement. This section explains why a constant rate can be inadequate, defines the nonhomogeneous Poisson process and derives its interval counts, introduces the power-law process, and works an example of a deteriorating pump.

17.4.1 Limits of a constant intensity

An HPP assumes that the rate of occurrence of failures is the same at every system age. A repairable system may instead deteriorate: mechanical wear accumulates, corrosion progresses, repeated thermal cycles damage insulation, and repairs do not fully restore the degraded components. These are the mechanisms of Chapter 4, acting on a system whose repairs leave part of the damage in place. A system may also improve: design defects are eliminated, maintenance procedures improve, software faults are corrected and weak components are replaced by upgraded ones. The improvement of a system under a program of corrective changes is called *reliability growth*.

In either case the expected number of failures in an interval of fixed length changes with the age of the system, and a model with a single constant rate predicts the same number at every age. Over a long horizon it understates the failures of a deteriorating system late in life and overstates those of an improving one. Representing such trends requires a rate of occurrence of failures that is a function of time.

17.4.2 Definition and interval counts

A *nonhomogeneous Poisson process* (NHPP) with intensity $w(t)$ is a counting process with three properties: $N(0) = 0$; independent increments; and, in a short interval $(t, t + \Delta t]$, a probability $w(t)\,\Delta t + o(\Delta t)$ of one failure and $o(\Delta t)$ of two or more. It differs from the HPP only in that the

increments are no longer stationary: the intensity $w(t) \geq 0$ varies with the age t. Its cumulative intensity is

$$W(t) = \int_0^t w(u)\,\mathrm{d}u, \tag{17.13}$$

and $w(t)$ is both its ROCOF and, since the increments are independent of the history, its conditional intensity: the intensity depends on the age of the system, on the calendar or the accumulated operating clock, but not on the particular sequence of earlier failure times.

The distribution of the number of failures in an interval $(t_a, t_b]$ follows from the same argument as for the HPP. Let $P_k(t) = \Pr[N(t) - N(t_a) = k]$ for $t \geq t_a$. For the count to be k at $t + \Delta t$, either it was k at t and no failure occurred in the next Δt, or it was $k-1$ and one failure occurred:

$$P_k(t + \Delta t) = P_k(t)\,[1 - w(t)\Delta t] + P_{k-1}(t)\,w(t)\Delta t + o(\Delta t),$$

with $P_{-1} \equiv 0$. Letting $\Delta t \to 0$, $\mathrm{d}P_k/\mathrm{d}t = -w(t)\,[P_k(t) - P_{k-1}(t)]$, with $P_0(t_a) = 1$ and $P_k(t_a) = 0$ for $k \geq 1$. The substitution $u = W(t) - W(t_a)$, with $\mathrm{d}u = w(t)\,\mathrm{d}t$, removes the time-varying coefficient: $\mathrm{d}P_k/\mathrm{d}u = -P_k + P_{k-1}$. Its solution is $P_k = u^k \mathrm{e}^{-u}/k!$, as direct differentiation confirms, $\frac{\mathrm{d}}{\mathrm{d}u}(u^k \mathrm{e}^{-u}/k!) = u^{k-1}\mathrm{e}^{-u}/(k-1)! - u^k \mathrm{e}^{-u}/k! = P_{k-1} - P_k$. Therefore

$$\Pr[N(t_b) - N(t_a) = k] = \frac{[W(t_b) - W(t_a)]^k}{k!}\,\mathrm{e}^{-[W(t_b) - W(t_a)]}, \qquad k = 0, 1, 2, \ldots : \tag{17.14}$$

the number of failures in an interval is Poisson with mean equal to the increase of the cumulative intensity over the interval, $N(t_b) - N(t_a) \sim \mathrm{Poisson}\big(W(t_b) - W(t_a)\big)$, and counts in disjoint intervals are independent. In particular, the probability of no failure in the interval is

$$\Pr[N(t_b) - N(t_a) = 0] = \mathrm{e}^{-[W(t_b) - W(t_a)]}. \tag{17.15}$$

With $w(t) = \lambda$ the increase is $\lambda(t_b - t_a)$ and the HPP results are recovered. The new feature is that the prediction depends on where the interval lies, through the current age t_a: under an HPP only the length of the interval matters, while under an NHPP two intervals of equal length can have different expected numbers of failures.

17.4.3 The power-law process

The most widely used NHPP is the *power-law process*. Its cumulative intensity is written in this book as

$$W(t) = \left(\frac{t}{\eta}\right)^b, \qquad w(t) = \frac{\mathrm{d}W}{\mathrm{d}t} = \frac{b}{\eta}\left(\frac{t}{\eta}\right)^{b-1}, \tag{17.16}$$

where $\eta > 0$ is a time-scale parameter, in hours, and $b > 0$ a dimensionless shape parameter. The scale has a direct reading: the expected cumulative number of failures reaches one at $t = \eta$, $W(\eta) = 1$. The same model is often written with a coefficient, $W(t) = at^b$ and $w(t) = abt^{b-1}$, with $a = \eta^{-b}$ in h^{-b}; the two forms are equivalent. The power-law process is the model behind the Duane and Crow–AMSAA reliability-growth models, where it describes the falling ROCOF of a system under development. Its cumulative intensity has the functional form of the cumulative hazard of a Weibull lifetime with shape b and scale η (Chapter 6), and Section 17.5.3 shows when the two coincide. The interfailure times of a power-law process are not, however, Weibull distributed, and they are not independent: only the time to the first failure has the Weibull distribution $\Pr(T_1 > t) = \mathrm{e}^{-(t/\eta)^b}$.

Table 17.5: The shape parameter of the power-law process, equation (17.16).

Shape	Intensity trend	Interpretation
$0 < b < 1$	decreasing	fewer failures per unit time as the system ages: an improving system
$b = 1$	constant	HPP behavior, $W(t) = t/\eta$, $\lambda = 1/\eta$
$b > 1$	increasing	more failures per unit time as the system ages: a deteriorating system

17.4.4 Improving and deteriorating systems

The shape parameter decides the direction of the trend, since $w(t) \propto t^{b-1}$. Table 17.5 lists the three cases: each row is a range of the shape, the resulting trend of the intensity and its interpretation.

Figure 17.5 draws three power-law processes with the same scale, $\eta = 1000\,\text{h}$, and shapes $b = 0.6$, 1 and 1.4. In both panels the horizontal axis is age in hours. The left panel is the cumulative intensity $W(t)$, with dots at the values $W(500)$ to $W(3000)$ listed in the caption; the right panel is the intensity $w(t)$, in $10^{-3}\,\text{h}^{-1}$. All three curves of the left panel pass through $W(1000) = 1$, and they separate afterward. The improving system ($b = 0.6$) accumulates failures more and more slowly, reaching 1.933 by 3000 h; the deteriorating system ($b = 1.4$) accumulates them faster and faster, reaching 4.656. The right panel shows the cause: the intensity of the improving system falls from $0.79 \times 10^{-3}\,\text{h}^{-1}$ at 500 h to $0.39 \times 10^{-3}\,\text{h}^{-1}$ at 3000 h, while that of the deteriorating system rises from 1.06×10^{-3} to $2.17 \times 10^{-3}\,\text{h}^{-1}$. At $t = \eta$ the three intensities are b/η, so they differ even where the cumulative curves cross.

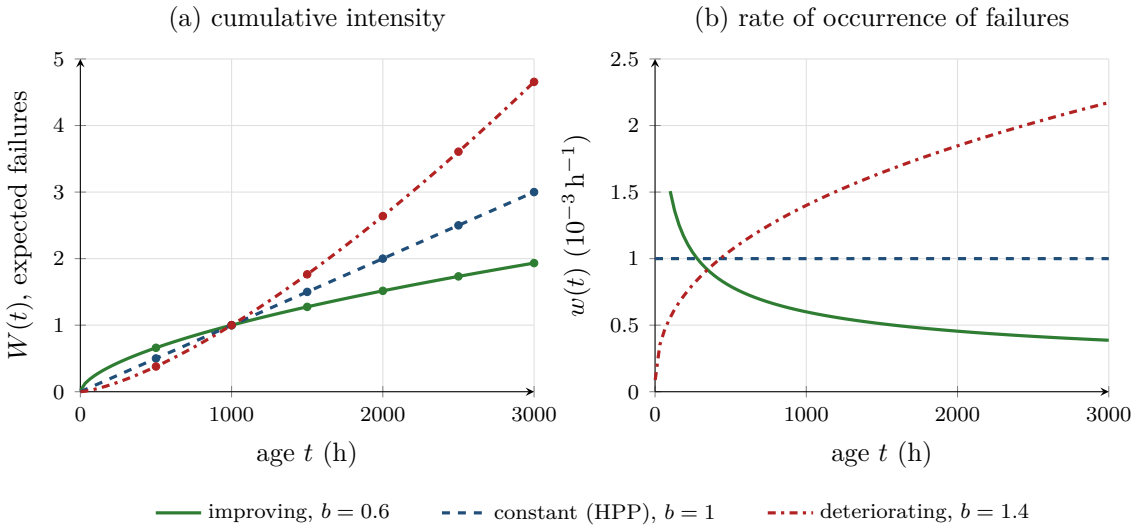


Figure 17.5: Power-law processes, equation (17.16), with $\eta = 1000\,\text{h}$ and $b = 0.6$, 1, 1.4. (a) $W(t) = (t/1000)^b$; dots at 500, 1000, ..., 3000 h: 0.660, 1, 1.275, 1.516, 1.733, 1.933 for $b = 0.6$; 0.5, 1, ..., 3 for $b = 1$; 0.379, 1, 1.764, 2.639, 3.607, 4.656 for $b = 1.4$. (b) $w(t) = (b/1000)(t/1000)^{b-1}$. All three have $W(1000) = 1$; the HPP is the case $b = 1$.

An increasing cumulative intensity does not by itself indicate deterioration. $W(t)$ counts expected failures, and a count can only increase; every curve in panel (a) rises. Deterioration is indicated by an increasing slope, the ROCOF $w(t)$ of panel (b), and improvement by a decreasing one. On a plot of the empirical $\hat{W}(t)$ the signature of deterioration is a curve that bends upward, and of improvement one that bends downward.

17.4.5 Worked example: a deteriorating pump

Suppose the failure process of a pump follows the power-law process

$$W(t) = \left(\frac{t}{1000}\right)^{1.4},$$

with t in operating hours, so that $\eta = 1000\,\mathrm{h}$ and $b = 1.4 > 1$: the intensity increases with age. The cumulative intensity at three ages is

$$W(1000) = 1, \qquad W(2000) = 2^{1.4} = 2.639, \qquad W(3000) = 3^{1.4} = 4.656.$$

The expected numbers of failures in two successive intervals of equal length are, by equation (17.14),

$$\mathbb{E}[N(2000) - N(1000)] = W(2000) - W(1000) = 2.639 - 1 = 1.639,$$
$$\mathbb{E}[N(3000) - N(2000)] = W(3000) - W(2000) = 4.656 - 2.639 = 2.017.$$

The intervals have the same duration, but the later one has 23 % more expected failures; that is the meaning of a deteriorating failure intensity. The ROCOF itself rises from $w(2000) = (1.4/1000) \times 2^{0.4} = 1.847 \times 10^{-3}\,\mathrm{h}^{-1}$ to $w(3000) = 2.173 \times 10^{-3}\,\mathrm{h}^{-1}$.

Now suppose the pump has reached an operating age of $t = 2000\,\mathrm{h}$, and the probability of no failure in the next 500 h, from 2000 to 2500 h, is required. By equation (17.15), with $W(2500) = 2.5^{1.4} = 3.607$,

$$\Pr(\Delta N = 0) = \mathrm{e}^{-[W(2500) - W(2000)]} = \mathrm{e}^{-(3.607 - 2.639)} = \mathrm{e}^{-0.968} = 0.380,$$

so the probability of at least one failure in those 500 h is 0.620; the probability of exactly one is $0.968 \times 0.380 = 0.368$, and of two or more 0.252. An HPP with the same scale, $W(t) = t/1000$, would give an expected count of 0.5 and a probability of at least one failure of $1 - \mathrm{e}^{-0.5} = 0.393$ for the same interval, and for any other interval of 500 h. Under the deteriorating model the current age of the system materially changes the probability of future failures, which a constant-rate model cannot represent.

17.5 Repair assumptions

The word repair covers interventions with very different effects on the future of a system. This section states the three idealized assumptions about what a repair does, formulates them with the virtual age of the system, separates the repair assumption from the choice of a failure-process model, and separates both from availability.

17.5.1 Repair as an intervention

A failed item may be replaced by an entirely new unit; restored to the condition it had immediately before the failure, for example by replacing one small part of a large machine; partially refurbished; or upgraded so that its future propensity to fail decreases. These interventions have different consequences for future reliability, and three idealized assumptions describe the range: perfect repair, minimal repair and imperfect repair.

The three can be written with one device, the *virtual age*. Let $\lambda(x)$ be the hazard function of a new item at age x, on the operating-time clock, and let V_n be the virtual age of the system just after its n-th repair, the age of a new item that would have the same condition, with $V_0 = 0$. Between the n-th and the $(n+1)$-th failure the system ages normally from V_n, and its conditional intensity is the hazard at the current virtual age,

$$\lambda^*(t \mid \mathcal{H}_t) = \lambda(V_n + t - T_n), \qquad T_n < t \le T_{n+1}, \tag{17.17}$$

with $T_0 = 0$. The repair assumption is a rule for V_n: perfect repair sets $V_n = 0$, minimal repair sets $V_n = T_n$, and imperfect repair sets it in between. Figure 17.6 draws the three rules. The horizontal axis is operating time, with three failure epochs marked; the vertical axis is the virtual age, both in arbitrary units. The same failure epochs are used for all three rules so that only the rules differ; in reality the epochs themselves would differ, because the rule changes the intensity. Under minimal repair the virtual age grows with operating time and is never reduced; under perfect repair it drops to zero at each failure; under the imperfect rule drawn, each repair removes half of the age gained since the previous repair.

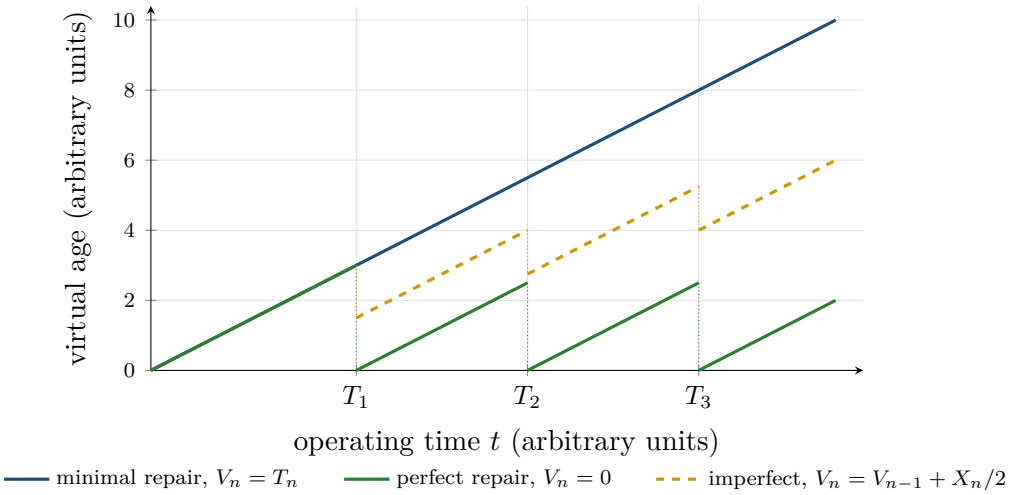


Figure 17.6: Virtual age under the three repair assumptions, equation (17.17), illustrative and in arbitrary units, with failures at $T_1 = 3$, $T_2 = 5.5$ and $T_3 = 8$; the three rules coincide before T_1. Minimal repair keeps the age ($V_n = T_n$); perfect repair resets it to zero; the imperfect rule drawn removes half of the age gained since the previous repair, giving $V_1 = 1.5$, $V_2 = 2.75$ and $V_3 = 4$.

17.5.2 Perfect repair

Under *perfect repair* the system is restored as good as new: its virtual age is reset to zero at every failure, $V_n = 0$. By equation (17.17) the conditional intensity after the j-th failure is $\lambda(t - T_j)$, the hazard of a new item evaluated at the time since the failure. The probability that the next interfailure time exceeds x is therefore, by equation (5.12),

$$\Pr(X_{j+1} > x) = \exp\left[-\int_0^x \lambda(u)\,\mathrm{d}u\right] = R(x), \tag{17.18}$$

the reliability function of a new item, whatever the history before the failure. When the restorations are independent, the interfailure times are independent and identically distributed with the lifetime distribution of a new item. Such a counting process is a *renewal process*, the subject of Chapter 18. Replacement of the failed item by a new one of the same design is the usual physical realization; repair of a component that is the only part that ages is another.

17.5.3 Minimal repair

Under *minimal repair* the repair restores the function but leaves the system at the effective age it had immediately before the failure, as bad as old: $V_n = T_n$. Then $V_n + t - T_n = t$, and equation (17.17) becomes

$$\lambda^*(t \mid \mathcal{H}_t) = \lambda(t)\,: \tag{17.19}$$

the conditional intensity is the hazard of a new item evaluated at the system's total age, whatever the number and times of the previous failures. Since it depends on time only, not on the history, the failure process is an NHPP with intensity $w(t) = \lambda(t)$ and cumulative intensity $W(t) = \Lambda(t)$, the cumulative hazard of the item. Minimal repair is the natural model for a complex system in which one of many parts fails and is replaced: the repaired system is no younger than before, because all its other parts kept their age.

This is the connection announced in Section 17.4.3. If the item's lifetime is Weibull with shape β and scale η, its cumulative hazard is $(t/\eta)^\beta$, and under minimal repair the failure process is the power-law process of equation (17.16) with $b = \beta$ and $\eta = \eta$. A wear-out item ($\beta > 1$) under minimal repair gives a deteriorating system ($b > 1$).

17.5.4 Imperfect repair

Under *imperfect repair* the system is restored to an intermediate condition, $0 < V_n < T_n$: younger than just before the failure, older than new. The conditional intensity of equation (17.17) then depends on the history through the virtual ages, and the failure process is neither a renewal process nor an NHPP but a history-dependent recurrence process. The rule drawn in Figure 17.6, $V_n = V_{n-1} + X_n/2$, removes half of the age gained in each operating interval; with $T_1 = 3$, $T_2 = 5.5$ and $T_3 = 8$ it gives $V_1 = 0 + 3/2 = 1.5$, $V_2 = 1.5 + 2.5/2 = 2.75$ and $V_3 = 2.75 + 2.5/2 = 4$, against 3, 5.5 and 8 under minimal repair and zero under perfect repair.

A repair can thus reduce the virtual age without resetting it to zero, and the fraction it removes is a property of the maintenance practice that can in principle be estimated. Imperfect-repair models of this kind, renewal theory and age-based replacement are developed in Chapter 18.

17.5.5 Repair assumption against failure-process model

The terms HPP and NHPP describe stochastic processes of failure events. The terms perfect, minimal and imperfect repair describe what a repair does to the condition of the system. The two are related, but they are not equivalent, and three cases show the relation.

- Perfect repair of an item with Weibull lifetimes, $\beta \neq 1$, gives a renewal process, which is not an NHPP: its conditional intensity depends on the time since the last failure.
- Minimal repair of an item with an age-dependent hazard gives an NHPP in the classical model, equation (17.19).
- If the item's lifetime is exponential, perfect and minimal repair give the same HPP. Under minimal repair $\lambda^*(t \mid \mathcal{H}_t) = \lambda(t) = \lambda$; under perfect repair the interfailure times are independent exponential variables with rate λ, which by Section 17.3.2 is an HPP with rate λ. The memoryless exponential distribution makes the age to which repair returns the item irrelevant.

The last case has a practical consequence. Observing a constant rate of occurrence of failures does not prove that maintenance performs perfect repair: a constant-rate HPP is compatible with several physical repair mechanisms, and the failure times alone do not identify which one acts. Only when the hazard of the item varies with age do the repair assumptions lead to different failure processes, and even then separating them from failure data alone requires many failures. In the vocabulary of state estimation, the repair mechanism is not observable from the event times without assumptions about the item's hazard; it becomes observable with additional evidence, such as the condition found at each repair.

Principle 17.2. *A failure-process model (HPP, NHPP, renewal) describes the sequence of failure events; a repair assumption (perfect, minimal, imperfect) describes what repair does to the system. The second determines the first, but the first does not in general determine the second.*

17.5.6 Failure intensity and availability

Suppose two plants both experience one equipment failure per 1000 operating hours. Plant A typically repairs the equipment in one hour, plant B in 24 h. Their failure processes may be identical, and their availabilities differ substantially. The long-run fraction of time operating is, for a stationary alternating sequence of up and down periods, the steady-state availability of Chapter 3,

$$A = \frac{\mathrm{MTTF}}{\mathrm{MTTF} + \mathrm{MTTR}} = \frac{\mathrm{MTTF}}{\mathrm{MTBF}}, \tag{17.20}$$

where MTTF is the mean operating time between failures, the mean up time, MTTR the mean down time per failure, and MTBF = MTTF + MTTR the mean time from failure to failure in this book's convention. A source that uses MTBF for the mean up time writes the same result as MTBF/(MTBF + MTTR). Table 17.6 evaluates equation (17.20) with MTTF = 1000 h for three repair durations. The columns are the mean repair time, the MTBF, the steady-state availability, and the expected down time per year of calendar time, $8760\,\mathrm{h} \times (1 - A)$. For plant A, $A = 1000/1001 = 0.99900$; for plant B, $A = 1000/1024 = 0.97656$, with 205 h of downtime per year against 8.8 h.

Table 17.6: Steady-state availability, equation (17.20), for a mean up time of 1000 h and three mean repair times. The rows differ only in repair; the failure process is the same.

MTTR (h)	MTBF (h)	Availability A	Down time per year (h)
1 (plant A)	1001	0.99900	8.8
8	1008	0.99206	69.5
24 (plant B)	1024	0.97656	205.3

These are steady-state results of an alternating renewal model of up and down periods, not consequences of the HPP alone. An HPP counts recurrent failures; it contains no repair durations, and no property of the counting process determines the availability. To compute availability a model of the operating state and of the repair durations is needed in addition, which is why the treatment of repairable systems proceeds from failure counting, in this chapter, to renewal and repair models, Markov models of the operating state, and availability and maintenance optimization in Chapters 18, 19 and 20.

17.6 Estimation, model selection and engineering applications

The models of the previous sections are useful only once their parameters are estimated from failure records and their adequacy is checked. This section derives the likelihood of an NHPP from failure timestamps and the closed-form estimates of the power-law process, tests whether a constant rate is adequate for the pump fleet, lists the conditions under which a counting-process model is credible, and applies the ideas to three engineering systems and to a worked comparison of predictions.

17.6.1 The NHPP likelihood from failure timestamps

Suppose a repairable system is observed from age zero to age T_{obs}, the end of the observation window, and failures are recorded at $0 < t_1 < t_2 < \cdots < t_n \le T_{\mathrm{obs}}$. Assume an NHPP with

intensity $w(t;\boldsymbol{\theta})$ and cumulative intensity $W(t;\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is the vector of parameters. The likelihood is built failure by failure. Given a failure at t_{j-1}, the next failure occurs at t when no failure occurs in (t_{j-1}, t), with probability $\mathrm{e}^{-[W(t)-W(t_{j-1})]}$ by equation (17.15), and one occurs at t, with rate $w(t)$; its density is $w(t)\,\mathrm{e}^{-[W(t)-W(t_{j-1})]}$. After the last failure, the record says that no failure occurred in $(t_n, T_{\mathrm{obs}}]$, with probability $\mathrm{e}^{-[W(T_{\mathrm{obs}})-W(t_n)]}$. Multiplying, with $t_0 = 0$ and $W(0) = 0$, the exponents telescope to $W(T_{\mathrm{obs}})$:

$$L(\boldsymbol{\theta}) = \mathrm{e}^{-W(T_{\mathrm{obs}};\boldsymbol{\theta})} \prod_{j=1}^{n} w(t_j;\boldsymbol{\theta}), \qquad \ell(\boldsymbol{\theta}) = \sum_{j=1}^{n} \ln w(t_j;\boldsymbol{\theta}) - W(T_{\mathrm{obs}};\boldsymbol{\theta}). \tag{17.21}$$

For K independent systems with the same model, each observed to T_{obs}, the log-likelihoods add: the first term runs over all r failures of the fleet and the second becomes $K\,W(T_{\mathrm{obs}};\boldsymbol{\theta})$. With $w = \lambda$ the expression reduces to $r\ln\lambda - \lambda K T_{\mathrm{obs}} = r\ln\lambda - \lambda T_{\mathrm{tot}}$, the HPP log-likelihood of Section 17.3.4.

The two terms have a direct interpretation. The first rewards parameter values that assign a high intensity near the observed failures. The second penalizes values that predict too many failures over the whole window. The observation window, including the failure-free interval after the last failure, is therefore an essential part of the data: two systems with the same failure times, one observed to the last failure and the other for a further 1000 h without failure, carry different evidence about the trend, in the same way that a censored record in Chapter 7 carried evidence about a lifetime.

17.6.2 Maximum likelihood for the power-law process

For the power-law process of equation (17.16), $\ln w(t) = \ln b - b\ln\eta + (b-1)\ln t$. For K systems observed to T_{obs} with r failures in total at ages t_j, equation (17.21) gives

$$\ell(\eta, b) = r\ln b - rb\ln\eta + (b-1)\sum_{j=1}^{r}\ln t_j - K\left(\frac{T_{\mathrm{obs}}}{\eta}\right)^{b}. \tag{17.22}$$

The maximum is found in two steps.

The scale. The derivative with respect to η is

$$\frac{\partial\ell}{\partial\eta} = -\frac{rb}{\eta} + \frac{Kb}{\eta}\left(\frac{T_{\mathrm{obs}}}{\eta}\right)^{b} = \frac{b}{\eta}\left[K\left(\frac{T_{\mathrm{obs}}}{\eta}\right)^{b} - r\right].$$

Setting it to zero gives $K\,W(T_{\mathrm{obs}}) = r$: at the estimate, the fitted expected number of failures over the window equals the observed number. Solving for the scale,

$$\hat\eta = T_{\mathrm{obs}}\left(\frac{K}{r}\right)^{1/\hat b}, \qquad \text{equivalently} \qquad \hat a = \hat\eta^{-\hat b} = \frac{r}{K\,T_{\mathrm{obs}}^{\hat b}}, \tag{17.23}$$

the second form being the coefficient of $W(t) = at^b$.

The shape. The derivative with respect to b is

$$\frac{\partial\ell}{\partial b} = \frac{r}{b} - r\ln\eta + \sum_{j=1}^{r}\ln t_j - K\left(\frac{T_{\mathrm{obs}}}{\eta}\right)^{b}\ln\frac{T_{\mathrm{obs}}}{\eta}.$$

At the estimate, $K(T_{\mathrm{obs}}/\hat\eta)^{\hat b} = r$, so the last term is $r\ln T_{\mathrm{obs}} - r\ln\hat\eta$ and the terms in $\ln\hat\eta$ cancel:

$$0 = \frac{r}{\hat b} + \sum_{j=1}^{r}\ln t_j - r\ln T_{\mathrm{obs}},$$

which gives the closed-form estimate

$$\hat{b} = \frac{r}{\sum_{j=1}^{r} \ln\left(T_{\text{obs}}/t_j\right)}. \tag{17.24}$$

These formulas assume at least one failure, complete recording of every failure timestamp, a common time origin and a common end of observation for all systems, observation that ends at a fixed time rather than at a failure, and the power-law NHPP itself. Systems with different observation windows require the likelihood of equation (17.21) system by system, and the maximum is then found numerically.

For the pumps, $K = 5$, $T_{\text{obs}} = 3000\,\text{h}$ and $r = 16$. The sum in the denominator is $\sum_j \ln(3000/t_j) = 16\ln 3000 - \sum_j \ln t_j = 128.1018 - 114.3725 = 13.7293$, so

$$\hat{b} = \frac{16}{13.7293} = 1.165, \qquad \hat{\eta} = 3000 \times \left(\frac{5}{16}\right)^{1/1.165} = 1106\,\text{h}, \qquad \hat{a} = 2.838 \times 10^{-4}\,\text{h}^{-1.165}.$$

The fitted ROCOF per pump is $w(t) = (\hat{b}/\hat{\eta})(t/\hat{\eta})^{\hat{b}-1}$, which rises from $1.037 \times 10^{-3}\,\text{h}^{-1}$ at 1000 h to $\hat{b}\,r/(KT_{\text{obs}}) = 1.243 \times 10^{-3}\,\text{h}^{-1}$ at 3000 h. Table 17.7 compares the two fitted models with the empirical estimate. The columns are the model, its estimated parameters, its maximized log-likelihood, and the cumulative intensity per pump at 1000, 2000 and 3000 h. Both models reproduce $\hat{W}(3000) = 3.2$, as the scale equation requires; they differ in how they distribute the failures within the window, and both lie below the empirical value at 1000 h, where the fleet had six early failures.

Table 17.7: The HPP and the power-law NHPP fitted to the five pumps, against the empirical mean cumulative function (Table 17.3). Log-likelihoods from equation (17.21).

Model	Estimates	ℓ	$W(1000)$	$W(2000)$	$W(3000)$
empirical	–	–	1.200	2.400	3.200
HPP	$\hat{\lambda} = 1.0667 \times 10^{-3}\,\text{h}^{-1}$	−125.491	1.067	2.133	3.200
power law	$\hat{b} = 1.165$, $\hat{\eta} = 1106\,\text{h}$	−125.313	0.889	1.995	3.200

17.6.3 Interpretation and uncertainty of the shape estimate

The shape estimate is read against the three cases of Table 17.5: $\hat{b} < 1$ suggests a decreasing failure intensity, $\hat{b} \approx 1$ is consistent with a constant one, and $\hat{b} > 1$ suggests an increasing one. For the pumps, $\hat{b} = 1.165$ points to a mild deterioration, while the interval counts 6, 6, 4 of Section 17.2.6 pointed to a mild improvement. The two summaries of the same data point in opposite directions because they weigh the data differently. The shape estimate depends on the ages through $\ln(T_{\text{obs}}/t_j)$, which grows without bound as t_j approaches zero and so weighs the earliest part of the window most heavily, and the fleet had no failure before 300 h; the interval counts weigh the three intervals equally and register the shortfall of the last one. A fitted estimate has uncertainty, and a small dataset can produce $\hat{b} > 1$ when the process has a constant intensity. Before deterioration is concluded, the uncertainty must be quantified and the fit assessed.

For the power-law process with a fixed end of observation the uncertainty of $\hat{b}$ has an exact form, derived in three steps. First, given that n failures occurred in $(0, T_{\text{obs}}]$, the likelihood of equation (17.21) divided by the Poisson probability $W(T_{\text{obs}})^n \mathrm{e}^{-W(T_{\text{obs}})}/n!$ of that count is $n! \prod_j [w(t_j)/W(T_{\text{obs}})]$: the failure ages behave as n independent draws from the distribution function $W(t)/W(T_{\text{obs}}) = (t/T_{\text{obs}})^b$, arranged in order. Second, for one such draw,

$U = \ln(T_{\text{obs}}/t)$ satisfies $\Pr(U > u) = \Pr(t < T_{\text{obs}}\mathrm{e}^{-u}) = \mathrm{e}^{-bu}$, so U is exponential with rate b and $2bU$ is chi-square with two degrees of freedom, as in Section 7.5.3. Third, summing over the r failures of the fleet, which are independent draws given the counts,

$$2b\sum_{j=1}^{r}\ln\frac{T_{\text{obs}}}{t_j} = \frac{2rb}{\hat{b}} \sim \chi^2_{2r}. \tag{17.25}$$

Two consequences follow. Since $\mathbb{E}[1/\chi^2_{2r}] = 1/(2r-2)$, $\mathbb{E}[\hat{b}] = rb/(r-1)$: the estimate is biased upward, and $(r-1)\hat{b}/r$ is unbiased, $15 \times 1.165/16 = 1.093$ for the pumps. And a $100(1-\alpha)\,\%$ confidence interval for b is

$$b_L = \hat{b}\,\frac{\chi^2_{2r,\,\alpha/2}}{2r}, \qquad b_U = \hat{b}\,\frac{\chi^2_{2r,\,1-\alpha/2}}{2r}. \tag{17.26}$$

For the pumps, with $\chi^2_{32,0.025} = 18.291$ and $\chi^2_{32,0.975} = 49.480$, the 95 % interval is $1.165 \times 18.291/32 = 0.666$ to $1.165 \times 49.480/32 = 1.802$. It contains $b = 1$, and it contains both a clearly improving and a clearly deteriorating process.

The same conclusion appears in the likelihood. Maximizing equation (17.22) over η for each fixed b gives the profile log-likelihood of the shape, which relative to its maximum is $\ell_p(b) - \ell_p(\hat{b}) = r[\ln(b/\hat{b}) - b/\hat{b} + 1]$, the same function of $b/\hat{b}$ as the exponential log-likelihood of Figure 7.4 is of $\lambda/\hat{\lambda}$. Figure 17.7 plots it for the pumps. The horizontal axis is the shape b and the vertical axis $\ell_p(b) - \ell_p(\hat{b})$. The curve peaks at $\hat{b} = 1.165$; the dotted line at -1.92 cuts it at 0.684 and 1.833, the 95 % likelihood-ratio interval, close to the exact interval; the HPP, $b = 1$, lies only 0.178 below the peak.

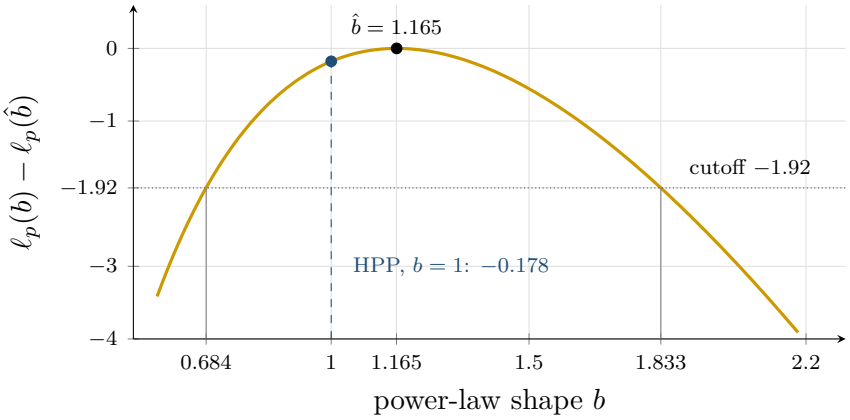


Figure 17.7: Profile log-likelihood of the power-law shape for the five pumps, $\ell_p(b) - \ell_p(\hat{b}) = r[\ln(b/\hat{b}) - b/\hat{b} + 1]$ with $r = 16$ and $\hat{b} = 1.165$. The 95 % likelihood-ratio interval, where the curve lies above -1.92, runs from 0.684 to 1.833 and contains the HPP, $b = 1$.

17.6.4 Testing the adequacy of the HPP

For the pump fleet the HPP estimates one average recurrence rate, $\hat{\lambda} = 16/15\,000$ h; the power-law NHPP adds one trend parameter. A practical analysis decides between them in seven steps:

1. plot the cumulative failure counts, or $\hat{W}(t)$, against operating age (Figure 17.4);
2. compare the observed counts in intervals of equal exposure;
3. fit an HPP;

4. fit a power-law NHPP;
5. examine whether the estimated trend differs meaningfully from $b = 1$;
6. check the residual patterns and the predictions of the fitted models against the data;
7. investigate the possible physical causes of any apparent trend.

Steps 1, 3 and 4 are done. For step 2, the counts 6, 6, 4 in three intervals of 5000 pump-hours each are compared with the $16/3 = 5.333$ expected in each under a constant rate. The Pearson statistic is $\sum(\text{observed} - \text{expected})^2/\text{expected} = (0.667^2 + 0.667^2 + 1.333^2)/5.333 = 0.50$ with two degrees of freedom, far below the 5 % critical value 5.99; the probability of a value at least this large under a constant rate is 0.78.

For step 5, equation (17.25) gives an exact test. Under the HPP, $b = 1$, the statistic $2\sum_j \ln(T_{\text{obs}}/t_j)$ is chi-square with $2r$ degrees of freedom; this is the trend test of the United States Military Handbook for reliability growth (MIL-HDBK-189). Small values indicate failures concentrated late in the window, deterioration; large values, failures concentrated early, improvement. For the pumps the statistic is $2 \times 13.729 = 27.46$, inside the two-sided 5 % acceptance region from $\chi^2_{32,0.025} = 18.29$ to $\chi^2_{32,0.975} = 49.48$; the two-sided probability of a deviation at least this large is 0.61. The likelihood-ratio test agrees: twice the gain in maximized log-likelihood, $2[-125.313 - (-125.491)] = 0.356$, lies far below the critical value 3.84 of the chi-square distribution with one degree of freedom, with a probability of 0.55. The HPP is adequate for these data: neither the apparent decline of the last interval nor the mild increase of the power-law fit is distinguishable from random variation with sixteen failures. A preliminary graphical check of the same question is the Duane plot, which graphs $\ln \hat{W}(t)$ against $\ln t$: for a power-law process $\ln W(t) = b \ln t - b \ln \eta$ is a straight line of slope b, and a slope visibly different from one indicates a trend.

Step 7 is engineering, not statistics. Possible causes of an increasing recurrence intensity include wear, changing duty cycles and inadequate repairs. Possible causes of a decreasing intensity include improved maintenance, the replacement of defective parts and design corrections. A fitted statistical trend alone does not identify which mechanism produced it; the failure modes and the condition found at each repair do.

17.6.5 Model credibility

Counting-process models have limitations that a fitted rate does not display. Table 17.8 lists the main ones: each row is an issue, and the second column says why it matters for the estimated intensity.

Each issue is a mismatch between the model and the physical system or between the model and the evidence. Heterogeneity is the counterpart for recurrent events of the mixed population of Chapter 5: a fleet of units with different constant intensities, some of which are retired early, can show a decreasing aggregate intensity that no unit has. Dependence through common causes violates the independence across systems on which the pooled likelihood rests, the subject of Chapters 13 and 14. The last row is the point of Section 17.5.5: a good fit to event times is not evidence about the repair mechanism.

> **Principle 17.3.** *A counting-process model is useful only when its time scale, its exposure, its events and its observation mechanism are clearly defined. A mathematically fitted rate is meaningful only when the recorded events and the clock correspond to the system behavior the analysis is meant to describe.*

Table 17.8: Issues that limit the credibility of a recurrent-failure analysis.

Issue	Why it matters
Changing operating exposure	calendar time may not represent the actual duty
Repair duration	failures generally cannot occur while equipment is unavailable and not exposed
Heterogeneous fleets	units with different intensities may produce a misleading aggregate trend
Preventive replacement	changes the virtual age and the recurrence behavior
Changing operating conditions	the intensity may depend on weather, load or environment
Common-cause failures	event counts across units may be statistically dependent
Incomplete failure records	undetected failures and censoring can bias the estimates
Repair effectiveness	an NHPP may fit the event times without describing the repair mechanism correctly

17.6.6 Engineering applications

Three systems show how the counting-process view enters engineering practice.

Transmission-line outages. Consider a transmission line i that experiences recurrent outages and restorations, with $N_i(t)$ its cumulative number of outages. Its conditional intensity may depend on its age, the weather, vegetation exposure and maintenance, $\lambda_i^*(t \mid \mathcal{H}_t, \mathbf{z}(t))$, where the covariate vector $\mathbf{z}(t)$ describes the relevant operating and environmental conditions; severe weather can raise the outage intensity for its duration. A stationary HPP underestimates outage occurrence during storms and overestimates it in benign conditions. A more appropriate model is an NHPP with a condition-dependent intensity,

$$w_i(t) = \lambda_{0,i} \exp\bigl[\boldsymbol{\gamma}^{\mathsf{T}} \mathbf{z}(t)\bigr], \tag{17.27}$$

a log-linear effect of the covariates on a baseline rate $\lambda_{0,i}$, the counterpart for recurrent events of the proportional hazards model of Section 5.6.3. As an illustration, let $\lambda_{0,i} = 0.5\,\mathrm{yr}^{-1}$ in benign weather and let storms, 200 h a year, multiply the intensity by $\exp(\boldsymbol{\gamma}^{\mathsf{T}}\mathbf{z}) = 30$. The expected outages per year are $0.5 \times 8560/8760 + 0.5 \times 30 \times 200/8760 = 0.489 + 0.342 = 0.831$. An HPP with that annual rate predicts $0.831 \times 24/8760 = 0.0023$ expected outages during a 24 h storm, while the covariate model predicts $0.5 \times 30 \times 24/8760 = 0.041$, eighteen times more, and in benign weather the HPP rate is 66 % too high. The parameters must be estimated from suitable outage and weather data, by the likelihood of equation (17.21) with W integrated along the recorded weather, and the shared weather exposure of neighboring lines makes their outages dependent, which connects recurrent outage analysis to the common-cause models of Chapters 13 and 14.

Data center cooling maintenance. Suppose the failure intensity of the chillers of the cooling plant, three units of which two are required, increases with operating age, and a power-law NHPP fitted to their records gives $\hat{b} > 1$, credibly so by the tests of Section 17.6.4. Failures then become more frequent as the chillers accumulate exposure, and the probability that two are down at once grows with them. Possible decisions include increasing condition monitoring, performing targeted maintenance, replacing components before severe deterioration, installing additional capacity, and investigating the distribution of load among the chillers, since on the

operating-time clock a chiller that carries more load ages faster. The best decision depends on the repair cost, the consequences of downtime, and whether maintenance actually reduces the future failure intensity, that is, on the repair assumption of Section 17.5. The recurrence model estimates the event behavior; a maintenance optimization model, Chapter 20, evaluates the interventions.

Battery energy storage systems. The recurrent events of a battery energy storage system include inverter trips, cooling system faults, communication interruptions, battery management alarms and protection system operations. These event classes have different mechanisms, and pooling every recorded alarm into one counting process can hide the information that matters. A better approach defines a separate counting process for each relevant class, $N_{\text{inverter}}(t)$, $N_{\text{cooling}}(t)$, $N_{\text{protection}}(t)$, and connects each to its operational outcome, its repair state and its effect on availability. As an illustration, let inverter trips follow a power-law process with $\eta = 1000\,\text{h}$ and $b = 1.4$ and communication interruptions one with $\eta = 1000\,\text{h}$ and $b = 0.6$. Over three successive 1000 h intervals the expected inverter trips are 1.000, 1.639 and 2.017, a rise of 102 %, and the expected communication interruptions 1.000, 0.516 and 0.417, a fall of 58 %. The pooled process, whose intensity is the sum of the two when they are independent, expects 2.000, 2.155 and 2.434, a rise of only 22 % that hides a strongly deteriorating inverter behind an improving communication link.

17.6.7 Expected counts under constant and power-law intensity

The choice of failure-process model changes the predictions on which maintenance is planned. Table 17.9 compares them. The model is the power-law process $W(t) = (t/\eta)^b$ with $\eta = 1000\,\text{h}$; the HPP is its case $b = 1$, $W(t) = t/\eta$, with rate $1/\eta = 10^{-3}\,\text{h}^{-1}$, and both have $W(\eta) = 1$, though for $b \neq 1$ their intensities at η, b/η and $1/\eta$, differ. For a system of current operating age t and a prediction horizon τ, the expected number of failures is $\Delta W = W(t+\tau) - W(t)$ and the probability of at least one is $1 - \text{e}^{-\Delta W}$, by equation (17.15). Each row is a shape and an age; the columns give the expected count and the probability of at least one failure for horizons of 500 and 1000 h. One row of the arithmetic, for $b = 1.6$, $t = 4000\,\text{h}$ and $\tau = 500\,\text{h}$: $\Delta W = (4500/1000)^{1.6} - (4000/1000)^{1.6} = 11.095 - 9.190 = 1.906$, and the probability of at least one failure is $1 - \text{e}^{-1.906} = 1 - 0.149 = 0.851$.

Table 17.9: Expected failures ΔW and probability of at least one failure over the next τ hours, power-law process with $\eta = 1000\,\text{h}$. The row $b = 1$ is the HPP with rate $10^{-3}\,\text{h}^{-1}$; the row $b = 1.4$, $t = 2000\,\text{h}$, $\tau = 500\,\text{h}$ is the worked example of Section 17.4.5.

Shape b	Current age t (h)	ΔW, $\tau = 500\,\text{h}$	$\Pr(\Delta N \geq 1)$, $\tau = 500\,\text{h}$	ΔW, $\tau = 1000\,\text{h}$	$\Pr(\Delta N \geq 1)$, $\tau = 1000\,\text{h}$
0.6	1000	0.275	0.241	0.516	0.403
0.6	2000	0.217	0.195	0.417	0.341
0.6	4000	0.168	0.155	0.329	0.280
1.0	any	0.500	0.393	1.000	0.632
1.4	1000	0.764	0.534	1.639	0.806
1.4	2000	0.968	0.620	2.017	0.867
1.4	4000	1.249	0.713	2.554	0.922
1.6	1000	0.913	0.599	2.031	0.869
1.6	2000	1.301	0.728	2.768	0.937
1.6	4000	1.906	0.851	3.943	0.981

The trend is the one the shape predicts. With $b = 1$ the HPP rows are identical at every age, 0.5 expected failures and a probability of 0.393 in 500 h, since only the length of the horizon matters. With $b = 1.4$ and $b = 1.6$ the expected number of failures in the next 500 h grows as the system ages: for $b = 1.6$ it doubles between 1000 and 4000 h, from 0.913 to 1.906, and the probability of at least one failure rises from 0.60 to 0.85. With $b = 0.6$ it falls, from 0.275 to 0.168, and an HPP would overstate the failures of the aging system. The scale matters as much as the shape: for $b = 1.4$, $t = 2000$ h and $\tau = 500$ h, the expected count is 2.554 with $\eta = 500$ h and 0.367 with $\eta = 2000$ h, against 1.0 and 0.25 for the HPP with the same scale. These calculations show why the failure-process model matters for maintenance planning: the same recent history extrapolated by an HPP and by a deteriorating NHPP gives different numbers of expected failures, spare parts and repair crews. The next chapter, Chapter 18, turns to what repair does to a system: renewal processes, perfect against minimal against imperfect repair with virtual age, and age-based replacement.

17.7 What to retain

Definition 17.1 (Repairable systems and counting processes)**.** A repairable system is described by its failure counting process $N(t)$; the expected cumulative number of failures $W(t) = \mathbb{E}[N(t)]$ and its slope, the rate of occurrence of failures $w(t)$, are properties of the failure process, distinct from the hazard of one lifetime. An HPP has a constant w and Poisson counts with mean λt; an NHPP lets w vary and gives Poisson counts with mean $W(t_b) - W(t_a)$, with the power-law process improving for $b < 1$ and deteriorating for $b > 1$; and what repair does to the system, and how long it takes, are separate modeling choices.

Table 17.10 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 17.10: The terms of repairable-system and counting-process analysis and where each is defined. Units assume time in hours.

Term	Expression	Units	Meaning	Defined in
Counting process	$N(t) = \sum_j \mathbf{1}\{T_j \le t\}$	–	failures in $(0, t]$	17.1.2
Failure epoch	T_j	h	time of the j-th failure	17.1.2
Interfailure time	$X_j = T_j - T_{j-1}$	h	time between failures	17.2.1
Operating-time clock	hour-meter reading	h	excludes downtime	17.2.2
Cumulative intensity	$W(t) = \mathbb{E}[N(t)]$	–	expected failures by t (MCF)	17.2.4
ROCOF	$w(t) = \mathrm{d}W/\mathrm{d}t$	h^{-1}	unconditional failure rate of the process	17.2.4
Conditional intensity	$\lambda^*(t \mid \mathcal{H}_t)$	h^{-1}	failure rate given the history	17.2.5
Empirical MCF	$\hat{W}(t) = \sum_{t_j \le t} d_j/n_j$	–	nonparametric estimate of W	17.2.6
HPP	$w = \lambda$, $N(t) \sim \mathrm{Poisson}(\lambda t)$	h^{-1}	constant ROCOF, exponential X_j	17.3.1
HPP rate estimate	$\hat{\lambda} = r/T_{\mathrm{tot}}$	h^{-1}	failures per unit exposure	17.3.4
NHPP	$\Delta N \sim \mathrm{Poisson}(W(t_b) - W(t_a))$	–	time-varying ROCOF	17.4.2
Power-law process	$W(t) = (t/\eta)^b$	–	$b < 1$ improving, $b > 1$ deteriorating	17.4.3
Virtual age	V_n	h	effective age after the n-th repair	17.5.1
Perfect repair	$V_n = 0$	–	as good as new; renewal process	17.5.2
Minimal repair	$V_n = T_n$, $w(t) = \lambda(t)$	–	as bad as old; NHPP	17.5.3
Imperfect repair	$0 < V_n < T_n$	–	in between; history-dependent	17.5.4
Steady-state availability	$\mathrm{MTTF}/(\mathrm{MTTF} + \mathrm{MTTR})$	–	needs repair durations	17.5.6
NHPP likelihood	$\mathrm{e}^{-W(T_{\mathrm{obs}})} \prod_j w(t_j)$	–	failure times and window	17.6.1
Power-law MLE	$\hat{b} = r/\sum_j \ln(T_{\mathrm{obs}}/t_j)$	–	closed-form shape	17.6.2
HPP trend test	$2\sum_j \ln(T_{\mathrm{obs}}/t_j) \sim \chi^2_{2r}$	–	tests $b = 1$	17.6.4

Further reading. The *NIST/SEMATECH e-Handbook of Statistical Methods* (the NIST Engineering Statistics Handbook): the homogeneous Poisson process, the NHPP power-law model, and reliability trend and growth plotting. Rausand, Barros and Høyland, *System Reliability Theory*, third edition: repairable systems, renewal processes and reliability modeling.

17.8 Exercises

Exercise 17.1. (a) What does $N(t)$ represent for a repairable system: the number of recorded failures by time t, the probability of surviving until t, or the number of completed repairs only? (The number of failures recorded by time t, the cumulative recurrent-failure count.) (b) What is the cumulative intensity: the hazard of a lifetime, $W(t) = \mathbb{E}[N(t)]$, or the mean repair time? ($W(t) = \mathbb{E}[N(t)]$, the expected number of failures by age t.) (c) Under an HPP, which distribution describes the number of failures over a fixed exposure interval: Poisson, always normal, or uniform? (Poisson, with mean equal to the rate times the exposure.) (d) For a power-law process with $W(t) = (t/\eta)^b$, what does $b > 1$ indicate? (An increasing rate of

occurrence of failures, since $w(t)$ is proportional to t^{b-1}.) (e) Does an NHPP necessarily model perfect repair? (No; the classical age-based NHPP corresponds to minimal repair, while perfect repair with nonexponential lifetimes generally gives a renewal process.) (f) Why does repair downtime matter when a recurrence rate is estimated? (It changes the time during which the equipment is exposed to failure; a rate estimate needs a consistent exposure clock, and calendar and operating time differ.)

Exercise 17.2. Four emergency generators are each observed for 2500 operating hours and have 10 failures in total; assume an HPP. (a) Estimate the rate and the mean operating time between failures, and compute the 95 % chi-square limits of the rate with $\chi^2_{20,0.025} = 9.591$ and $\chi^2_{22,0.975} = 36.781$. ($\hat{\lambda} = 10/10\,000 = 1.0 \times 10^{-3}\,\mathrm{h}^{-1}$; 1000 h; 4.80×10^{-4} to $1.84 \times 10^{-3}\,\mathrm{h}^{-1}$.) (b) For one generator, compute the expected number of failures in its next 1000 operating hours, the probability of none, and the probability of two or more. (1.0; $\mathrm{e}^{-1} = 0.368$; $1 - 2\mathrm{e}^{-1} = 0.264$.) (c) With a mean repair time of 20 h, compute the MTBF in this book's convention and the steady-state availability. (1020 h; $1000/1020 = 0.980$.) (d) If the 10 000 operating hours had been recorded as 10 400 calendar hours, downtime included, what rate would be computed, and why should it not be used in the failure-occurrence model? (9.62×10^{-4} per calendar hour, 3.8 % lower; it counts repair periods, in which the generators could not fail, as exposure.)

Exercise 17.3. A pump fleet follows a power-law process with $b = 1.6$ and $\eta = 2000$ h. (a) Compute $W(t)$ and $w(t)$ at 2000 and 4000 h. (1 and 3.031; 8.0×10^{-4} and $1.213 \times 10^{-3}\,\mathrm{h}^{-1}$.) (b) Compute the expected number of failures per pump in $(2000, 3000]$ and in $(4000, 5000]$. (0.913 and 1.301.) (c) For a pump of age 4000 h, compute the probability of no failure in the next 500 h. ($\mathrm{e}^{-(3.660-3.031)} = \mathrm{e}^{-0.629} = 0.533$, so 0.467 for at least one.) (d) An analyst fits instead the HPP with the same expected count at 4000 h. Give its rate and its probability of at least one failure in the same 500 h, and state the direction of its error. ($3.031/4000 = 7.58 \times 10^{-4}\,\mathrm{h}^{-1}$; $1 - \mathrm{e}^{-0.379} = 0.315$; it understates the risk of the aging pump by a third.)

Exercise 17.4. A single repairable system is observed to $T_{\mathrm{obs}} = 5000$ h and fails at 900, 1900, 2700, 3400, 4000, 4500 and 4800 h. (a) Compute $\sum_j \ln(T_{\mathrm{obs}}/t_j)$ and the power-law estimates $\hat{b}$ and $\hat{\eta}$. (4.054; $\hat{b} = 7/4.054 = 1.727$; $\hat{\eta} = 5000 \times 7^{-1/1.727} = 1620$ h.) (b) Compute the unbiased shape estimate and the 95 % interval of equation (17.26), with $\chi^2_{14,0.025} = 5.629$ and $\chi^2_{14,0.975} = 26.119$. ($6 \times 1.727/7 = 1.480$; 0.694 to 3.222.) (c) Test the HPP with the statistic $2\sum_j \ln(T_{\mathrm{obs}}/t_j)$ at the two-sided 5 % level. (8.107 lies between 5.629 and 26.119: the HPP is not rejected, although $\hat{b} = 1.73$.) (d) Compute the fitted ROCOF at 5000 h and the expected failures in the next 1000 h under the power-law fit and under the HPP fit, and state what the comparison implies for a maintenance decision. ($w(5000) = \hat{b}\,r/T_{\mathrm{obs}} = 2.42 \times 10^{-3}\,\mathrm{h}^{-1}$ against $1.4 \times 10^{-3}\,\mathrm{h}^{-1}$; $7[(6000/5000)^{1.727} - 1] = 2.59$ against 1.40; the decision depends on a trend that seven failures cannot establish, so the condition of the system at its repairs should be examined before acting on it.)

Chapter 18

Renewal Processes and Maintenance Models

18.1 The effect of repair on future failures

Chapter 17 described a repairable item by its counting process $N(t)$, the number of failures by time t, and by the cumulative intensity $W(t) = \mathbb{E}[N(t)]$ and its derivative, the rate of occurrence of failures (ROCOF) $w(t) = \mathrm{d}W/\mathrm{d}t$. It distinguished the homogeneous Poisson process (HPP), whose ROCOF is constant, from the nonhomogeneous Poisson process (NHPP), whose ROCOF changes with the age of the system. It left one engineering question open: when a failed item is repaired, its reliability may return to that of a new item, or the item may keep some of the deterioration it had accumulated. The answer changes the mathematical model, the predicted number of failures and the best maintenance policy. The question of this chapter is therefore

what a repair or a replacement does to the future failure behavior of an item, and when preventive replacement is economically justified.

The chapter answers it in five steps. It defines three idealized repair assumptions, perfect, minimal and imperfect repair, through the age that an item has after the repair. It develops the renewal process that perfect repair produces: the renewal function, the renewal equation and the elementary renewal theorem. It shows that minimal repair produces an NHPP whose ROCOF equals the hazard of a new item. It represents imperfect repair by a virtual age and the two Kijima models. Finally, it uses the renewal-reward theorem to derive the long-run cost rate of an age-based preventive replacement policy, optimizes the replacement age of a cooling-water pump, and adds repair downtime and availability.

Jargon: perfect, minimal and imperfect repair, renewal process, renewal function, renewal density, convolution, virtual age, restoration parameter, regeneration point, renewal-reward theorem, age replacement, long-run cost rate, mean down time

A *perfect repair* restores an item to the condition of a new one ("as good as new"); a *minimal repair* restores its function and leaves its condition as it was just before the failure ("as bad as old"); an *imperfect repair* leaves it in an intermediate condition. A *renewal process* is the counting process of failures of an item whose successive operating times are independent and identically distributed, which perfect repair produces. Its *renewal function* $M(t)$ is the expected number of renewals in $[0, t]$, dimensionless, and its *renewal density* $m(t) = \mathrm{d}M/\mathrm{d}t$, in h^{-1}, is the ROCOF of the renewal process. The *convolution* F^{*n} is the distribution of a sum of n independent lifetimes. The *virtual age* V_n, in hours, is the age of a new item that would have the same future failure behavior as the item has after its n-th repair. The *restoration parameter* q, between zero and one, states what fraction of the age is retained at a repair. A *regeneration point* is an instant at which a process restarts with the same probabilistic future, and a *renewal cycle* is the interval between two of them. The *renewal-reward theorem* gives the long-run average cost per unit time as the expected cost of a cycle divided by its expected length. *Age replacement* replaces an item at failure or when it reaches a preventive replacement age T_p, whichever comes first, at cost c_f or c_p; its *long-run cost rate* $C(T_p)$ is in money per operating hour. The *mean down time* of a replacement is the mean time the item is out of service for it.

Key idea

Repair effectiveness, not the lifetime distribution alone, fixes the future failure process: perfect repair produces a renewal process, minimal repair an NHPP whose ROCOF equals the hazard of a new item, and imperfect repair a process driven by a virtual age. The renewal-reward theorem turns one replacement cycle into a long-run cost rate, and preventive replacement lowers that rate only when the hazard increases with age and a failure costs more than a planned replacement.

This section states what repair means in terms of the age of an item and fixes the lifetime distribution of a new pump on which every later section builds.

18.1.1 What repair means mathematically

Consider a cooling-water pump whose hazard increases with age, because its bearings wear. At an operating age of 2000 h a bearing fails. Three different maintenance actions are possible, and they leave the pump in three different conditions. A replacement of the whole pump by a new unit returns it to age zero. A repair that restores the flow and changes nothing else, such as freeing a jammed shaft, leaves it at the age it had, 2000 h. A repair that replaces the failed bearing and leaves the motor, the seals, the housing and the piping as they were removes part of the accumulated wear, and the pump then behaves as a younger one.

These three outcomes are described by an *effective age*, the age that a new pump would need to have to face the same future hazard. Figure 18.1 shows the three cases. Each horizontal bar is the scale of effective age from 0 to 2000 h, and the dot marks the effective age immediately after the repair at 2000 h: zero for perfect repair, an illustrative 900 h for imperfect repair, and 2000 h for minimal repair. The text at the right of each bar gives the hazard at that effective age, computed from the lifetime model of the next subsection. The hazard after the repair ranges from zero to $1.0 \times 10^{-3}\,\mathrm{h}^{-1}$, so the three repairs promise very different futures for the same pump.

These are idealizations. Replacing a bearing does not reverse the age of the motor, the

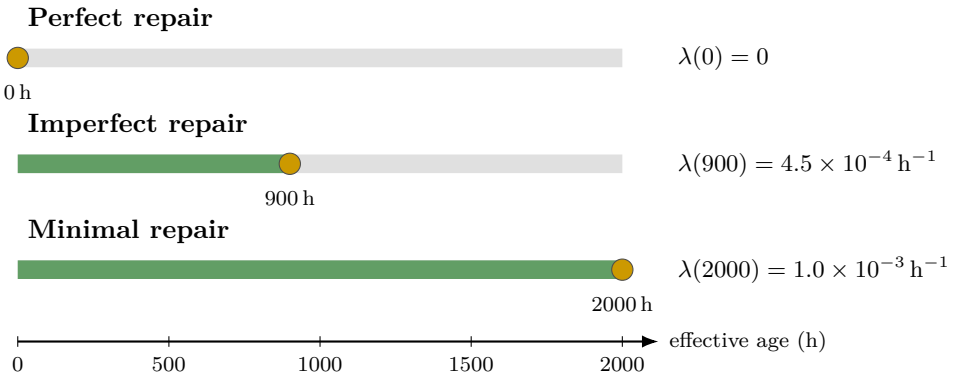


Figure 18.1: Effective age of a cooling-water pump immediately after a repair at an operating age of 2000 h, under three repair assumptions. The imperfect-repair value of 900 h is illustrative. The hazards are from the Weibull model of Section 18.1.2, $\lambda(t) = 2t/\eta^2$ with $\eta = 2000$ h.

seals or the housing, and the pump does not literally become younger. An effective-age model nevertheless describes how a maintenance action changes the probability of future failures, which is what a maintenance decision needs. The rest of the chapter makes the three cases precise: perfect repair in Section 18.2, minimal repair in Section 18.3 and imperfect repair in Section 18.4.

18.1.2 The baseline lifetime distribution

All three repair models start from the lifetime of a new item. Let T be the time to first failure of a new pump, in operating hours, with cumulative failure probability $F(t) = \Pr(T \le t)$, reliability function $R(t) = 1 - F(t)$, failure density $f(t) = \mathrm{d}F/\mathrm{d}t$, hazard $\lambda(t) = f(t)/R(t)$ and cumulative hazard $\Lambda(t) = \int_0^t \lambda(u)\,\mathrm{d}u = -\ln R(t)$, as in Chapter 5. This distribution is called the *baseline* lifetime distribution: it is the reference that the repair models modify.

For the running example the baseline lifetime is the Weibull distribution of Chapter 6,

$$R(t) = \exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right], \qquad \beta = 2, \quad \eta = 2000\,\mathrm{h}, \tag{18.1}$$

an illustrative model of a wearing pump. The characteristic life η is the age by which 63.2 % of new pumps have failed, not the mean. The mean follows from the Weibull mean of Chapter 6,

$$\mathrm{MTTF} = \eta\,\Gamma\!\left(1 + \frac{1}{\beta}\right) = 2000\,\Gamma(1.5) = 2000 \times \frac{\sqrt{\pi}}{2} = 1772.45\,\mathrm{h}. \tag{18.2}$$

The Weibull hazard is $\lambda(t) = (\beta/\eta)(t/\eta)^{\beta-1}$, and with $\beta = 2$ it is linear in age,

$$\lambda(t) = \frac{2t}{\eta^2} = 5 \times 10^{-7}\,t\,\mathrm{h}^{-1}, \tag{18.3}$$

with t in hours. Because $\beta > 1$, the hazard increases with age:

$$\lambda(1000) = 5 \times 10^{-7} \times 1000 = 5 \times 10^{-4}\,\mathrm{h}^{-1}, \qquad \lambda(2000) = 1.0 \times 10^{-3}\,\mathrm{h}^{-1}.$$

Under this model a pump that has reached 2000 h has twice the hazard of a pump aged 1000 h. The cumulative hazard is $\Lambda(t) = (t/\eta)^2$, which is 1 at $t = 2000$ h. The maintenance question can now be stated precisely: after a repair, at which age does the pump re-enter equation (18.3)?

18.2 Perfect repair and renewal processes

Perfect repair restarts the lifetime of the item at every failure. This section defines the renewal process that results, derives its renewal function and the renewal equation, compares the renewal function with the cumulative intensity of an NHPP, and states the elementary renewal theorem for the long-run failure frequency.

18.2.1 Perfect repair and the ordinary renewal process

A perfect repair restores the item to the same probabilistic condition as a new unit: after the repair, the time to the next failure has the baseline distribution, whatever happened before. Replacement of the failed item by a new identical one is the usual physical realization. Let $X_1, X_2, \ldots$ be the successive operating times, X_1 from the start to the first failure and X_n from the $(n-1)$-th restoration to the n-th failure. Under perfect repair they are assumed independent and identically distributed with the baseline distribution,

$$X_n \overset{\text{iid}}{\sim} F, \qquad n = 1, 2, \ldots. \tag{18.4}$$

Time is measured on the operating-time clock of Chapter 17, and the repair itself is taken as instantaneous; downtime is added in Section 18.6.4.

The n-th failure, or *renewal*, occurs at

$$S_n = X_1 + X_2 + \cdots + X_n, \qquad S_0 = 0, \tag{18.5}$$

and the number of renewals by time t is

$$N(t) = \max\{n : S_n \le t\} = \sum_{n=1}^{\infty} \mathbf{1}\{S_n \le t\}. \tag{18.6}$$

The counting process $N(t)$ defined by equations (18.4)–(18.6) is an *ordinary renewal process*; "ordinary" means that the first operating time has the same distribution as the others, because the item starts new. The word renewal refers to the probabilistic restart of the item after each perfect restoration: at every S_n the future of the process has the same distribution as at time zero.

18.2.2 The renewal function

The expected number of renewals in $[0, t]$ is the *renewal function*,

$$M(t) = \mathbb{E}[N(t)]. \tag{18.7}$$

It is the cumulative intensity $W(t)$ of Chapter 17 for the particular case of a renewal process; this chapter keeps the letter M for it so that the perfect-repair count is never confused with the NHPP mean function $W(t)$. Taking the expectation of the sum of indicators in equation (18.6), and using $\mathbb{E}[\mathbf{1}\{A\}] = \Pr(A)$,

$$M(t) = \sum_{n=1}^{\infty} \Pr(S_n \le t). \tag{18.8}$$

The time of the n-th renewal is a sum of n independent lifetimes, so its distribution is the n-fold *convolution* of F with itself, defined recursively by $F^{*1} = F$ and

$$F^{*n}(t) = \Pr(S_n \le t) = \int_0^t F^{*(n-1)}(t-u)\, f(u)\, \mathrm{d}u,$$

where the integral conditions on the first lifetime, $X_1 = u$, and asks that the remaining $n-1$ lifetimes fit into $t-u$. Substituting into equation (18.8),

$$M(t) = \sum_{n=1}^{\infty} F^{*n}(t). \tag{18.9}$$

This is the exact renewal function. Its derivative is the *renewal density* $m(t) = \mathrm{d}M/\mathrm{d}t$, in h^{-1}: $m(t)\,\mathrm{d}t$ is, to first order, the probability of a renewal in $(t, t+\mathrm{d}t]$, so $m(t)$ is the ROCOF of the renewal process.

Equation (18.9) is exact but rarely convenient. The convolutions have closed forms for the exponential distribution, whose n-fold convolution is the gamma (Erlang) distribution, and for a few others; for a Weibull lifetime with $\beta \neq 1$ none of them has a closed form, and the series must be evaluated numerically. A more useful characterization is an integral equation.

18.2.3 The renewal equation

The renewal equation is derived by conditioning on the time of the first renewal, X_1. Two cases are possible.

- If $X_1 = u > t$, no renewal has occurred by t, and $N(t) = 0$.
- If $X_1 = u \le t$, one renewal has occurred at u. Because the repair is perfect, the process restarts at u as a new, independent copy of itself, and the number of renewals in $(u, t]$ has the distribution of $N(t-u)$. Hence $\mathbb{E}[N(t) \mid X_1 = u] = 1 + M(t-u)$.

Averaging over the density $f(u)$ of the first lifetime,

$$M(t) = \int_0^t \left[1 + M(t-u)\right] f(u)\,\mathrm{d}u + \int_t^{\infty} 0 \cdot f(u)\,\mathrm{d}u,$$

and since $\int_0^t f(u)\,\mathrm{d}u = F(t)$,

$$M(t) = F(t) + \int_0^t M(t-u)\, f(u)\,\mathrm{d}u. \tag{18.10}$$

For a lifetime distribution without a density the integral is written $\int_0^t M(t-u)\,\mathrm{d}F(u)$. Equation (18.10) is the *renewal equation*. The first term, $F(t)$, is the expected contribution of the first failure; the integral accounts for every subsequent failure, which happens in a renewed copy of the process started at the first failure. Its recursive structure, with M on both sides, is the mathematical expression of the repeated resetting. Differentiating both sides with respect to t, and using $M(0) = 0$, gives the same equation for the renewal density,

$$m(t) = f(t) + \int_0^t m(t-u)\, f(u)\,\mathrm{d}u. \tag{18.11}$$

The renewal equation is solved numerically by discretizing time. On a grid $t_i = i\,\Delta t$ the integral is replaced by a sum over the increments $F(t_{k+1}) - F(t_k)$, and $M(t_i)$ is obtained from the values already computed at earlier grid points. For the pump of equation (18.1), a grid of $\Delta t = 1\,\mathrm{h}$ gives $M(1000) = 0.231$, $M(2000) = 0.754$ and $M(4000) = 1.894$; a Monte Carlo simulation of two million renewal histories agrees to the third decimal. These values are compared with minimal repair in Section 18.3.3.

18.2.4 The renewal function and the NHPP mean function

For a nonexponential lifetime the renewal function is not the cumulative hazard of the baseline distribution:

$$M(t) \neq \Lambda(t) = -\ln R(t) \quad \text{in general.} \tag{18.12}$$

For the pump, $\Lambda(2000) = 1$ while $M(2000) = 0.754$. The distinction matters because, as Section 18.3.2 shows, $\Lambda(t)$ is the expected number of failures under minimal repair, the mean function $W(t)$ of the power-law NHPP of Chapter 17, whereas $M(t)$ is the expected number of failures under perfect repair. The two describe different repair mechanisms applied to the same item.

The exponential lifetime is the one case in which the two coincide. With $F(t) = 1 - \mathrm{e}^{-\lambda t}$ and $f(u) = \lambda\,\mathrm{e}^{-\lambda u}$, the trial solution $M(t) = \lambda t$ satisfies the renewal equation. The integral in equation (18.10) is

$$\int_0^t \lambda(t-u)\,\lambda\,\mathrm{e}^{-\lambda u}\,\mathrm{d}u = \lambda t(1-\mathrm{e}^{-\lambda t}) - \left[1 - \mathrm{e}^{-\lambda t}(1+\lambda t)\right] = \lambda t - 1 + \mathrm{e}^{-\lambda t},$$

where the second term is $\lambda^2 \int_0^t u\,\mathrm{e}^{-\lambda u}\,\mathrm{d}u$, evaluated by parts. Adding $F(t) = 1 - \mathrm{e}^{-\lambda t}$ gives λt, as required, and the solution of the renewal equation is unique among functions bounded on finite intervals. The renewal process with exponential lifetimes is therefore the HPP, with $M(t) = \lambda t = \Lambda(t)$: a constant-hazard item gives the same expected count under every repair assumption, a point taken up in Section 18.3.4.

18.2.5 The elementary renewal theorem

Over a long horizon the renewal function grows linearly. Suppose the operating times have a finite mean, $\mathbb{E}[X] = \mathrm{MTTF}$. The *elementary renewal theorem* states that

$$\lim_{t\to\infty} \frac{M(t)}{t} = \frac{1}{\mathrm{MTTF}}. \tag{18.13}$$

It is a standard result of renewal theory. Its sample-path version shows why it holds. At time t the last renewal has occurred and the next has not, $S_{N(t)} \le t < S_{N(t)+1}$; dividing by $N(t)$,

$$\frac{S_{N(t)}}{N(t)} \le \frac{t}{N(t)} < \frac{S_{N(t)+1}}{N(t)+1} \cdot \frac{N(t)+1}{N(t)}.$$

As $t \to \infty$, $N(t) \to \infty$, and by the strong law of large numbers each average of operating times tends to MTTF; hence $t/N(t) \to \mathrm{MTTF}$, and $N(t)/t \to 1/\mathrm{MTTF}$ with probability one. The elementary renewal theorem is the corresponding statement for the expectation $M(t) = \mathbb{E}[N(t)]$, which requires a further argument and is taken here as standard.

The *long-run failure frequency* under perfect repair is therefore

$$f_{\mathrm{F}} = \frac{1}{\mathrm{MTTF}}. \tag{18.14}$$

For the pump, from equation (18.2),

$$f_{\mathrm{F}} = \frac{1}{1772.45\,\mathrm{h}} = 5.642 \times 10^{-4}\,\mathrm{h}^{-1},$$

or 0.564 failures per 1000 operating hours in the long run. This assumes instantaneous restoration on the operating-time clock. Downtime must be added before converting to failures per calendar year: with a mean down time of 24 h per failure, each cycle lasts on average $1772.45+24 = 1796.45$ calendar hours, and a pump in continuous service fails $8760/1796.45 = 4.88$ times per year rather than $8760/1772.45 = 4.94$.

A sharper statement is also standard: when the lifetime has a density and a finite variance σ^2,

$$M(t) \approx \frac{t}{\mathrm{MTTF}} + \frac{1}{2}\left(\frac{\sigma^2}{\mathrm{MTTF}^2} - 1\right) \qquad \text{for large } t. \tag{18.15}$$

For the Weibull pump, $\sigma^2 = \eta^2[\Gamma(2) - \Gamma(1.5)^2] = \eta^2(1 - \pi/4)$, so $\sigma = 926.5\,\text{h}$, $\sigma^2/\text{MTTF}^2 = 0.273$ and the constant is -0.363. A wearing item, whose lifetimes are less variable than exponential ones ($\sigma < \text{MTTF}$), lies below the line t/MTTF by a fixed amount; the renewal density $m(t)$ correspondingly rises from zero, overshoots slightly and settles at $1/\text{MTTF}$, as Figure 18.2 of the next section shows.

18.3 Minimal repair

Minimal repair is the opposite idealization: the item is returned to operation in the condition it had just before the failure. This section defines it, shows that it produces an NHPP whose ROCOF equals the hazard of a new item, compares it numerically with perfect repair, and explains why the exponential lifetime cannot tell the two apart.

18.3.1 Minimal repair: as bad as old

A minimal repair restores the function of the item without changing its effective age. If the item has accumulated an operating age t when it fails, its virtual age after the repair is

$$V(t) = t, \tag{18.16}$$

and the hazard that governs its next failure is the baseline hazard at that age, $\lambda(t)$, exactly as for a new item that had survived to t without failing. For the Weibull pump, $\lambda(t) = 2t/\eta^2$: a pump minimally repaired at 2000 h resumes with the hazard $1.0 \times 10^{-3}\,\text{h}^{-1}$ of a 2000 h pump, not with the zero hazard of a new one.

Minimal repair describes interventions that touch a small part of a large item: freeing a jammed shaft, resetting a tripped protection, replacing a fuse or a single gasket in an otherwise aged machine. The failed part is put right, but the wear state of everything else, which dominates the hazard, is unchanged.

18.3.2 Minimal repair produces an NHPP

Under minimal repair the probability of a failure in a short interval $(t, t + \mathrm{d}t]$, given the entire history of failures before t, is $\lambda(t)\,\mathrm{d}t$: whatever the number and times of earlier failures, the item at t is in the condition of a survivor of age t. A conditional failure intensity that depends on time but not on the history is the defining property of the NHPP of Chapter 17. Minimal repair of an item with a deterministic age-dependent hazard therefore produces an NHPP with ROCOF

$$w(t) = \lambda(t). \tag{18.17}$$

The ROCOF is a rate of a counting process and the hazard is a property of a lifetime distribution; equation (18.17) states that, under minimal repair, the first takes the numerical values of the second, not that they are the same object.

The expected number of failures in $[0, t]$ is the integral of the ROCOF,

$$W(t) = \int_0^t w(u)\,\mathrm{d}u = \int_0^t \lambda(u)\,\mathrm{d}u = \Lambda(t) = -\ln R(t). \tag{18.18}$$

Under minimal repair the expected number of failures is the cumulative hazard of the baseline lifetime. For the Weibull model,

$$W(t) = \left(\frac{t}{\eta}\right)^{\beta}, \tag{18.19}$$

which is the power-law NHPP of Chapter 17, with its shape equal to the Weibull shape β and its scale equal to the Weibull scale η. For the pump, $W(t) = (t/2000)^2$:

$$W(2000) = 1, \qquad W(4000) = 4, \qquad W(6000) = 9.$$

The expected cumulative count grows with the square of age, and the expected number of failures in successive 2000 h intervals is 1, 3 and 5. This is the consequence of minimal repair combined with an increasing hazard: every repair returns the pump to the rising part of its own hazard curve.

18.3.3 Minimal repair compared with perfect repair

For the same baseline lifetime, perfect repair resets the failure-age clock after every failure, and minimal repair leaves it running. The expected failure counts can therefore differ by a large factor. Figure 18.2 compares them for the pump. In both panels the horizontal axis is operating time in hours. In panel (a) the vertical axis is the expected number of failures: the solid curve is the renewal function $M(t)$ of perfect repair, computed from the renewal equation; the dashed curve is the minimal-repair count $W(t) = (t/2000)^2$; and the dotted line is the asymptote of equation (18.15). In panel (b) the vertical axis is the ROCOF in $10^{-4}\,\mathrm{h}^{-1}$: the renewal density $m(t)$, the minimal-repair ROCOF $w(t) = \lambda(t)$, and the horizontal line 1/MTTF.

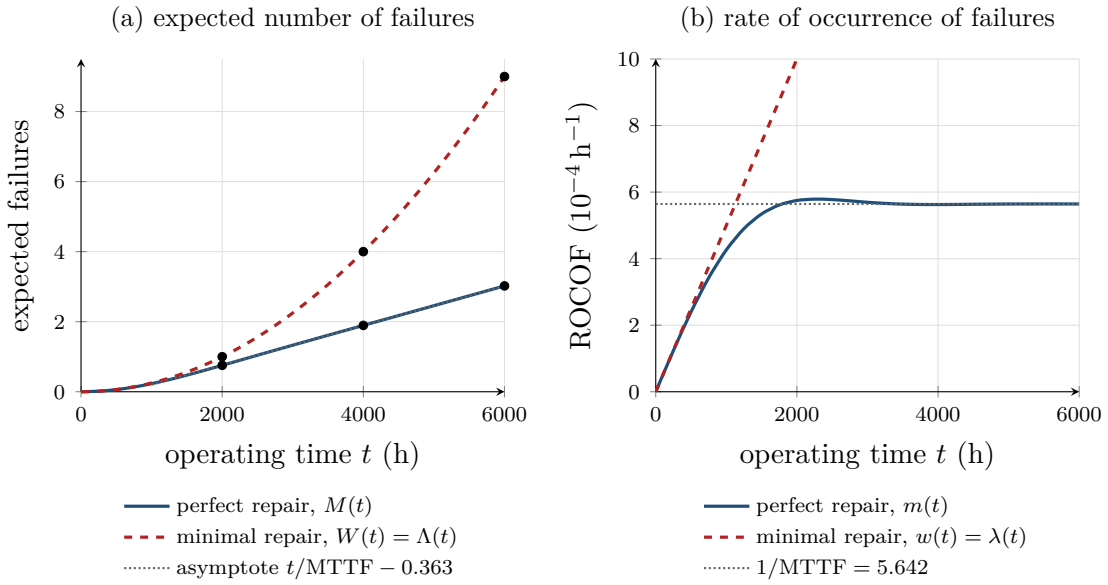


Figure 18.2: Perfect against minimal repair for the Weibull pump, $\beta = 2$, $\eta = 2000$ h, illustrative. (a) Expected number of failures: renewal function $M(t)$ from the renewal equation (18.10), solved numerically; minimal-repair count $W(t) = (t/\eta)^2$; dots at 2000, 4000 and 6000 h are the values of Table 18.1. (b) ROCOF: renewal density $m(t)$, which settles at 1/MTTF; minimal-repair ROCOF $w(t) = 2t/\eta^2$, which leaves the panel at $t = 2000$ h and keeps rising.

The two repair assumptions agree only at the start. Up to about 1000 h few pumps have failed even once, and the counts are close: 0.231 against 0.250. They then separate. Under perfect repair the count bends into the straight line of slope 1/MTTF, because every replacement brings a pump back to the low-hazard part of its life; the renewal density in panel (b) rises from zero, overshoots to a maximum of $5.79 \times 10^{-4}\,\mathrm{h}^{-1}$ near 2300 h and settles at $1/\mathrm{MTTF} = 5.64 \times 10^{-4}\,\mathrm{h}^{-1}$. Under minimal repair the count keeps curving upward, because the ROCOF $w(t) = 5 \times 10^{-7}\,t$ grows without limit. Table 18.1 lists the values. Its columns are the operating time, the perfect-repair count $M(t)$, the first-order approximation t/MTTF of equation (18.13), the minimal-repair count $W(t)$ and the ratio of the two counts. By 6000 h minimal repair produces three times as many expected failures as perfect repair, 9 against 3.02, and between 4000 and 6000 h it produces 5 against 1.13.

The comparison has a practical consequence. The power-law NHPP of Chapter 17 is the model of minimal repair; it is not a model of perfect replacement of a component with a

Table 18.1: Expected number of failures of the Weibull pump ($\beta = 2$, $\eta = 2000$ h) under perfect repair, $M(t)$, and under minimal repair, $W(t) = (t/\eta)^2$. $M(t)$ is the numerical solution of the renewal equation (18.10); MTTF = 1772.45 h.

Time t (h)	$M(t)$, perfect	t/MTTF	$W(t)$, minimal	$W(t)/M(t)$
1000	0.231	0.564	0.250	1.08
2000	0.754	1.128	1.000	1.33
3000	1.330	1.693	2.250	1.69
4000	1.894	2.257	4.000	2.11
6000	3.022	3.385	9.000	2.98

nonexponential Weibull lifetime. For a lifetime with increasing hazard, minimal repair never yields fewer expected failures than perfect repair, since a perfectly repaired item is always at least as young as a minimally repaired one, and the difference grows with time.

Common mistake

Using the power-law NHPP $W(t) = (t/\eta)^\beta$ to count the failures of a Weibull component that is replaced by a new one at every failure. That NHPP is the minimal-repair model, and it overstates the count of a replaced wearing component: for the pump it predicts 9 failures in 6000 h where perfect replacement gives 3.02. Under replacement the count is the renewal function $M(t)$, approximately t/MTTF over a long horizon.

18.3.4 The exponential lifetime as a special case

Suppose the baseline hazard is constant, $\lambda(t) = \lambda$, so that $R(t) = \mathrm{e}^{-\lambda t}$. Under perfect repair, independent exponential operating times produce an HPP with rate λ, as Section 18.2.4 verified from the renewal equation. Under minimal repair the ROCOF is $w(t) = \lambda(t) = \lambda$, constant, and the NHPP is again an HPP with rate λ. Therefore

$$M(t) = W(t) = \lambda t. \tag{18.20}$$

Not only the expected counts agree: both processes are the same HPP, so every probability of the failure times agrees. This is a consequence of the memoryless property of the exponential distribution (Chapter 6). A surviving item of any age has the same future as a new one, so it does not matter whether a repair returns it to age zero or leaves it at its current age.

The result does not mean that perfect and minimal repairs are physically equivalent. It means that an exponential failure model cannot distinguish their effects on subsequent failure times, and conversely that observing a constant ROCOF does not show that repairs are perfect. For aging equipment the effectiveness of repair matters, and it matters more the faster the hazard rises.

Principle 18.1. *The repair assumption determines the failure process: perfect repair of an item produces a renewal process with expected count $M(t)$, and minimal repair produces an NHPP with ROCOF equal to the baseline hazard and expected count $\Lambda(t)$. The two coincide only for a constant hazard, so a constant observed ROCOF says nothing about how effective the repairs are.*

18.4 Imperfect repair and virtual age

Most real maintenance lies between the two extremes. This section represents imperfect repair by a virtual age, defines the Kijima Type I and Type II models, works a comparison of the two for the pump, and discusses how a repair model is chosen.

18.4.1 Virtual age and the conditional ROCOF

A technician who replaces the failed bearing of a pump and leaves the other components unchanged removes some of the deterioration, not all of it. The *virtual age* formalizes the partial restoration. Let V_n be the virtual age immediately after the n-th repair, with $V_0 = 0$ for a new item, and let the virtual age grow at the rate of operating time between repairs. At an operating time t between the n-th repair, at S_n, and the next failure, the virtual age is

$$V(t) = V_n + (t - S_n), \qquad n = N(t). \tag{18.21}$$

The model assumes that the item fails as a new item of age $V(t)$ would. Its conditional ROCOF, the failure intensity given the history $\mathcal{H}_t$ of failures and repairs up to t, is

$$w_{\mathrm{C}}(t \mid \mathcal{H}_t) = \lambda(V(t)), \tag{18.22}$$

and, equivalently, the operating time to the next failure has the conditional reliability of a new item that has survived to age V_n,

$$\Pr(X_{n+1} > x \mid V_n = v) = \frac{R(v + x)}{R(v)}. \tag{18.23}$$

For the Weibull baseline the conditional ROCOF is

$$w_{\mathrm{C}}(t \mid \mathcal{H}_t) = \frac{\beta}{\eta}\left(\frac{V(t)}{\eta}\right)^{\beta-1},$$

and for an increasing hazard a repair that reduces the virtual age reduces the propensity to fail afterward. The intensity depends on the history through V_n, which depends on all earlier operating times, so the process is neither a renewal process nor an NHPP. Perfect repair is the case $V_n = 0$, and minimal repair the case $V_n = S_n$, for which $V(t) = t$. What remains to be specified is how a repair updates the virtual age.

18.4.2 Kijima Type I and Type II models

The two most common rules are the Kijima Type I and Type II models. Both use a *restoration parameter*

$$0 \le q \le 1, \tag{18.24}$$

which measures repair effectiveness: $q = 0$ is perfect repair, $q = 1$ is minimal repair, and $0 < q < 1$ is imperfect repair, with a smaller q meaning a more effective restoration. Some software reports its complement $1 - q$, the fraction of age a repair removes, as the restoration factor; the two conventions must not be mixed.

Remark 18.1. In this chapter q is the Kijima restoration parameter, a property of a repair. It is not a failure probability, and in particular not the basic-event probability q_i of the fault-tree chapters.

Kijima Type I. The repair removes the fraction $1 - q$ of the age accumulated during the most recent operating interval and none of the virtual age retained from earlier intervals:

$$V_n = V_{n-1} + q\,X_n. \tag{18.25}$$

Unrolling the recursion from $V_0 = 0$ gives $V_n = q\,(X_1 + \cdots + X_n) = q\,S_n$: the virtual age is a fixed fraction of the total operating age. Damage that a repair leaves in place stays there; a later repair addresses only what accumulated since the previous one.

Kijima Type II. The repair reduces the whole virtual age accumulated before the intervention, old and new:

$$V_n = q\,(V_{n-1} + X_n). \tag{18.26}$$

Unrolling gives $V_n = qX_n + q^2X_{n-1} + \cdots + q^nX_1$: each earlier interval is discounted by a further factor q at every subsequent repair. If the operating times are bounded by $x_{\max}$, the virtual age stays below the geometric-series bound $q\,x_{\max}/(1-q)$, so a Type II item settles into a steady pattern of failures, while a Type I item with $q > 0$ ages without limit, more slowly than under minimal repair.

Both models contain the two limiting cases: with $q = 0$ both give $V_n = 0$, perfect repair, and with $q = 1$ both give $V_n = S_n$, minimal repair. They differ only for intermediate restoration, and the difference grows with the number of repairs.

18.4.3 Worked example: three repairs of the pump

Suppose the pump of Section 18.1.2 fails after the successive operating intervals

$$X_1 = 1000\,\text{h}, \qquad X_2 = 800\,\text{h}, \qquad X_3 = 700\,\text{h},$$

and that each failure is followed by an imperfect repair with $q = 0.5$, starting from $V_0 = 0$. Repair durations are neglected, so the cumulative operating age is also the calendar age.

Kijima Type I. By equation (18.25),

$$V_1 = 0 + 0.5 \times 1000 = 500\,\text{h}, \quad V_2 = 500 + 0.5 \times 800 = 900\,\text{h}, \quad V_3 = 900 + 0.5 \times 700 = 1250\,\text{h},$$

which is 0.5 times the cumulative ages 1000, 1800 and 2500 h.

Kijima Type II. By equation (18.26),

$$V_1 = 0.5 \times 1000 = 500\,\text{h}, \quad V_2 = 0.5\,(500 + 800) = 650\,\text{h}, \quad V_3 = 0.5\,(650 + 700) = 675\,\text{h},$$

which is also $0.5 \times 700 + 0.25 \times 800 + 0.125 \times 1000 = 350 + 200 + 125 = 675$ h from the unrolled form.

Table 18.2 collects the results. Each row is a repair; the columns are the operating interval that ended in the failure, the cumulative operating age at the repair, and the virtual age after the repair under each model. Both models use exactly the same failure intervals and the same restoration parameter, and they agree after the first repair, where the past contains a single interval; afterward they diverge, and after the third repair the Type I virtual age is nearly twice the Type II value.

Table 18.2: Virtual age of the pump after each of three imperfect repairs with restoration parameter $q = 0.5$, illustrative operating intervals.

Repair n	interval X_n (h)	cumulative age S_n (h)	Type I V_n (h)	Type II V_n (h)
first	1000	1000	500	500
second	800	1800	900	650
third	700	2500	1250	675

Figure 18.3 draws the virtual age against the age of the pump. The horizontal axis is the cumulative operating age, equal to the calendar age when repairs take no time; the vertical axis is the virtual age. Between repairs every curve rises with slope one; at each repair it drops by

the amount the repair removes. The thin dotted diagonal is minimal repair, which never drops, and the thin gray sawtooth is perfect repair, which drops to zero. The two Kijima models coincide up to the second failure at 1800 h, where Type II drops lower, to 650 h against 900 h, because it also discounts the 500 h retained from the first interval. The segments after 2500 h continue the operation into the fourth interval. The figure shows the long-run difference as well: the post-repair values of the Type I curve lie on the line $V = q\,t = 0.5\,t$ and keep climbing, while those of the Type II curve stay in a band.

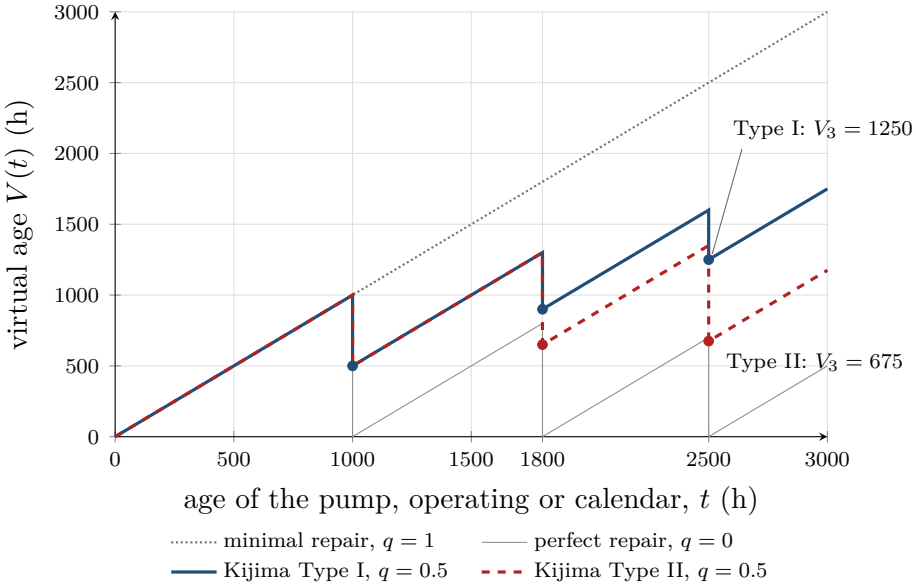


Figure 18.3: Virtual age against age for the pump with failures after 1000, 800 and 700 h of operation and instantaneous repairs: Kijima Type I, $V_n = V_{n-1} + qX_n$, and Type II, $V_n = q(V_{n-1} + X_n)$, with $q = 0.5$, against the limiting cases of minimal and perfect repair. Dots: virtual ages after the repairs, Table 18.2. Illustrative values.

The virtual ages translate into hazards through equation (18.22). For the Weibull pump with $\beta = 2$ the hazard is proportional to the virtual age, $\lambda(V) = 2V/2000^2$, so immediately after the third repair

$$\lambda_{\text{Type I}} = \frac{2 \times 1250}{2000^2} = 6.25 \times 10^{-4}\,\text{h}^{-1}, \qquad \lambda_{\text{Type II}} = \frac{2 \times 675}{2000^2} = 3.375 \times 10^{-4}\,\text{h}^{-1}.$$

Table 18.3 extends the comparison to the two limiting cases and adds, as an illustrative addition, the probability that the pump fails within the next 500 h of operation, from equation (18.23): $1 - R(V_3 + 500)/R(V_3)$. One row of the arithmetic, for Type I: the increase of the cumulative hazard over the interval is $(1750/2000)^2 - (1250/2000)^2 = 0.7656 - 0.3906 = 0.375$, and the failure probability is $1 - \text{e}^{-0.375} = 0.3127$. The four assumptions span a factor of eight in this probability, from 0.061 to 0.497, for the same pump with the same history.

The Type II model predicts a lower future hazard because it assumes that each repair reduces deterioration retained from earlier intervals as well as that of the most recent one. This is a model assumption, not evidence that Type II repairs are more effective in practice: the same data and the same number $q = 0.5$ mean different physical repairs in the two models.

18.4.4 Choosing a repair model

Table 18.4 summarizes the repair assumptions of this chapter. Each row is an assumption; the columns give the virtual age after the n-th repair, the failure process that results, how its

Table 18.3: The pump immediately after its third repair, at an operating age of 2500 h, under four repair assumptions: virtual age, hazard, and probability of failure within the next 500 h. Weibull baseline $\beta = 2$, $\eta = 2000$ h; Kijima models with $q = 0.5$.

Repair assumption	virtual age V_3 (h)	hazard $\lambda(V_3)$ (h^{-1})	failure in next 500 h
perfect, $q = 0$	0	0	0.0606
Kijima Type II, $q = 0.5$	675	3.375×10^{-4}	0.2065
Kijima Type I, $q = 0.5$	1250	6.25×10^{-4}	0.3127
minimal, $q = 1$	2500	1.25×10^{-3}	0.4972

expected number of failures is obtained, and the corresponding value of the Kijima restoration parameter.

Table 18.4: Repair assumptions, the virtual age each leaves, and the failure process each produces.

Assumption	virtual age V_n	failure process	expected failures	Kijima q
perfect (as good as new)	0	renewal process	$M(t)$, renewal equation (18.10)	0
minimal (as bad as old)	S_n	NHPP, $w(t) = \lambda(t)$	$W(t) = \Lambda(t)$	1
imperfect, Type I	$V_{n-1} + qX_n = qS_n$	history-dependent, $w_{\mathrm{C}} = \lambda(V(t))$	numerical evaluation or simulation	$0 < q < 1$
imperfect, Type II	$q(V_{n-1} + X_n)$	history-dependent, $w_{\mathrm{C}} = \lambda(V(t))$	numerical evaluation or simulation	$0 < q < 1$

The choice among the rows is an engineering judgment supported by data. The maintenance records and the physics of the repair should establish whether repairs mainly address the deterioration accumulated since the previous repair, which suggests Type I; whether they remove substantial preexisting deterioration, which suggests Type II; whether they replace entire assemblies that dominate the hazard, which suggests perfect repair; or whether they leave most of the previous wear unchanged, which suggests minimal repair. A single item can need different models for different repair actions: a bearing replacement may be close to Type II, a seal adjustment close to minimal.

Failure and repair histories can also be used to estimate q together with the baseline parameters, by maximizing a likelihood built from equation (18.23). Such estimates require care. The inferred restoration parameter can depend strongly on the assumed baseline distribution, a Weibull and a lognormal baseline fitted to the same history can imply different repair effectiveness, and, as Section 18.3.4 showed, with a nearly constant hazard the data carry almost no information about q at all. Goodness-of-fit tests that compare perfect, minimal and imperfect repair on the same history are the tool for checking the choice.

18.5 Renewal-reward theory and age replacement

The previous sections predicted failures for a given repair assumption. This section turns to maintenance decisions: it states the renewal-reward theorem, defines what counts as a renewal cycle, and derives the long-run cost rate of an age-based preventive replacement policy.

18.5.1 The renewal-reward theorem

A replacement policy must balance two costs. Preventive replacement at a scheduled age is usually cheaper and less disruptive than replacement after a failure, but replacing too often discards useful life. Comparing policies requires the long-run average cost per unit time, and the renewal-reward theorem computes it from a single cycle.

Consider a process that consists of successive cycles, the j-th of length Y_j and with cost K_j incurred during it, and suppose that the pairs $(Y_1, K_1), (Y_2, K_2), \ldots$ are independent and identically distributed, with finite expected length $\mathbb{E}[Y]$ and finite expected cost $\mathbb{E}[K]$; the length and the cost of one cycle may be dependent. Let $K(t)$ be the total cost accumulated up to time t. The *renewal-reward theorem* states that the long-run average cost rate is

$$\lim_{t\to\infty} \frac{K(t)}{t} = \frac{\mathbb{E}[K]}{\mathbb{E}[Y]}, \tag{18.27}$$

with probability one, and that the same limit holds for the expected cost, $\mathbb{E}[K(t)]/t$. When cost is in reais and time in hours, the rate is in reais per hour. The theorem is a standard result of renewal theory.

The justification follows the argument of Section 18.2.5. Let $N_c(t)$ be the number of cycles completed by t. If the cost of each cycle is incurred at its end, as a replacement cost is,

$$\frac{K(t)}{t} = \frac{1}{t}\sum_{j=1}^{N_c(t)} K_j = \underbrace{\frac{1}{N_c(t)}\sum_{j=1}^{N_c(t)} K_j}_{\to\ \mathbb{E}[K]} \cdot \underbrace{\frac{N_c(t)}{t}}_{\to\ 1/\mathbb{E}[Y]} .$$

The first factor is an average of independent, identically distributed cycle costs and tends to $\mathbb{E}[K]$ by the strong law of large numbers; the second tends to $1/\mathbb{E}[Y]$ by the renewal argument, since the cycle ends form a renewal process with mean interval $\mathbb{E}[Y]$. Over a long horizon the number of cycles grows in proportion to time, and the total cost grows in proportion to the number of cycles. Costs incurred during a cycle rather than at its end change the total by at most the cost of the cycle in progress, which is negligible relative to t.

The power of the theorem is that it is not necessary to follow every future failure and replacement. A description of one typical cycle, its expected cost and its expected length, gives the long-run performance. The same theorem with "reward" equal to uptime gives the long-run availability, as Section 18.6.4 shows.

18.5.2 What counts as a renewal cycle

A cycle must begin at a regeneration point, an instant at which the future of the process is probabilistically identical to its future at the start of every other cycle. Installation of a new item is such an instant. A perfect repair is one; a minimal or a Kijima imperfect repair is not, because the item's future after it depends on its age or history. Under a policy with minimal repairs the cycle must therefore run from one replacement by a new item to the next, with the minimal repairs as events inside it.

For an *age-replacement policy* each cycle begins with a new component and ends when either

1. the component fails and is replaced, or
2. it reaches the specified age T_p without failing and is replaced preventively.

After either replacement the next cycle starts with a new component. With T the lifetime of the component in the cycle, the cycle length is

$$Y = \min(T, T_p), \tag{18.28}$$

a random quantity, because a failure may occur before the planned age. The policy does not guarantee operation to T_p.

Figure 18.4 shows four cycles of such a policy on one time axis. The boxes are operating intervals of successive new components; a red line at the end of a box marks a corrective replacement after a failure, at cost c_f, and a blue line a preventive replacement at age T_p, at cost c_p. The braces give the length of each cycle. The first and fourth components fail before reaching T_p, so their cycles are shorter than T_p; the second and third survive and are replaced at exactly T_p.

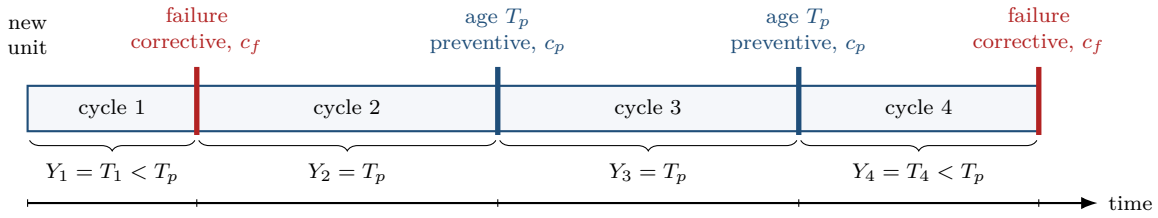


Figure 18.4: Four renewal cycles of an age-replacement policy. Each cycle starts with a new component and ends at its failure (corrective replacement, cost c_f) or at age T_p (preventive replacement, cost c_p), whichever comes first; the cycle length is $Y = \min(T, T_p)$. Illustrative lifetimes $T_1 = 430\,\text{h}$ and $T_4 = 610\,\text{h}$ with $T_p = 765\,\text{h}$; replacement times neglected.

18.5.3 The age-replacement policy

Let c_p be the cost of a preventive replacement and c_f the cost of a corrective replacement after a failure, both in reais, with

$$c_f > c_p. \tag{18.29}$$

The corrective cost is higher because a failure can involve emergency labor, secondary damage to equipment, production loss or interruption of service, in addition to the replacement itself. The policy is

replace the component when it fails, or preventively when it reaches age T_p, whichever happens first.

Two simplifications are made for now: both types of replacement restore the component to an as-good-as-new condition, so that every replacement is a regeneration point, and both take negligible time compared with the operating life. The downtime simplification is relaxed in Section 18.6.4. The decision variable is T_p, with $T_p = \infty$, no preventive replacement, as one of the candidates.

18.5.4 Expected cost and expected length of a cycle

Expected cost. A cycle ends in a preventive replacement if the component survives to T_p, which has probability $\Pr(T > T_p) = R(T_p)$, and in a corrective replacement if it fails before T_p, which has probability $\Pr(T \le T_p) = F(T_p)$. The expected cost per cycle is therefore

$$\mathbb{E}[K(T_p)] = c_p\, R(T_p) + c_f\, F(T_p), \tag{18.30}$$

or, since $R = 1 - F$,

$$\mathbb{E}[K(T_p)] = c_p + (c_f - c_p)\, F(T_p). \tag{18.31}$$

The second form reads as the preventive cost, paid in every cycle, plus the extra penalty of a failure, weighted by the probability that the failure comes first.

Expected length. For any nonnegative random variable Z, $\mathbb{E}[Z] = \int_0^\infty \Pr(Z > t)\,\mathrm{d}t$, the identity that gave MTTF $= \int_0^\infty R(t)\,\mathrm{d}t$ in Chapter 5. The cycle length $\min(T, T_p)$ exceeds t exactly when $T > t$ and $T_p > t$, so $\Pr(\min(T, T_p) > t)$ equals $R(t)$ for $t < T_p$ and zero for $t \geq T_p$. Hence

$$\mathbb{E}[Y(T_p)] = \mathbb{E}[\min(T, T_p)] = \int_0^{T_p} R(t)\,\mathrm{d}t. \tag{18.32}$$

This integral is the mean operating time per cycle; in survival analysis it is called the restricted mean survival time up to T_p. It increases with T_p and tends to MTTF as $T_p \to \infty$.

18.5.5 The long-run cost rate

The renewal-reward theorem (18.27), applied with equations (18.30) and (18.32), gives the long-run cost rate of the age-replacement policy,

$$C(T_p) = \frac{\mathbb{E}[K(T_p)]}{\mathbb{E}[Y(T_p)]} = \frac{c_p\,R(T_p) + c_f\,F(T_p)}{\displaystyle\int_0^{T_p} R(t)\,\mathrm{d}t}, \tag{18.33}$$

in reais per operating hour. This is the fundamental age-replacement cost formula under the assumptions of Section 18.5.3. The optimal replacement age is

$$T_p^* = \arg\min_{T_p > 0} C(T_p), \tag{18.34}$$

where the candidates include $T_p = \infty$, for which the policy reduces to corrective replacement only and $C(\infty) = c_f/\mathrm{MTTF}$, by equation (18.14): failures at the long-run frequency $1/\mathrm{MTTF}$, each costing c_f. The optimum need not be at a finite age.

The two ends of the formula show the trade-off. As $T_p \to 0$ the numerator tends to c_p and the denominator to zero, so $C(T_p)$ grows without bound: replacing almost new components wastes nearly all of their life. As $T_p \to \infty$ the cost tends to c_f/MTTF. Between the two, a minimum below c_f/MTTF exists when failures become likely enough with age that avoiding them is worth the discarded life, a condition made precise in Section 18.6.2.

> **Principle 18.2.** *By the renewal-reward theorem the long-run cost rate of a maintenance policy is the expected cost of one renewal cycle divided by its expected length. For age replacement it is* $C(T_p) = [c_p R(T_p) + c_f F(T_p)]/\int_0^{T_p} R(t)\,\mathrm{d}t$*, and a cycle must begin at a regeneration point.*

18.6 Optimal age replacement of the cooling-water pump

This section applies the cost rate to the pump: it evaluates the cost at several replacement ages, derives and interprets the optimality condition, shows when preventive replacement is not worthwhile, adds downtime and availability, tabulates the optimum for several cost ratios and shapes, and sets out a practical workflow.

18.6.1 Worked example: the cost rate of the pump

Consider the pump of equation (18.1), $R(t) = \exp[-(t/2000)^2]$, with illustrative costs

$$c_p = \mathrm{R\$}\,500, \qquad c_f = \mathrm{R\$}\,4000.$$

The corrective cost includes both the replacement expense and the expected operational consequences of an unscheduled failure. The cost rate (18.33) is

$$C(T_p) = \frac{500\,\mathrm{e}^{-(T_p/2000)^2} + 4000\,\left[1 - \mathrm{e}^{-(T_p/2000)^2}\right]}{\displaystyle\int_0^{T_p} \mathrm{e}^{-(t/2000)^2}\,\mathrm{d}t}. \tag{18.35}$$

The denominator. For $\beta = 2$ the survival integral has a closed form through the error function, $\operatorname{erf}(z) = (2/\sqrt{\pi}) \int_0^z \mathrm{e}^{-u^2}\,\mathrm{d}u$. Substituting $u = t/\eta$, $\mathrm{d}t = \eta\,\mathrm{d}u$,

$$\int_0^{T_p} \mathrm{e}^{-(t/\eta)^2}\,\mathrm{d}t = \eta \int_0^{T_p/\eta} \mathrm{e}^{-u^2}\,\mathrm{d}u = \frac{\eta\sqrt{\pi}}{2}\,\operatorname{erf}\!\left(\frac{T_p}{\eta}\right),$$

and with $\eta = 2000\,\mathrm{h}$

$$\mathbb{E}[Y(T_p)] = 1000\sqrt{\pi}\,\operatorname{erf}\!\left(\frac{T_p}{2000}\right) = 1772.45\,\operatorname{erf}\!\left(\frac{T_p}{2000}\right)\ \mathrm{h}. \tag{18.36}$$

Since $\operatorname{erf}(z) \to 1$ as $z \to \infty$, the expected cycle length tends to the MTTF of equation (18.2), as it must. For other shapes the integral is evaluated numerically.

Comparison of replacement ages. Table 18.5 evaluates the cost rate at seven replacement ages and for no preventive replacement. Its columns are the replacement age, the probability that the failure comes first, $F(T_p)$, the expected cycle length from equation (18.36), the expected cycle cost from equation (18.31), and the cost rate. One row of the arithmetic, at $T_p = 750\,\mathrm{h}$: $F(750) = 1 - \mathrm{e}^{-0.140625} = 0.1312$; $\operatorname{erf}(0.375) = 0.4041$, so $\mathbb{E}[Y] = 1772.45 \times 0.4041 = 716.3\,\mathrm{h}$; $\mathbb{E}[K] = 500 + 3500 \times 0.1312 = \mathrm{R\$}\,959.15$; and $C(750) = 959.15/716.3 = \mathrm{R\$}\,1.339$ per operating hour.

Table 18.5: Long-run cost rate of age replacement for the pump, Weibull $\beta = 2$, $\eta = 2000\,\mathrm{h}$, $c_p = \mathrm{R\$}\,500$, $c_f = \mathrm{R\$}\,4000$, illustrative. The row 765 h is the optimum.

Replacement age T_p	$F(T_p)$	$\mathbb{E}[Y]$ (h)	$\mathbb{E}[K]$ (R$)	$C(T_p)$ (R$/h)
500 h	6.06 %	489.8	712.05	1.454
750 h	13.12 %	716.3	959.15	1.339
765 h	13.61 %	729.3	976.37	1.339
1000 h	22.12 %	922.6	1274.20	1.381
1250 h	32.34 %	1104.7	1631.78	1.477
1500 h	43.02 %	1260.5	2005.76	1.591
2000 h	63.21 %	1493.6	2712.42	1.816
never ($T_p = \infty$)	100 %	1772.5	4000.00	2.257

Figure 18.5 draws the cost rate as a continuous function of the replacement age. The horizontal axis is T_p in hours and the vertical axis is the cost rate in reais per operating hour. The solid curve is $C(T_p)$ from equation (18.35); the dots are the rows of Table 18.5; the dashed horizontal line is the cost without preventive replacement, $c_f/\mathrm{MTTF} = \mathrm{R\$}\,2.257$ per hour, which the curve approaches from below at large T_p; and the dash-dotted line is $(c_f - c_p)\,\lambda(T_p)$, whose role is explained in Section 18.6.2. The curve falls steeply from very short replacement ages, reaches a minimum and rises again.

Numerical minimization of equation (18.35) gives

$$T_p^* = 765\,\mathrm{h}, \qquad C(T_p^*) = \mathrm{R\$}\,1.339 \text{ per operating hour},$$

against R$ 2.257 per hour without preventive replacement, a reduction of 41 %. At the optimum 13.6 % of the cycles still end in a failure. The minimum is flat: the cost stays within 1 % of its minimum for replacement ages from about 660 to 890 h, with $C(650) = \mathrm{R\$}\,1.355$ and $C(900) = \mathrm{R\$}\,1.355$ per hour almost equal on the two sides of the optimum, so that a replacement scheduled at 750 h costs R$ 1.339 per hour as well, and practical constraints on scheduling cost little.

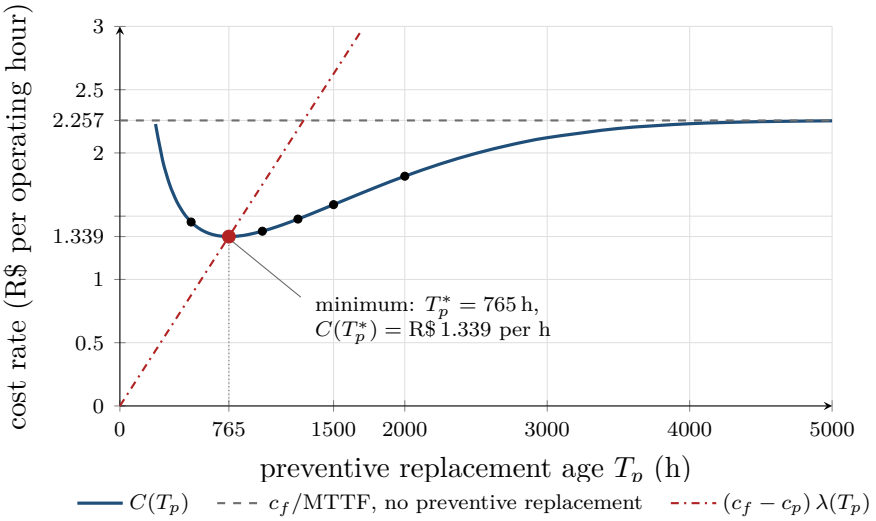


Figure 18.5: Long-run cost rate of age replacement for the pump, $C(T_p) = [500\,R(T_p) + 4000\,F(T_p)]/\int_0^{T_p} R(t)\,\mathrm{d}t$ with $R(t) = \exp[-(t/2000)^2]$, against the replacement age; illustrative costs and parameters. Dots: Table 18.5. Dashed: cost without preventive replacement. Dash-dotted: marginal failure penalty $(c_f - c_p)\lambda(T_p) = 1.75 \times 10^{-3}\,T_p$, which crosses $C(T_p)$ at its minimum.

The optimum is much earlier than the mean lifetime of 1772 h. The reason is the ratio of the costs, $c_f/c_p = 4000/500 = 8$: an unplanned failure costs eight times a planned replacement, and the policy accepts frequent preventive replacements, discarding on average $1772 - 729 = 1043$ h of potential life per cycle, to avoid the more expensive outcome. The result does not mean that every industrial pump should be replaced at 765 h. It is specific to the assumed Weibull parameters, to the assumption that every replacement is perfect, to the costs, and to the system boundary within which those costs were counted.

18.6.2 The optimality condition

Write the cost rate as $C(T) = k(T)/y(T)$, with numerator and denominator

$$k(T) = c_p + (c_f - c_p)\,F(T), \qquad y(T) = \int_0^T R(t)\,\mathrm{d}t,$$

where T stands for T_p to lighten the notation. Their derivatives are $k'(T) = (c_f - c_p)\,f(T)$ and $y'(T) = R(T)$, the latter by the fundamental theorem of calculus. By the quotient rule,

$$C'(T) = \frac{k'(T)\,y(T) - k(T)\,y'(T)}{y(T)^2} = \frac{(c_f - c_p)\,f(T)\,y(T) - k(T)\,R(T)}{y(T)^2}.$$

Since $f(T) = \lambda(T)\,R(T)$, the factor $R(T)$ can be taken out,

$$C'(T) = \frac{R(T)}{y(T)^2}\,\varphi(T), \qquad \varphi(T) = (c_f - c_p)\,\lambda(T)\,y(T) - k(T), \tag{18.37}$$

and the sign of C' is the sign of φ. At an interior optimum $C'(T_p^*) = 0$, so $\varphi(T_p^*) = 0$; dividing by $y(T_p^*)$ and recognizing $k/y = C$,

$$(c_f - c_p)\,\lambda(T_p^*) = C(T_p^*). \tag{18.38}$$

This is the interior optimality condition of age replacement.

Interpretation. Extending the replacement age of a surviving component from T_p to $T_p + \mathrm{d}t$ has two effects. It exposes the component to failure for a further $\mathrm{d}t$, which happens with probability $\lambda(T_p)\,\mathrm{d}t$ and turns a planned replacement into a corrective one, at an extra cost $c_f - c_p$: the left side of equation (18.38) is this marginal expected penalty per hour. It also adds $\mathrm{d}t$ of service, which at the long-run average rate is worth $C(T_p)\,\mathrm{d}t$: the right side. At the optimum the marginal expected failure penalty of keeping the component equals the average cost rate of the policy. Before T_p^* the penalty is smaller and it pays to keep the component; after T_p^* it is larger and it pays to replace it. The condition also gives the minimum cost directly from the hazard: for the pump, $(c_f - c_p)\,\lambda(765) = 3500 \times 5 \times 10^{-7} \times 765 = \mathrm{R\$}\,1.339$ per hour, the value of Table 18.5. In Figure 18.5 the dash-dotted line $(c_f - c_p)\lambda(T_p)$ crosses the cost curve exactly at its minimum.

Existence and uniqueness. The condition alone guarantees neither a unique optimum nor a finite one; the behavior of φ and of C at the boundaries settles both. Differentiating φ, and using $\lambda R = f$ again,

$$\varphi'(T) = (c_f - c_p)\big[\lambda'(T)\,y(T) + \lambda(T)\,R(T)\big] - (c_f - c_p)\,f(T) = (c_f - c_p)\,\lambda'(T)\,y(T).$$

For a strictly increasing hazard, φ is therefore increasing, starting from $\varphi(0) = -c_p < 0$. The cost rate decreases while $\varphi < 0$ and increases once $\varphi > 0$, so it has at most one interior minimum, and it has one exactly when φ becomes positive, that is, when

$$\lim_{T\to\infty} (c_f - c_p)\,\lambda(T)\,\mathrm{MTTF} > c_f, \tag{18.39}$$

since $y(T) \to \mathrm{MTTF}$ and $k(T) \to c_f$. For a Weibull lifetime with $\beta > 1$ the hazard grows without bound, the condition always holds, and a finite optimum exists for every cost ratio above one, although for a mild shape and a small ratio it may lie far out and save little (Table 18.7). For a hazard that levels off, the optimum is finite only if the limiting hazard is high enough compared with $c_f/[(c_f - c_p)\mathrm{MTTF}]$. For a constant hazard, $\lambda' = 0$, φ stays at $-c_p$, and the cost rate decreases at every age, the case of the next subsection.

18.6.3 When preventive replacement is not worthwhile

Suppose the component lifetime is exponential, $R(t) = \mathrm{e}^{-\lambda t}$, $F(t) = 1 - \mathrm{e}^{-\lambda t}$. The denominator of the cost rate is

$$\int_0^{T_p} \mathrm{e}^{-\lambda t}\,\mathrm{d}t = \frac{1 - \mathrm{e}^{-\lambda T_p}}{\lambda},$$

and substituting into equation (18.33),

$$C(T_p) = \frac{c_p\,\mathrm{e}^{-\lambda T_p} + c_f\,(1 - \mathrm{e}^{-\lambda T_p})}{(1 - \mathrm{e}^{-\lambda T_p})/\lambda}.$$

Dividing each term of the numerator by the denominator,

$$C(T_p) = \lambda\,c_f + \frac{\lambda\,c_p\,\mathrm{e}^{-\lambda T_p}}{1 - \mathrm{e}^{-\lambda T_p}}. \tag{18.40}$$

For $c_p > 0$ the second term is positive, and it decreases as T_p increases, because $\mathrm{e}^{-\lambda T_p}/(1 - \mathrm{e}^{-\lambda T_p}) = 1/(\mathrm{e}^{\lambda T_p} - 1)$ is decreasing. Therefore

$$\lim_{T_p\to\infty} C(T_p) = \lambda\,c_f, \qquad T_p^* = \infty. \tag{18.41}$$

In this model the optimal age-replacement policy is not to replace preventively at all. The reason is the memoryless property: replacing a working exponential component by a new identical

one does not change its future hazard, so the scheduled replacement adds the cost c_p without reducing the probability of failure. As a numerical check, an exponential component with the same MTTF as the pump, $\lambda = 1/1772.45\,\mathrm{h}^{-1} = 5.642 \times 10^{-4}\,\mathrm{h}^{-1}$, and the same costs has $\lambda c_f = \mathrm{R\$}\,2.257$ per hour; replacing it at 765 h adds $\lambda c_p/(\mathrm{e}^{0.4316} - 1) = \mathrm{R\$}\,0.523$ per hour, a total of R$ 2.779, 23 % more than never replacing it.

Principle 18.3. *Age-based preventive replacement lowers the long-run cost only when the hazard increases with age and a failure costs more than a planned replacement. A component with a constant hazard should not be replaced on a schedule solely to reduce its failure hazard.*

The conclusion rests on assumptions that should be checked before it is applied: a constant hazard; replacement with an identical component; no systematic upgrade during the replacement; no change in failure consequences with age; no regulatory or inspection requirement; and no other benefit from the preventive intervention. Condition monitoring, inspection, software updates, lubrication or other interventions may still be worthwhile when the baseline failure process is approximately exponential, because they act on something other than the age of the item. The result applies specifically to the age-replacement policy and the cost model derived here.

18.6.4 Repair downtime and availability

So far the replacements were instantaneous. In a real cooling plant a preventive replacement may take several hours and an emergency replacement much longer. Let d_p and d_f be the mean down times of a preventive and of a corrective replacement, in hours, and keep the assumption that every replacement restores the component to an as-good-as-new condition. A cycle now consists of an operating period followed by a down period. The expected down time per cycle is

$$\bar{d}(T_p) = d_p\,R(T_p) + d_f\,F(T_p), \tag{18.42}$$

the expected operating time per cycle is $\mathbb{E}[Y(T_p)] = \int_0^{T_p} R(t)\,\mathrm{d}t$ as before, and the expected calendar length of a cycle is $\mathbb{E}[Y(T_p)] + \bar{d}(T_p)$. Applying the renewal-reward theorem with the operating time of the cycle as the "reward" gives the long-run availability,

$$A(T_p) = \frac{\mathbb{E}[Y(T_p)]}{\mathbb{E}[Y(T_p)] + \bar{d}(T_p)}, \tag{18.43}$$

which is exact under the regenerative-cycle assumptions and finite mean durations. It is the steady-state availability of Chapter 3, with the cycle of the policy in place of a single failure and repair.

Cost per calendar hour against cost per operating hour. The cost rate (18.33) is per operating hour. Suppose c_p and c_f are direct maintenance costs that exclude the consequences of the outage. The average maintenance expenditure per calendar hour is then

$$C_{\mathrm{cal}}(T_p) = \frac{c_p\,R(T_p) + c_f\,F(T_p)}{\mathbb{E}[Y(T_p)] + \bar{d}(T_p)}. \tag{18.44}$$

If each hour of unavailability also costs c_d, in reais per hour, for lost production or unserved load, the expected cost per cycle gains the term $c_d\,\bar{d}(T_p)$ and

$$C_{\mathrm{tot}}(T_p) = \frac{c_p\,R(T_p) + c_f\,F(T_p) + c_d\,\bar{d}(T_p)}{\mathbb{E}[Y(T_p)] + \bar{d}(T_p)}. \tag{18.45}$$

The appropriate objective depends on what is to be minimized: cost per calendar hour, cost per operating hour, annual expected expense, unavailability, lost production, or life-cycle cost and net present value. These objectives need not share the same optimal policy. Downtime costs must also not be counted twice: if the consequence of an outage is already in c_f, as it was in the worked example, it must not be added again through c_d.

Common mistake

Adding a downtime cost $c_d\,\bar{d}(T_p)$ to a corrective cost c_f that already includes the production lost during the outage. Each consequence belongs in exactly one term of the cycle cost.

Availability example. Let $d_p = 4\,\text{h}$ and $d_f = 24\,\text{h}$, illustrative. At the cost-optimal age $T_p = 765\,\text{h}$, $F(765) = 0.136$ and $\mathbb{E}[Y(765)] = 729.3\,\text{h}$. The expected down time per cycle is

$$\bar{d} = 4 \times 0.864 + 24 \times 0.136 = 3.456 + 3.264 = 6.72\,\text{h},$$

and the availability is

$$A(765) = \frac{729.3}{729.3 + 6.72} = 0.99087 \approx 99.09\,\%.$$

Table 18.6 extends the example to other replacement ages, as an illustrative addition. Its columns are the replacement age, the expected operating and down time per cycle, the availability, the cost per operating hour $C(T_p)$, the cost per calendar hour $C_{\text{cal}}(T_p)$, and the total cost rate $C_{\text{tot}}(T_p)$ with an illustrative $c_d = \text{R\$}\,100$ per down hour, for direct costs $c_p = \text{R\$}\,500$ and $c_f = \text{R\$}\,4000$ taken as excluding the outage. Each objective has its own optimum: the cost per operating hour is lowest at 765 h and the cost per calendar hour at 764 h, nearly the same age, because the down time is small compared with the cycle; the availability is highest, 99.099 %, at about 910 h, because preventive down time also counts against it; and the total cost with $c_d = \text{R\$}\,100$ is lowest, R\$ 2.235 per hour, at about 819 h. The replacement policy affects availability through both the frequency of failures and the duration of restoration.

Table 18.6: Down time, availability and three cost rates of age replacement for the pump, with $d_p = 4\,\text{h}$, $d_f = 24\,\text{h}$, $c_p = \text{R\$}\,500$, $c_f = \text{R\$}\,4000$ and, in the last column, $c_d = \text{R\$}\,100$ per down hour; illustrative. Cost rates in R\$/h.

T_p (h)	$\mathbb{E}[Y]$ (h)	$\bar{d}$ (h)	$A(T_p)$	$C(T_p)$	$C_{\text{cal}}(T_p)$	$C_{\text{tot}}(T_p)$
500	489.8	5.21	98.947 %	1.454	1.439	2.491
765	729.3	6.72	99.087 %	1.339	1.327	2.240
910	850.6	7.74	99.099 %	1.357	1.345	2.246
1000	922.6	8.42	99.095 %	1.381	1.369	2.273
1500	1260.5	12.60	99.010 %	1.591	1.576	2.566
2000	1493.6	16.64	98.898 %	1.816	1.796	2.898
never	1772.5	24.00	98.664 %	2.257	2.227	3.563

18.6.5 Cost rate and optimal age for several cost ratios and shapes

The worked example fixes $\beta = 2$ and $c_f/c_p = 8$. How the optimum depends on these two inputs is seen by evaluating the model over a grid. The model is equation (18.33) with the Weibull reliability $R(t) = \exp[-(t/\eta)^{\beta}]$, $\eta = 2000\,\text{h}$ and $c_p = \text{R\$}\,500$, a new component being installed at every replacement and the expected cycle length evaluated by numerical integration. Table 18.7 gives, for shapes 1, 1.5, 2 and 3 and corrective costs of 2, 4, 8 and 16 times the preventive cost, the optimal age T_p^*, the minimum cost rate $C(T_p^*)$, the cost rate at a fixed age

of 750 h, the cost rate without preventive replacement $C(\infty) = c_f/\text{MTTF}$, and the saving of the optimum over no preventive replacement. One row of the arithmetic, for $\beta = 2$ and $c_f/c_p = 4$, that is c_f = R\$ 2000, at the optimum $T_p^* = 1188$ h: $F(1188) = 0.2973$, $\mathbb{E}[Y] = 1061.9$ h, $\mathbb{E}[K] = 500 + 1500 \times 0.2973 =$ R\$ 945.96 and $C = 945.96/1061.9 =$ R\$ 0.891 per hour; the optimality condition gives the same number, $1500 \times 5 \times 10^{-7} \times 1188 = 0.891$.

Table 18.7: Optimal replacement age and cost rates of age replacement for Weibull lifetimes with $\eta = 2000$ h and $c_p =$ R\$ 500, for four shapes and four cost ratios; illustrative. Cost rates in R\$ per operating hour. The row $\beta = 2$, $c_f/c_p = 8$ is the pump of Table 18.5.

β	c_f/c_p	T_p^* (h)	$C(T_p^*)$	$C(750)$	$C(\infty) = c_f/\text{MTTF}$	saving
1	2	∞	0.500	1.050	0.500	0
1	4	∞	1.000	1.550	1.000	0
1	8	∞	2.000	2.550	2.000	0
1	16	∞	4.000	4.550	4.000	0
1.5	2	4349	0.553	0.879	0.554	0.2 %
1.5	4	1678	1.030	1.178	1.108	7.0 %
1.5	8	902	1.763	1.776	2.215	20.4 %
1.5	16	532	2.900	2.973	4.431	34.6 %
2	2	2182	0.545	0.790	0.564	3.3 %
2	4	1188	0.891	0.973	1.128	21.1 %
2	8	765	1.339	1.339	2.257	40.7 %
2	16	519	1.947	2.072	4.514	56.9 %
3	2	1621	0.493	0.710	0.560	12.0 %
3	4	1108	0.691	0.780	1.120	38.3 %
3	8	832	0.909	0.918	2.240	59.4 %
3	16	645	1.169	1.196	4.479	73.9 %

Three trends appear. First, with $\beta = 1$, the exponential case, preventive replacement never pays: the optimum is $T_p^* = \infty$, and replacing at 750 h adds R\$ 0.55 per hour whatever the cost ratio, the second term of equation (18.40). Reducing the shape from 2 to 1 therefore makes frequent preventive replacement unattractive, because the hazard no longer increases with age. Second, for every shape above one, a larger cost ratio shortens the optimal age and raises the saving: for $\beta = 2$ the optimum falls from 2182 h at ratio 2 to 519 h at ratio 16, and the saving rises from 3.3 % to 56.9 %. Third, a steeper hazard makes preventive replacement more valuable at a given ratio, with savings at ratio 8 of 20 %, 41 % and 59 % for shapes 1.5, 2 and 3. At $\beta = 1.5$ and ratio 2 a finite optimum exists, as equation (18.39) guarantees, but it lies at 4349 h, more than twice the scale, and saves 0.2 %: in practice there is no case for preventive replacement there.

Figure 18.6 shows the cost curves behind the rows with $c_f/c_p = 8$. The horizontal axis is the replacement age and the vertical axis the cost rate; each curve is one shape, and the dots mark the minima of Table 18.7. The exponential curve, $C(T_p) = 2 + 0.25/(\mathrm{e}^{T_p/2000} - 1)$ from equation (18.40), decreases at every age toward its limit of R\$ 2 per hour; the three wear-out curves dip to a minimum that is deeper and better defined the larger the shape.

18.6.6 A practical workflow

For a real maintenance study of a pump, a chiller, an inverter or a transformer, the methods of this chapter combine into a defensible analysis in eight steps.

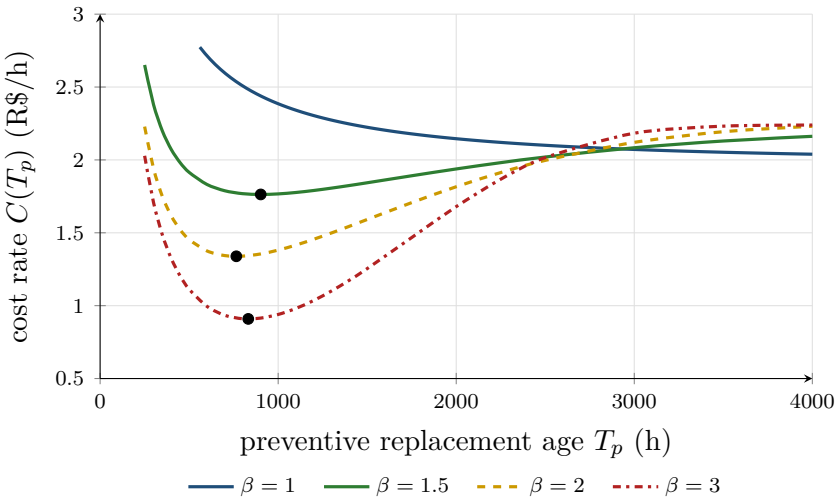


Figure 18.6: Long-run cost rate of age replacement against the replacement age for Weibull lifetimes with $\eta = 2000$ h and shapes 1, 1.5, 2 and 3, with $c_p =$ R$\,500 and $c_f =$ R$\,4000; illustrative. The $\beta = 1$ curve is $2 + 0.25/(\mathrm{e}^{T_p/2000} - 1)$, drawn from $T_p = 560$ h; the others are equation (18.33) evaluated numerically. Dots: minima at 902, 765 and 832 h. Without preventive replacement the costs are 2.000, 2.215, 2.257 and 2.240 per hour.

1. **Define the maintained item.** State whether the maintenance action restores an entire asset, a subsystem or an individual component.
2. **Establish the failure mechanism.** Determine whether the failure behavior is driven by age, cycles, load, contamination or another physical stress (Chapter 4).
3. **Collect recurrent failure and maintenance histories.** Preserve timestamps, exposure on a consistent clock, repair actions and post-repair observations.
4. **Evaluate repair effectiveness.** Compare the perfect, minimal and imperfect restoration assumptions against the evidence.
5. **Select a stochastic model.** Use a renewal process, an NHPP or a virtual-age model, as the evidence and the physics of the repair indicate.
6. **Define consequences and costs.** Include maintenance expense, outage duration, operational disruption and any other material impact, each counted once.
7. **Optimize the maintenance policy.** Evaluate the preventive replacement age, an inspection interval, a condition threshold or a replacement decision.
8. **Test sensitivity and uncertainty.** Determine whether the preferred policy is robust to uncertainty in the model parameters and the costs.

The order matters because the recurrence model that fits the failure history best does not by itself identify the best maintenance policy. A policy changes the future failure process of the system, so it must be evaluated with a model of how the proposed maintenance action affects reliability: an age-replacement optimum computed under perfect repair is wrong if the replacements are in fact imperfect, and a precise replacement interval is not defensible without evidence for the lifetime distribution, the repair effectiveness and the costs. The flatness of the optimum in Figure 18.5 helps here, since moderate errors in T_p^* cost little, but an error in the shape, which decides between the rows of Table 18.7 with and without a saving, does not.

The models of this chapter follow one item through its cycles. The next chapter represents the states of repairable systems with several units explicitly, as continuous-time Markov models (Chapter 19).

18.7 What to retain

Definition 18.1 (Repair and maintenance models)**.** What a repair does to the age of an item determines its future failure process: perfect repair produces a renewal process, minimal repair an NHPP whose ROCOF is the baseline hazard, and imperfect repair a virtual-age process. By the renewal-reward theorem the long-run cost of age replacement is $C(T_p) = [c_p R(T_p) + c_f F(T_p)] / \int_0^{T_p} R(t)\,\mathrm{d}t$, whose optimum satisfies $(c_f - c_p)\lambda(T_p^*) = C(T_p^*)$ and is finite only when the hazard rises enough with age.

Table 18.8 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 18.8: The terms of renewal processes, repair models and age replacement, and where each is defined. Units assume time in operating hours and money in reais.

Term	Expression	Units	Meaning	Defined in
Perfect repair	$X_n \overset{\text{iid}}{\sim} F$, $V_n = 0$	–	as good as new	18.2.1
Renewal process	$N(t) = \max\{n : S_n \le t\}$	–	failures under perfect repair	18.2.1
Renewal function	$M(t) = \sum_n F^{*n}(t)$	–	expected renewals in $[0, t]$	18.2.2
Renewal density	$m(t) = \mathrm{d}M/\mathrm{d}t$	h^{-1}	ROCOF of a renewal process	18.2.2
Renewal equation	$M = F + \int_0^t M(t-u) f(u)\,\mathrm{d}u$	–	condition on first renewal	18.2.3
Elementary renewal theorem	$M(t)/t \to 1/\mathrm{MTTF}$	h^{-1}	long-run failure frequency f_F	18.2.5
Minimal repair	$V(t) = t$	–	as bad as old	18.3.1
Minimal-repair NHPP	$w(t) = \lambda(t)$, $W(t) = \Lambda(t)$	h^{-1}, –	ROCOF equals baseline hazard	18.3.2
Virtual age	$V(t) = V_n + (t - S_n)$	h	effective age	18.4.1
Conditional ROCOF	$w_\mathrm{C}(t \mid \mathcal{H}_t) = \lambda(V(t))$	h^{-1}	intensity given history	18.4.1
Kijima Type I	$V_n = V_{n-1} + q X_n$	h	last interval reduced	18.4.2
Kijima Type II	$V_n = q(V_{n-1} + X_n)$	h	whole age reduced	18.4.2
Restoration parameter	$q \in [0, 1]$	–	0 perfect, 1 minimal	18.4.2
Renewal-reward theorem	$K(t)/t \to \mathbb{E}[K]/\mathbb{E}[Y]$	R\$/h	cost rate from one cycle	18.5.1
Regeneration point	–	–	probabilistic restart	18.5.2
Age replacement	$Y = \min(T, T_p)$	h	replace at failure or T_p	18.5.2
Long-run cost rate	$C(T_p) = \mathbb{E}[K]/\mathbb{E}[Y]$	R\$/h	policy objective	18.5.5
Optimality condition	$(c_f - c_p)\lambda(T_p^*) = C(T_p^*)$	R\$/h	marginal penalty = average rate	18.6.2
Availability with downtime	$\mathbb{E}[Y]/(\mathbb{E}[Y] + \bar{d})$	–	long-run up fraction	18.6.4

Further reading. Rausand, Barros and Høyland, *System Reliability Theory*, third edition: renewal theory, repairable systems and maintenance strategies. Modarres, Kaminskiy and Krivtsov, *Reliability Engineering and Risk Analysis*: repairable systems and maintenance optimization. *Imperfect Maintenance Models, from Theory to Practice*: Kijima virtual-age formulations and imperfect-repair modeling. *Imperfect Repair Models: Sequential Goodness-of-Fit Testing*: distinguishing perfect, minimal and imperfect repair processes.

18.8 Exercises

Exercise 18.1. (a) State what perfect repair assumes about the successive operating times. (They are independent and identically distributed with the lifetime distribution of a new component, which makes the failure process an ordinary renewal process.) (b) State which failure process is classically associated with minimal repair of an item with a deterministic age-dependent hazard. (A nonhomogeneous Poisson process, whose ROCOF follows the baseline hazard, since minimal repair preserves the effective age.) (c) State what $q = 0$ represents in a Kijima model. (Perfect restoration: both types reduce the virtual age to zero; q is a restoration parameter, not a failure probability.) (d) State what the renewal-reward theorem provides. (The long-run average cost or reward per unit time, as the expected cycle cost divided by the expected cycle length.) (e) State why age replacement is unattractive for an exponential lifetime. (A new identical component has the same constant hazard as the old one, so replacing a working item lowers no hazard and adds the cost c_p.) (f) State why repair downtime and replacement costs need separate treatment. (They affect availability and the economic objective differently: downtime lengthens the calendar cycle and reduces availability, while cost is the economic penalty of the cycle.)

Exercise 18.2. A pump has a Weibull lifetime with $\beta = 2$ and $\eta = 3000\,\mathrm{h}$. (a) Compute the MTTF and the long-run failure frequency under perfect repair. ($3000\,\Gamma(1.5) = 2658.7\,\mathrm{h}$; $f_\mathrm{F} = 3.761 \times 10^{-4}\,\mathrm{h}^{-1}$, or 0.376 per 1000 h.) (b) Compute the expected number of failures in the first 6000 h under minimal repair, and under perfect repair from Table 18.1, noting that for a fixed shape $M(t)$ depends only on t/η; check the second with equation (18.15). ($W(6000) = (6000/3000)^2 = 4$; $M(6000) = 1.894$, the tabulated value at $t/\eta = 2$; the asymptote gives $6000/2658.7 - 0.363 = 1.893$.) (c) The pump fails after operating intervals of 1200, 900 and 600 h, each followed by an imperfect repair with $q = 0.4$. Compute the virtual ages under Kijima Type I and Type II. (Type I: 480, 840, 1080 h; Type II: 480, 552, 460.8 h. The Type II virtual age falls at the third repair, because $0.4\,(552 + 600) < 552$.) (d) Compute the hazard immediately after the third repair under each model and under minimal repair, and the probability of failure in the next 500 h. (Type I: $2 \times 1080/3000^2 = 2.40 \times 10^{-4}\,\mathrm{h}^{-1}$ and 0.137; Type II: $1.02 \times 10^{-4}\,\mathrm{h}^{-1}$ and 0.076; minimal, $V = 2700\,\mathrm{h}$: $6.0 \times 10^{-4}\,\mathrm{h}^{-1}$ and 0.279.)

Exercise 18.3. The pump of Section 18.6.1 ($\beta = 2$, $\eta = 2000\,\mathrm{h}$) has $c_p = \mathrm{R\$}\,500$ and a lower corrective cost, $c_f = \mathrm{R\$}\,2000$. (a) Compute $C(1000)$, $C(1500)$ and $C(\infty)$, using $F(1000) = 0.2212$, $\mathbb{E}[Y(1000)] = 922.6\,\mathrm{h}$, $F(1500) = 0.4302$ and $\mathbb{E}[Y(1500)] = 1260.5\,\mathrm{h}$. (831.8/922.6 = R\$ 0.902; 1145.3/1260.5 = R\$ 0.909; 2000/1772.45 = R\$ 1.128 per hour.) (b) Show that $T_p = 1188\,\mathrm{h}$ satisfies the optimality condition (18.38), given $F(1188) = 0.2973$ and $\mathbb{E}[Y(1188)] = 1061.9\,\mathrm{h}$. ($(c_f - c_p)\lambda(1188) = 1500 \times 5 \times 10^{-7} \times 1188 = 0.891$ and $C(1188) = 945.96/1061.9 = 0.891$.) (c) Compute the saving of the optimum over no preventive replacement, and the cost of keeping the replacement age of 765 h found for $c_f = \mathrm{R\$}\,4000$. ($1 - 0.891/1.128 = 21\,\%$; $C(765) = (500 + 1500 \times 0.1361)/729.3 = \mathrm{R\$}\,0.966$ per hour, 8 % above the optimum.)

Exercise 18.4. (a) A component has an exponential lifetime with the same MTTF as the pump, 1772.45 h, and the costs $c_p = \mathrm{R\$}\,500$ and $c_f = \mathrm{R\$}\,4000$. Compute $C(765)$ and $C(\infty)$ from equation (18.40) and state the optimal policy. ($\lambda T_p = 0.4316$; $C(765) = 2.257 + 0.523 = \mathrm{R\$}\,2.779$ per hour; $C(\infty) = \lambda c_f = \mathrm{R\$}\,2.257$; never replace preventively.) (b) For the Weibull pump replaced at $T_p = 1000\,\mathrm{h}$, with $d_p = 4\,\mathrm{h}$ and $d_f = 24\,\mathrm{h}$, compute the expected down time per cycle, the availability and the cost per calendar hour, using $F(1000) = 0.2212$ and $\mathbb{E}[Y(1000)] = 922.6\,\mathrm{h}$. ($\bar{d} = 4 \times 0.7788 + 24 \times 0.2212 = 8.42\,\mathrm{h}$; $A = 922.6/931.0 = 99.10\,\%$; $C_\mathrm{cal} = 1274.2/931.0 = \mathrm{R\$}\,1.369$ per hour.) (c) With a downtime cost $c_d = \mathrm{R\$}\,100$ per hour and direct costs excluding the outage, compare C_tot at 765 and 1000 h, given $\bar{d}(765) = 6.72\,\mathrm{h}$. ($C_\mathrm{tot}(765) = (976.37 + 672.2)/736.0 = \mathrm{R\$}\,2.240$; $C_\mathrm{tot}(1000) = (1274.2 + 842.4)/931.0 = \mathrm{R\$}\,2.273$; 765 h is the better of the two, and the optimum of C_tot is about 819 h.)

Chapter 19

Markov Models and State Transitions

19.1 From failure counting to system states

Chapter 17 described the recurrent failures of a repairable item by counting processes, and Chapter 18 examined how the assumption made about each repair shapes the future failure behavior. Both chapters followed the sequence of failure events in time. A redundant system is characterized, in addition, by its condition at each instant: a cooling system with two redundant pumps can have both pumps operating, one operating, or neither operating, and it moves between these conditions as pumps fail and are repaired. The question of this chapter is

how the probability that a repairable system is operating, degraded or failed at any time is calculated, accounting for the failures and the repairs of its components.

Markov models answer it. They describe the probabilities of occupying the different states of a system and how those probabilities evolve over time, and they are the standard tool for the availability analysis of redundant systems, of systems with limited repair resources and of systems with several operating configurations. The chapter states the Markov property and the continuous-time Markov chain, solves the single repairable component in this framework, writes the general state equations in matrix form, solves a system of two repairable pumps in transient and in steady state, derives its failure frequency and mean down time, and ends with extensions, applications and checks of validity.

Jargon: state, Markov property, CTMC, transition rate, generator matrix, Kolmogorov forward equations, transient availability, stationary distribution, failure frequency, mean down time, absorbing state

A *state* is one element of a set of mutually exclusive and exhaustive conditions in which the model places the system, such as "two pumps operating". A process has the *Markov property* when its future evolution, given its present state, does not depend on the path by which it reached that state. A *continuous-time Markov chain* (CTMC) is a process with a discrete set of states, transitions at any instant of continuous time and the Markov property. A *transition rate* q_{ij}, in h^{-1}, is the instantaneous rate at which the process moves from state i to state j while it is in i. The *generator matrix* (transition-rate matrix) $\mathbf{Q}$ collects these rates. The *Kolmogorov forward equations* are the differential equations that the state probabilities $P_i(t)$ satisfy. *Transient availability* is the instantaneous availability $A(t)$ computed from them, as a function of time and of the initial state. The *stationary distribution* $\boldsymbol{\pi}$ is the set of state probabilities that the equations leave unchanged. The *failure frequency* f_{F}, in h^{-1} or yr^{-1}, is the long-run rate at which the system enters its failed states, and the *mean down time* (MDT), in hours, is the mean duration of one visit to them. An *absorbing state* is a state that the process cannot leave.

Key idea

A Markov model represents a repairable system by a finite set of states and constant transition rates between them; the Kolmogorov forward equations turn the rates into state probabilities as functions of time, from which availability, failure frequency and mean down time all follow.

This section introduces the example used throughout the chapter, two pumps in parallel, defines its states and names the quantities that the rest of the chapter computes.

19.1.1 Two pumps in parallel and their aggregate states

Two identical pumps, A and B, are installed in parallel in the cooling-water system. Either pump alone can supply the required cooling-water flow, so the required function is met while at least one pump operates. Each pump fails while it operates and is repaired after it fails. The condition of the system is summarized by the number of operational pumps at time t,

$$X(t) \in \{0, 1, 2\}. \tag{19.1}$$

Table 19.1 lists the three states. In state 2 both pumps operate and the system is available and fully redundant; in state 1 one pump operates and the system is available but degraded, since one further failure stops it; in state 0 neither pump operates and the system is unavailable.

Table 19.1: States of the system of two parallel pumps, either of which can supply the required flow.

State	Operational pumps	System condition
2	2	available, fully redundant
1	1	available, degraded (no remaining redundancy)
0	0	unavailable

The state is an *aggregate* state: it records how many pumps operate, not which ones. That is sufficient when the two pumps are identical, because the future of the system then depends only on the count. Two pumps with different failure or repair rates require the four states (A up, B up), (A up, B down), (A down, B up) and (A down, B down). The choice illustrates the

distinction between the physical system, the model and the evidence. The physical pumps have many attributes, such as bearing wear, temperature and vibration; the model keeps only the number operating; and the evidence is the operating log, which records the times of failures and of completed repairs and from which the rates of the model are estimated.

19.1.2 State probabilities, availability and unavailability

The central quantities are the state probabilities

$$P_2(t) = \Pr\big(X(t) = 2\big), \qquad P_1(t) = \Pr\big(X(t) = 1\big), \qquad P_0(t) = \Pr\big(X(t) = 0\big). \tag{19.2}$$

Because the three states are mutually exclusive and exhaustive, at every time

$$P_2(t) + P_1(t) + P_0(t) = 1. \tag{19.3}$$

The system is available whenever at least one pump operates, so its instantaneous availability is

$$A(t) = P_2(t) + P_1(t), \tag{19.4}$$

and its instantaneous unavailability is

$$1 - A(t) = P_0(t). \tag{19.5}$$

These are point quantities in the sense of Chapter 3: $A(t)$ is the probability that the system is available at the instant t, allowing any number of failures and restorations before it. The objective of the chapter is to compute the three state probabilities as functions of time, and their limits as $t \to \infty$.

Remark 19.1. The unavailability is written $1 - A(t)$, and its steady-state value $1 - A$, throughout this chapter. The letter Q is reserved for the generator matrix $\mathbf{Q}$ of Section 19.3.5, which is a matrix of transition rates and not the top-event probability Q_{top} or the system failure probability Q_{sys} of the fault-tree chapters.

19.2 The Markov property, continuous-time Markov chains and transition rates

The state probabilities can be computed with ordinary differential equations only when the future of the system depends on its present state alone. This section states that property, defines the continuous-time Markov chain and its transition rates, and fixes the failure and repair rates of one pump.

19.2.1 The Markov property

Let $\mathcal{H}_t$ denote the history of the process before time t, the states it occupied at all earlier times. A stochastic process $X(t)$ has the Markov property when, for every t, every $\Delta t > 0$ and every pair of states i and j,

$$\Pr\big(X(t + \Delta t) = j \mid X(t) = i,\ \mathcal{H}_t\big) = \Pr\big(X(t + \Delta t) = j \mid X(t) = i\big). \tag{19.6}$$

Given the present state, the future is conditionally independent of the earlier history. When the right-hand side depends on Δt but not on t, the process is *time-homogeneous*; all the models of this chapter are.

The property does not assert that the history is physically irrelevant. It asserts that the chosen state contains all the information needed to predict the future *under the model*. If

the failure probability of a pump depends on its accumulated wear, the statement that it is operational does not provide enough information, and its effective age, temperature, accumulated operating hours or another relevant variable must be added to the state. The same physical system can therefore be Markovian under one state description and not under another.

Principle 19.1. *The Markov property is a property of the state representation, not of the physical system alone. A model is Markovian when its state contains every variable on which the future transition rates depend.*

19.2.2 Continuous-time Markov chains and transition rates

A continuous-time Markov chain has a discrete state space, but its transitions can occur at any instant of continuous time. For a small interval Δt and $j \neq i$,

$$\Pr(X(t+\Delta t) = j \mid X(t) = i) = q_{ij}\,\Delta t + o(\Delta t), \tag{19.7}$$

where $o(\Delta t)$ stands for terms that vanish faster than Δt, $o(\Delta t)/\Delta t \to 0$. The constant $q_{ij} \geq 0$ is the transition rate from state i to state j. Its units are inverse time, here h^{-1}. For example, $q_{21} = 0.002\,\mathrm{h}^{-1}$ means that while the system occupies state 2, its instantaneous rate of transition to state 1 is 0.002 per hour. It does not mean that exactly 0.2 % of systems in state 2 move to state 1 in each hour: a rate is not a probability, and a finite-time transition probability follows only from the state equations, which account for the competing transitions and for the returns. The probability of leaving state 2 within one hour is $1 - \mathrm{e}^{-0.002} = 0.001998$, close to 0.002 only because the interval is short compared with the mean time spent in the state.

The total departure rate from state i is

$$\nu_i = \sum_{j \neq i} q_{ij}. \tag{19.8}$$

Let $G_i(s)$ be the probability that the process, having entered state i, is still there s hours later. Over a further short interval it stays with probability $1 - \nu_i\,\Delta s + o(\Delta s)$, so $G_i(s + \Delta s) = G_i(s)\,[1 - \nu_i\,\Delta s] + o(\Delta s)$, and in the limit $\mathrm{d}G_i/\mathrm{d}s = -\nu_i\,G_i$ with $G_i(0) = 1$. The holding time (sojourn time) in state i is therefore exponential, $G_i(s) = \mathrm{e}^{-\nu_i s}$, with mean

$$\mathbb{E}[\text{holding time in } i] = \frac{1}{\nu_i}, \qquad \nu_i > 0. \tag{19.9}$$

When the process leaves i, it goes to j with probability q_{ij}/ν_i, independently of the holding time; this is the standard description of a CTMC as a sequence of exponential holding times and jumps. In state 2 of the pump system, with $\nu_2 = q_{21} = 0.002\,\mathrm{h}^{-1}$, the mean holding time is 500 h. Exponential holding times are a defining feature of a time-homogeneous CTMC; a model whose holding times are not exponential is a semi-Markov model, a more general class. A state with $\nu_i = 0$ is absorbing: once entered, it is never left.

19.2.3 Failure and repair transition rates

For one pump, the failure rate while it operates is λ, and the repair rate, the rate at which an active repair is completed, is μ. With exponential operating and repair durations,

$$\mathrm{MTTF} = \frac{1}{\lambda}, \qquad \mathrm{MTTR} = \frac{1}{\mu}, \tag{19.10}$$

where the MTTF is the mean operating time until failure and the MTTR is the mean repair downtime, the full restoration time of Chapter 3. As an illustrative example,

$$\lambda = 0.001\,\mathrm{h}^{-1}, \qquad \mu = 0.05\,\mathrm{h}^{-1},$$

so that

$$\mathrm{MTTF} = \frac{1}{0.001\,\mathrm{h}^{-1}} = 1000\,\mathrm{h}, \qquad \mathrm{MTTR} = \frac{1}{0.05\,\mathrm{h}^{-1}} = 20\,\mathrm{h}.$$

The pump alternates between operating and failed periods with exponential up and down durations. Both rates are estimated from the operating log in the same way as a constant failure rate: the estimate of λ is the number of failures divided by the total operating time, and the estimate of μ is the number of completed repairs divided by the total time spent under repair. The following sections first derive the availability of one pump in the Markov framework and then turn to the redundant pair.

19.3 The single repairable component and the Kolmogorov equations

The single repairable pump is the smallest Markov model, and its equations can be solved by hand. This section writes its state equations, solves them for the transient and the steady-state availability, compares availability with reliability, and then restates the same model in the matrix form that generalizes to any number of states.

19.3.1 The two-state model and its forward equations

A single repairable pump has two states: U, operational, and D, down for repair. Figure 3.2 of Chapter 3 draws the model: the pump moves from U to D at rate λ and from D to U at rate μ. Let

$$P_{\mathrm{U}}(t) = \Pr\bigl(X(t) = \mathrm{U}\bigr), \qquad P_{\mathrm{D}}(t) = \Pr\bigl(X(t) = \mathrm{D}\bigr), \qquad P_{\mathrm{U}}(t) + P_{\mathrm{D}}(t) = 1.$$

Over a short interval Δt, the pump is up at $t + \Delta t$ either because it was up at t and did not fail, or because it was down at t and its repair was completed:

$$P_{\mathrm{U}}(t + \Delta t) = P_{\mathrm{U}}(t)\,(1 - \lambda\,\Delta t) + P_{\mathrm{D}}(t)\,\mu\,\Delta t + o(\Delta t). \tag{19.11}$$

The first term represents equipment that remains operational; the second represents equipment that completes its repair and returns to operation. Subtracting $P_{\mathrm{U}}(t)$, dividing by Δt and letting $\Delta t \to 0$,

$$\frac{\mathrm{d}P_{\mathrm{U}}}{\mathrm{d}t} = -\lambda\,P_{\mathrm{U}} + \mu\,P_{\mathrm{D}}, \tag{19.12}$$

and by the same argument for the down state,

$$\frac{\mathrm{d}P_{\mathrm{D}}}{\mathrm{d}t} = \lambda\,P_{\mathrm{U}} - \mu\,P_{\mathrm{D}}. \tag{19.13}$$

These are the Kolmogorov forward equations of the two-state model. Equation (19.12) is equation (3.10) of Chapter 3, written for the state probability instead of the availability. The two right-hand sides sum to zero, so $P_{\mathrm{U}} + P_{\mathrm{D}}$ stays equal to one.

19.3.2 Transient availability

Substituting $P_{\mathrm{D}} = 1 - P_{\mathrm{U}}$ in equation (19.12) gives

$$\frac{\mathrm{d}P_{\mathrm{U}}}{\mathrm{d}t} = -\lambda\,P_{\mathrm{U}} + \mu\,(1 - P_{\mathrm{U}}) = \mu - (\lambda + \mu)\,P_{\mathrm{U}},$$

a first-order linear equation. Its constant solution is $\mu/(\lambda + \mu)$, and the deviation from that constant decays as $\mathrm{e}^{-(\lambda+\mu)t}$. Matching the initial value $P_{\mathrm{U}}(0)$,

$$P_{\mathrm{U}}(t) = \frac{\mu}{\lambda + \mu} + \left[P_{\mathrm{U}}(0) - \frac{\mu}{\lambda + \mu}\right]\mathrm{e}^{-(\lambda+\mu)t}. \tag{19.14}$$

For a pump that is operating at $t = 0$, $P_{\mathrm{U}}(0) = 1$, the bracket is $\lambda/(\lambda + \mu)$, and

$$A(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu}\,\mathrm{e}^{-(\lambda+\mu)t}, \tag{19.15}$$

which is equation (3.11) of Chapter 3. This is the transient availability of the pump: the probability that it is operational at time t, allowing for failures and repairs before that time. Equation (19.14) adds the dependence on the initial state. A pump that starts under repair, $P_{\mathrm{U}}(0) = 0$, has $A(t) = [\mu/(\lambda + \mu)]\,[1 - \mathrm{e}^{-(\lambda+\mu)t}]$, which rises from zero toward the same limit.

For the example pump, $\lambda + \mu = 0.051\,\mathrm{h}^{-1}$, a time constant of $1/0.051 = 19.61\,\mathrm{h}$, close to the MTTR. At $t = 24\,\mathrm{h}$,

$$A(24\,\mathrm{h}) = \frac{0.05}{0.051} + \frac{0.001}{0.051}\,\mathrm{e}^{-0.051\times 24} = 0.980392 + 0.019608 \times 0.294052 = 0.986158,$$

and at $t = 100\,\mathrm{h}$ the exponential term has fallen to 0.0061 of its initial value and $A(100\,\mathrm{h}) = 0.980512$.

19.3.3 Availability against reliability

For the same pump, initially operational and with the exponential failure rate λ, the reliability, the probability of no failure in $[0, t]$, is

$$R(t) = \mathrm{e}^{-\lambda t}. \tag{19.16}$$

Reliability and transient availability are different quantities. Reliability asks whether the pump has avoided failure throughout the mission; availability asks whether it is operational at the specified instant, even if it failed and was repaired before. For this reason $A(t) \geq R(t)$ at every t, inequality (3.9).

Figure 19.1 plots the two functions for the example pump over 1000 h. The horizontal axis is the time since the pump was put into service in the operating state; the vertical axis is a probability. The solid green curve is the transient availability of equation (19.15), the dashed red curve the reliability of equation (19.16), and the dotted horizontal line the steady-state availability 0.980392. The availability falls from 1 to its steady-state value within a few tens of hours and then stays there; the reliability keeps falling, to 0.9763 at 24 h, 0.9048 at 100 h and 0.3679 at 1000 h, one MTTF. At 1000 h the pump is operational with probability 0.98, although it has failed at least once before that time with probability 0.63.

19.3.4 Steady-state availability

As $t \to \infty$ the exponential term of equation (19.15) vanishes, whatever the initial state in equation (19.14), and the availability tends to the steady-state availability

$$A = \lim_{t\to\infty} A(t) = \frac{\mu}{\lambda + \mu} = \frac{1/\mathrm{MTTR}}{1/\mathrm{MTTF} + 1/\mathrm{MTTR}} = \frac{\mathrm{MTTF}}{\mathrm{MTTF} + \mathrm{MTTR}}, \tag{19.17}$$

the result of equation (3.7). The same value follows directly from the forward equations: in steady state $\mathrm{d}P_{\mathrm{U}}/\mathrm{d}t = 0$, so $\lambda\, P_{\mathrm{U}} = \mu\, P_{\mathrm{D}}$, the probability flow from U to D equals the flow back. For the example,

$$A = \frac{0.05}{0.001 + 0.05} = \frac{0.05}{0.051} = 0.980392,$$

or 98.039 %, and the steady-state unavailability is

$$1 - A = \frac{\lambda}{\lambda + \mu} = \frac{0.001}{0.051} = 0.019608,$$

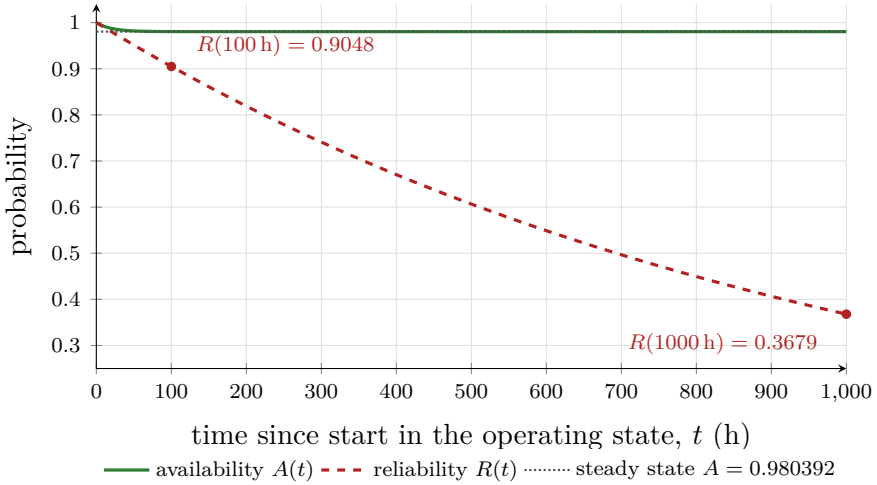


Figure 19.1: Transient availability $A(t) = \mu/(\lambda+\mu) + [\lambda/(\lambda+\mu)]\,\mathrm{e}^{-(\lambda+\mu)t}$ (equation (19.15)) and reliability $R(t) = \mathrm{e}^{-\lambda t}$ (equation (19.16)) of one repairable pump with $\lambda = 0.001\,\mathrm{h}^{-1}$ and $\mu = 0.05\,\mathrm{h}^{-1}$, operating at $t = 0$; illustrative values. Availability credits restorations; reliability does not.

about 1.961 %. A mean repair duration of 20 h thus costs almost 2 % of the long-run availability, although the mean operating time between failures is 1000 h: the unavailability is close to MTTR/MTTF = 0.02. Over a year of 8760 h in continuous service the single pump is expected to be down $8760 \times 0.019608 = 171.8$ h. This figure is the reference against which the redundant pair of Section 19.4 is compared.

19.3.5 The generator matrix

The transition rates of a CTMC are collected in a square matrix, the generator matrix or transition-rate matrix $\mathbf{Q}$, with one row and one column per state. Its off-diagonal entry in row i and column j is the transition rate q_{ij} from i to j; its diagonal entry is minus the total departure rate,

$$q_{ij} \geq 0 \quad (i \neq j), \qquad q_{ii} = -\nu_i = -\sum_{j \neq i} q_{ij}. \tag{19.18}$$

Each row of $\mathbf{Q}$ therefore sums to zero, which is a necessary property of a valid generator. For the two-state pump, with the states ordered (U, D),

$$\mathbf{Q} = \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix}. \tag{19.19}$$

The first row says that from U the process leaves at total rate λ, all of it toward D; the second that from D it leaves at rate μ, all of it toward U.

Remark 19.2. The generator is written in bold, $\mathbf{Q}$, and its entries carry two indices, q_{ij}. It is not the top-event probability Q_{top} of the fault-tree chapters of Part III, and its entries are rates in h^{-1}, not the basic-event probabilities q_i of fault-tree quantification. The book uses the row-vector convention: state probabilities form a row vector that multiplies $\mathbf{Q}$ from the left.

19.3.6 The forward equations in matrix form

For a chain with any finite number of states, let $\mathbf{P}(t) = [P_1(t), P_2(t), \ldots]$ be the row vector of state probabilities, ordered as the rows of $\mathbf{Q}$. The argument of equation (19.11) applies to every

state j: the process is in j at $t+\Delta t$ either because it was in j and did not leave, or because it was in another state i and jumped to j,

$$P_j(t+\Delta t) = P_j(t)\,(1-\nu_j\,\Delta t) + \sum_{i\neq j} P_i(t)\,q_{ij}\,\Delta t + o(\Delta t).$$

Subtracting $P_j(t)$, dividing by Δt, letting $\Delta t \to 0$ and using $q_{jj} = -\nu_j$,

$$\frac{\mathrm{d}P_j}{\mathrm{d}t} = \sum_i P_i(t)\,q_{ij}, \qquad \text{that is,} \qquad \frac{\mathrm{d}\mathbf{P}(t)}{\mathrm{d}t} = \mathbf{P}(t)\,\mathbf{Q}. \tag{19.20}$$

These are the Kolmogorov forward equations in matrix form. Each component reads: rate of change of the probability of state j equals the probability flow into j minus the probability flow out of j. Multiplying by a column of ones, $\mathbf{1}$, and using $\mathbf{Q1} = \mathbf{0}$ (each row sums to zero), $\mathrm{d}(\mathbf{P1})/\mathrm{d}t = \mathbf{PQ1} = 0$: the total probability is conserved.

The equations are linear with constant coefficients, and their formal solution is

$$\mathbf{P}(t) = \mathbf{P}(0)\,\mathrm{e}^{\mathbf{Q}t}, \qquad \mathrm{e}^{\mathbf{Q}t} = \sum_{n=0}^{\infty} \frac{(\mathbf{Q}t)^n}{n!}, \tag{19.21}$$

where $\mathrm{e}^{\mathbf{Q}t}$ is the matrix exponential. Its entry (i,j) is the transition probability $\Pr(X(t) = j \mid X(0) = i)$, and differentiating the series term by term confirms $\mathrm{d}\,\mathrm{e}^{\mathbf{Q}t}/\mathrm{d}t = \mathrm{e}^{\mathbf{Q}t}\mathbf{Q}$. This is the general method for transient state probabilities: the initial state enters through $\mathbf{P}(0)$, the model through $\mathbf{Q}$.

For the two-state pump the matrix exponential has a closed form. The generator of equation (19.19) satisfies $\mathbf{Q}^2 = -(\lambda+\mu)\,\mathbf{Q}$, as direct multiplication shows (the first row of $\mathbf{Q}^2$ is $[\lambda^2+\lambda\mu,\ -\lambda^2-\lambda\mu]$), so $\mathbf{Q}^n = [-(\lambda+\mu)]^{n-1}\mathbf{Q}$ for $n \geq 1$, and the series sums to

$$\mathrm{e}^{\mathbf{Q}t} = \mathbf{I} + \frac{1-\mathrm{e}^{-(\lambda+\mu)t}}{\lambda+\mu}\,\mathbf{Q} = \frac{1}{\lambda+\mu}\begin{bmatrix} \mu+\lambda\,\mathrm{e}^{-(\lambda+\mu)t} & \lambda-\lambda\,\mathrm{e}^{-(\lambda+\mu)t} \\ \mu-\mu\,\mathrm{e}^{-(\lambda+\mu)t} & \lambda+\mu\,\mathrm{e}^{-(\lambda+\mu)t} \end{bmatrix}. \tag{19.22}$$

For a pump initially operational, $\mathbf{P}(0) = [1,\ 0]$ selects the first row, whose first entry is the transient availability of equation (19.15). The eigenvalues of $\mathbf{Q}$ are 0 and $-(\lambda+\mu)$; the zero eigenvalue gives the constant steady-state part and the negative one the decaying transient, and every finite chain in this chapter has the same structure.

19.3.7 Scope of the matrix formulation

A two-state model can be solved with elementary differential equations, as in Section 19.3.2. Industrial systems, however, may require dozens or thousands of states, and the matrix formulation generalizes without change to multiple redundant pumps, generators and electrical supply configurations, limited repair crews, standby units that must be activated, degraded operating states, and multiple failure and restoration pathways. Each of these adds states or transitions, that is, rows and entries of $\mathbf{Q}$, while equations (19.20) and (19.21) remain the same.

When closed-form expressions are inconvenient, the system of differential equations is solved numerically, either by evaluating the matrix exponential or by integrating equation (19.20) with a standard solver for ordinary differential equations. The number of states grows quickly with the number of components (a system of n two-state components distinguished individually has 2^n states), which is why aggregate states such as the pump count of Section 19.1.1 are used whenever the components are interchangeable.

19.4 Two repairable pumps: the three-state model

The redundant pair of Section 19.1.1 is the first model in which the matrix formulation is needed. This section derives its transition rates, writes its generator and its Kolmogorov equations, computes its transient state probabilities, and interprets the point unavailability.

19.4.1 Transition rates of the three-state model

The two pumps are identical, each with failure rate $\lambda = 0.001\,\mathrm{h}^{-1}$ while it operates and repair rate $\mu = 0.05\,\mathrm{h}^{-1}$ while it is under repair. For now two repair crews are available, so two failed pumps can be repaired simultaneously. The states are those of Table 19.1, and each transition rate is derived from the behavior of the individual pumps.

From state 2 to state 1. Both pumps operate, and either can fail; the two independent failures compete. Over a short interval the probability that pump A fails is $\lambda\,\Delta t$, and similarly for pump B, so the probability that at least one fails is $1 - (1 - \lambda\,\Delta t)^2 = 2\lambda\,\Delta t - \lambda^2\Delta t^2$. The probability that both fail in the same interval, $\lambda^2\Delta t^2$, is of order Δt^2 and vanishes in the rate, so there is no direct transition from 2 to 0, and

$$q_{21} = 2\lambda. \tag{19.23}$$

From state 1 to state 2, and from state 1 to state 0. One pump is failed and one operates. The repair of the failed pump is completed at rate μ, and the remaining pump fails at rate λ:

$$q_{12} = \mu, \qquad q_{10} = \lambda. \tag{19.24}$$

From state 0 to state 1. Both pumps are failed. With two crews, each pump has its own repair in progress at rate μ, and either repair may finish first; the time to the first of two independent exponential completions is exponential with the sum of the rates, so

$$q_{01} = 2\mu. \tag{19.25}$$

This last rate is an assumption about the repair resources, not a mathematical necessity. If only one crew can work on one pump at a time, it changes, as Section 19.5.3 shows.

Figure 19.2 draws the model. The three boxes are the states, labeled by the number of operational pumps: green for full redundancy, gold for the degraded state, red for the unavailable state. The red arrows above the boxes are failures, from left to right, with rates 2λ and λ; the green arrows below are repair completions, from right to left, with rates μ and 2μ. The brace marks the two states in which the system is available. All four transitions connect neighboring states, which makes the model a birth–death process, the class of chains in which the state changes by one unit at a time.

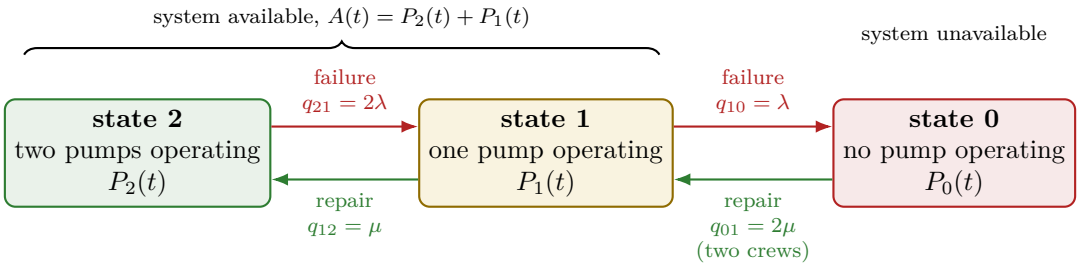


Figure 19.2: Three-state Markov model of two identical repairable pumps in parallel, either of which supplies the required flow. Each operating pump fails at rate λ; each failed pump is repaired at rate μ, with two crews so that two repairs can proceed at once. With one crew the repair rate out of state 0 is μ instead of 2μ.

19.4.2 Generator matrix and Kolmogorov equations

With the states ordered $(2, 1, 0)$, the rates of equations (19.23)–(19.25) and the diagonal rule of equation (19.18) give the generator

$$\mathbf{Q} = \begin{bmatrix} -2\lambda & 2\lambda & 0 \\ \mu & -(\mu+\lambda) & \lambda \\ 0 & 2\mu & -2\mu \end{bmatrix}. \tag{19.26}$$

Substituting $\lambda = 0.001\,\mathrm{h}^{-1}$ and $\mu = 0.05\,\mathrm{h}^{-1}$,

$$\mathbf{Q} = \begin{bmatrix} -0.002 & 0.002 & 0 \\ 0.05 & -0.051 & 0.001 \\ 0 & 0.10 & -0.10 \end{bmatrix} \mathrm{h}^{-1}. \tag{19.27}$$

The rows sum to $-0.002 + 0.002 + 0 = 0$, $0.05 - 0.051 + 0.001 = 0$ and $0 + 0.10 - 0.10 = 0$, and every off-diagonal entry is nonnegative; the generator is valid.

With $\mathbf{P}(t) = [P_2(t),\ P_1(t),\ P_0(t)]$, the forward equations $\mathrm{d}\mathbf{P}/\mathrm{d}t = \mathbf{P}\mathbf{Q}$ expand column by column into

$$\frac{\mathrm{d}P_2}{\mathrm{d}t} = -2\lambda\, P_2 + \mu\, P_1, \tag{19.28}$$

$$\frac{\mathrm{d}P_1}{\mathrm{d}t} = 2\lambda\, P_2 - (\lambda+\mu)\, P_1 + 2\mu\, P_0, \tag{19.29}$$

$$\frac{\mathrm{d}P_0}{\mathrm{d}t} = \lambda\, P_1 - 2\mu\, P_0. \tag{19.30}$$

Each equation follows the same accounting principle: the rate of change of a state probability is the probability inflow minus the probability outflow. P_1 increases when one of the two operating pumps fails in state 2 (inflow $2\lambda P_2$) or when one of the two failed pumps is repaired in state 0 (inflow $2\mu P_0$); it decreases when the remaining pump fails or the failed pump is repaired (outflow $(\lambda+\mu)P_1$). The structure is that of a balance equation, in which the balanced quantity is probability mass rather than energy, mass or electric charge. Adding the three equations,

$$\frac{\mathrm{d}}{\mathrm{d}t}(P_2 + P_1 + P_0) = 0,$$

so the total probability remains one, as equation (19.3) requires.

19.4.3 Transient state probabilities and availability

Both pumps operate initially, $\mathbf{P}(0) = [1,\ 0,\ 0]$, and the state probabilities are $\mathbf{P}(t) = \mathbf{P}(0)\,\mathrm{e}^{\mathbf{Q}t}$. With two crews the solution has a closed form, because the two pumps then fail and are repaired independently of one another: each follows the two-state model of Section 19.3.1, up at time t with probability $P_{\mathrm{U}}(t)$ of equation (19.15) and down with probability $P_{\mathrm{D}}(t) = 1 - P_{\mathrm{U}}(t)$. Counting the operational pumps,

$$P_2(t) = P_{\mathrm{U}}(t)^2, \qquad P_1(t) = 2\,P_{\mathrm{U}}(t)\,P_{\mathrm{D}}(t), \qquad P_0(t) = P_{\mathrm{D}}(t)^2. \tag{19.31}$$

These satisfy equation (19.28): $\mathrm{d}P_2/\mathrm{d}t = 2P_{\mathrm{U}}\,(-\lambda P_{\mathrm{U}} + \mu P_{\mathrm{D}}) = -2\lambda P_2 + \mu P_1$, and similarly for the other two. Since $P_{\mathrm{D}}(t) = [\lambda/(\lambda+\mu)]\,[1 - \mathrm{e}^{-(\lambda+\mu)t}]$, the unavailability is

$$1 - A(t) = P_0(t) = \left(\frac{\lambda}{\lambda+\mu}\right)^2 \left[1 - \mathrm{e}^{-(\lambda+\mu)t}\right]^2, \tag{19.32}$$

which contains the exponentials $\mathrm{e}^{-(\lambda+\mu)t}$ and $\mathrm{e}^{-2(\lambda+\mu)t}$; the eigenvalues of the generator of equation (19.26) are 0, $-(\lambda+\mu) = -0.051\,\mathrm{h}^{-1}$ and $-2(\lambda+\mu) = -0.102\,\mathrm{h}^{-1}$.

A general form, valid for any repair rate q_{01} out of state 0, is obtained from the eigenvalues. The nonzero eigenvalues s_1 and s_2 of the generator are the roots of its characteristic polynomial divided by s,

$$s^2 + (3\lambda + \mu + q_{01})\, s + (2\lambda^2 + 2\lambda q_{01} + \mu q_{01}) = 0, \tag{19.33}$$

whose coefficients are minus the trace of $\mathbf{Q}$ and the sum of its principal 2×2 minors. Then $P_0(t) = c_0 + c_1 \mathrm{e}^{s_1 t} + c_2 \mathrm{e}^{s_2 t}$, and the three constants follow from the initial conditions $P_0(0) = 0$, $P_0'(0) = \lambda P_1(0) - q_{01} P_0(0) = 0$ and $P_0''(0) = \lambda P_1'(0) = \lambda \cdot 2\lambda P_2(0) = 2\lambda^2$, obtained by differentiating equation (19.30):

$$P_0(t) = 2\lambda^2 \left[\frac{1}{s_1 s_2} + \frac{\mathrm{e}^{s_1 t}}{s_1 (s_1 - s_2)} + \frac{\mathrm{e}^{s_2 t}}{s_2 (s_2 - s_1)} \right]. \tag{19.34}$$

For two crews, $q_{01} = 2\mu = 0.1\,\mathrm{h}^{-1}$, the polynomial is $s^2 + 0.153\, s + 0.005202$ with roots -0.102 and $-0.051\,\mathrm{h}^{-1}$, and equation (19.34) reduces to equation (19.32).

Table 19.2 gives the state probabilities at seven times. Its columns are the time, the three state probabilities in percent, and the availability $A(t) = P_2(t) + P_1(t)$. At $t = 24\,\mathrm{h}$, $P_\mathrm{U} = 0.986158$ from Section 19.3.2, so $P_2 = 0.986158^2 = 0.972507$, $P_1 = 2 \times 0.986158 \times 0.013842 = 0.027301$ and $P_0 = 0.013842^2 = 0.000192$, and

$$A(24\,\mathrm{h}) = 0.972507 + 0.027301 = 0.999808,$$

or 99.9808 %. At 500 h, $A(500\,\mathrm{h}) = 0.999616$, or 99.9616 %, and the state probabilities have reached their limits to the precision shown: they approach a steady distribution over a few time constants of $1/0.051 = 19.6\,\mathrm{h}$.

Table 19.2: Transient state probabilities and availability of two repairable pumps, two crews, both pumps operating at $t = 0$ ($\lambda = 0.001\,\mathrm{h}^{-1}$, $\mu = 0.05\,\mathrm{h}^{-1}$; illustrative), from $\mathbf{P}(t) = \mathbf{P}(0)\,\mathrm{e}^{\mathbf{Q}t}$.

Time (h)	Both operating, P_2	One operating, P_1	Both failed, P_0	Availability $A(t)$
0	100.0000 %	0 %	0 %	100.0000 %
10	98.4395 %	1.5544 %	0.00614 %	99.9939 %
24	97.2507 %	2.7301 %	0.01916 %	99.9808 %
50	96.4173 %	3.5500 %	0.03268 %	99.9673 %
100	96.1403 %	3.8217 %	0.03798 %	99.9620 %
200	96.1170 %	3.8445 %	0.03844 %	99.9616 %
500	96.1169 %	3.8447 %	0.03845 %	99.9616 %

Figure 19.3 plots the transient availability over the first 200 h. The horizontal axis is the time since both pumps were put into service; the vertical axis is the availability, on an axis that spans only 0.99915 to 1. The solid blue curve is the two-crew model of this section, equation (19.32); the dashed red curve is the one-crew model of Section 19.5.3, computed from equation (19.34) with $q_{01} = \mu$; the dotted lines are the two steady-state values, and the dots mark $t = 24\,\mathrm{h}$. Both curves start at 1, fall as the probability of having lost one pump and then the second builds up, and level off at their steady-state values within about 200 h. The one-crew curve falls further and more slowly, because it lingers longer in state 0; at 24 h the two availabilities are 0.999808 and 0.999735.

19.4.4 Point unavailability and the probability of an outage

The probability $P_0(24\,\mathrm{h}) = 0.000192$ is the probability that the system is unavailable *exactly at* hour 24. It is not the probability that the system experienced at least one period of complete

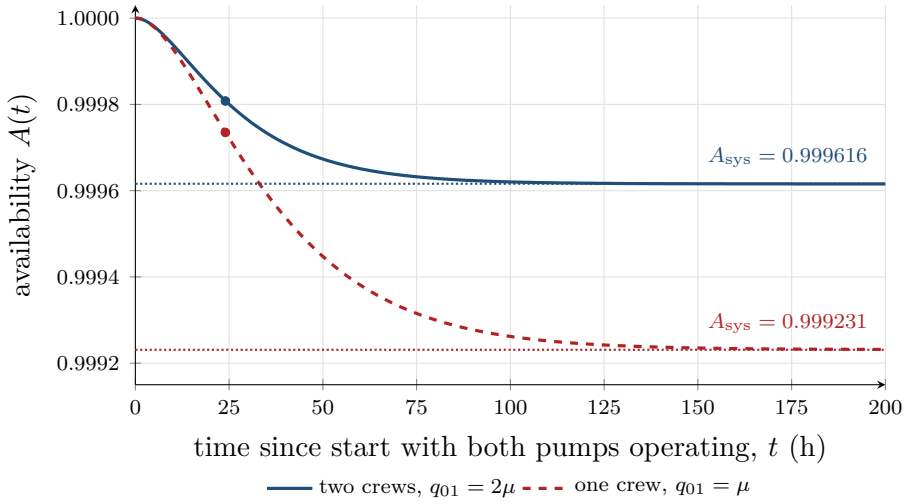


Figure 19.3: Transient availability of two repairable pumps in parallel, both operating at $t = 0$, with $\lambda = 0.001\,\text{h}^{-1}$ and $\mu = 0.05\,\text{h}^{-1}$ (illustrative). Two crews: $A(t) = 1 - 0.000384468\,[1 - \text{e}^{-0.051t}]^2$ (equation (19.32)). One crew: $A(t) = 1 - 0.000768640 - 0.002407786\,\text{e}^{-0.058589t} + 0.003176426\,\text{e}^{-0.044411t}$ (equation (19.34), t in hours). Dotted lines: the steady-state values. Dots: $A(24\,\text{h}) = 0.999808$ (two crews) and 0.999735 (one crew).

loss of pumping during the first 24 h: the system could have entered state 0 earlier and returned to state 1 before hour 24. The distinction is central to the analysis of repairable systems, and the second probability is computed with a modified model in which state 0 is made absorbing, $q_{01} = 0$. The probability of being in state 0 in that model is the probability of having reached it at least once by t, and its complement is the reliability of the redundant system with repair, the probability of no complete outage in $[0, t]$.

For the example, the absorbing model gives a probability of at least one complete outage in 24 h of 0.000393, about twice $P_0(24\,\text{h}) = 0.000192$, because an outage that starts in the first 24 h often ends before hour 24. Over a year of 8760 h the probability of at least one complete outage is 0.281, while the point unavailability at any instant of that year is below 0.0004. The mean time to the first complete outage, starting with both pumps operating, follows from the mean times T_2 and T_1 to reach state 0 from states 2 and 1. From state 2 the system first spends a mean of $1/(2\lambda)$ and then moves to state 1, so $T_2 = 1/(2\lambda) + T_1$; from state 1 it spends a mean of $1/(\lambda + \mu)$ and then moves to state 2 with probability $\mu/(\lambda + \mu)$ or to state 0 otherwise, so $T_1 = 1/(\lambda + \mu) + [\mu/(\lambda + \mu)]\,T_2$. Solving,

$$T_1 = \frac{2\lambda + \mu}{2\lambda^2}, \qquad T_2 = \frac{3\lambda + \mu}{2\lambda^2} = \frac{0.003 + 0.05}{2 \times 10^{-6}} = 26\,500\,\text{h}, \tag{19.35}$$

about three years, against a mean of $1/(2\lambda) = 500\,\text{h}$ to the first failure of either pump. Redundancy with repair makes a complete outage rare; it does not make it impossible, and its probability over a long interval is far larger than the point unavailability suggests.

Common mistake

Reading the point unavailability $P_0(t)$ as the probability that an outage has occurred by t. The first is a probability about one instant; the second is a first-passage probability, which requires the model with state 0 made absorbing and is always at least as large.

19.5 Steady-state availability, repair resources and failure frequency

The transient solution settles into a stationary regime, and most availability requirements are stated for it. This section derives the stationary distribution of the pump system, shows how a limited repair crew changes it, converts it into expected annual downtime, separates failure frequency from unavailability, and closes with a worked table over a range of rates and crew numbers.

19.5.1 The stationary distribution

In steady state the probabilities no longer change, $\mathrm{d}\mathbf{P}/\mathrm{d}t = \mathbf{0}$. The stationary distribution $\boldsymbol{\pi} = [\pi_2,\ \pi_1,\ \pi_0]$ therefore solves

$$\boldsymbol{\pi}\,\mathbf{Q} = \mathbf{0}, \qquad \pi_2 + \pi_1 + \pi_0 = 1. \tag{19.36}$$

The first condition is three linear equations of which only two are independent, since their sum, $\boldsymbol{\pi}\mathbf{Q}\mathbf{1}$ with $\mathbf{1}$ a column of ones, vanishes identically because every row of $\mathbf{Q}$ sums to zero; the normalization supplies the third. For a finite chain in which every state can be reached from every other (an irreducible chain) the solution exists, is unique, and is the limit of $\mathbf{P}(t)$ as $t \to \infty$ from any initial state.

For a birth–death chain the equations reduce to a local balance of probability flow across each boundary between neighboring states. In steady state the flow from state 2 to state 1 equals the flow back,

$$2\lambda\,\pi_2 = \mu\,\pi_1 \quad\Longrightarrow\quad \pi_1 = \frac{2\lambda}{\mu}\,\pi_2,$$

and the flow from state 1 to state 0 equals the flow back,

$$\lambda\,\pi_1 = 2\mu\,\pi_0 \quad\Longrightarrow\quad \pi_0 = \frac{\lambda}{2\mu}\,\pi_1 = \left(\frac{\lambda}{\mu}\right)^2 \pi_2.$$

With the ratio of failure rate to repair rate

$$\rho = \frac{\lambda}{\mu}, \tag{19.37}$$

these read $\pi_1 = 2\rho\,\pi_2$ and $\pi_0 = \rho^2\,\pi_2$. The normalization gives $\pi_2(1 + 2\rho + \rho^2) = \pi_2(1+\rho)^2 = 1$, so

$$\pi_2 = \frac{1}{(1+\rho)^2}, \qquad \pi_1 = \frac{2\rho}{(1+\rho)^2}, \qquad \pi_0 = \frac{\rho^2}{(1+\rho)^2}. \tag{19.38}$$

Substituting into the middle column of equation (19.36), $2\lambda\pi_2 - (\lambda+\mu)\pi_1 + 2\mu\pi_0 = [2\lambda - 2\rho(\lambda+\mu) + 2\mu\rho^2]\,\pi_2 = 0$, confirms the solution. The steady-state availability of the redundant system is

$$A_{\mathrm{sys}} = 1 - \pi_0 = 1 - \left(\frac{\lambda}{\lambda+\mu}\right)^2, \tag{19.39}$$

since $\rho/(1+\rho) = \lambda/(\lambda+\mu)$. This is exactly the availability of a parallel system of two independent components, each with the unavailability $\lambda/(\lambda+\mu)$ of Section 19.3.4: with two crews the pumps are independent, as in equation (19.31), and π_2, π_1, π_0 are the binomial probabilities of two, one and zero pumps up.

19.5.2 Steady-state availability of the redundant pumps

For $\lambda = 0.001\,\mathrm{h}^{-1}$ and $\mu = 0.05\,\mathrm{h}^{-1}$, $\rho = 0.001/0.05 = 0.02$ and $(1+\rho)^2 = 1.0404$, so

$$\pi_2 = \frac{1}{1.0404} = 0.961169, \qquad \pi_1 = \frac{0.04}{1.0404} = 0.038447, \qquad \pi_0 = \frac{0.0004}{1.0404} = 0.00038447,$$

and

$$A_{\mathrm{sys}} = 1 - 0.00038447 = 0.99961553,$$

or 99.9616 %, with a steady-state unavailability of 0.03845 %. These are the limits reached by the transient solution in Table 19.2. Redundancy raises the steady-state availability from 98.04 % for one pump to 99.9616 % for the pair, and divides the unavailability by $0.019608/0.00038447 = 51$. The result requires independent failure processes, independent repair processes, and enough repair capacity to work on both failed pumps at once. The next subsection relaxes the last assumption.

19.5.3 Limited repair resources

Suppose the pumps keep their failure rate $\lambda = 0.001\,\mathrm{h}^{-1}$ and their individual repair rate $\mu = 0.05\,\mathrm{h}^{-1}$, but the facility has only one repair crew. When both pumps are failed the crew works on one at a time, so the transition from state 0 to state 1 occurs at rate $q_{01} = \mu$ instead of 2μ, and the generator becomes

$$\mathbf{Q}_{\text{one crew}} = \begin{bmatrix} -2\lambda & 2\lambda & 0 \\ \mu & -(\mu+\lambda) & \lambda \\ 0 & \mu & -\mu \end{bmatrix}. \tag{19.40}$$

Only one rate has changed, in the last row. The first boundary balance is unchanged, $2\lambda\pi_2 = \mu\pi_1$, so $\pi_1 = 2\rho\,\pi_2$; the second is now $\lambda\pi_1 = \mu\pi_0$, so $\pi_0 = \rho\,\pi_1 = 2\rho^2\,\pi_2$. Normalizing, $\pi_2(1 + 2\rho + 2\rho^2) = 1$ and

$$\pi_2 = \frac{1}{1+2\rho+2\rho^2}, \qquad \pi_1 = \frac{2\rho}{1+2\rho+2\rho^2}, \qquad \pi_0 = \frac{2\rho^2}{1+2\rho+2\rho^2}. \tag{19.41}$$

For $\rho = 0.02$,

$$\pi_0 = \frac{2 \times 0.0004}{1 + 0.04 + 0.0008} = \frac{0.0008}{1.0408} = 0.00076864,$$

and $A_{\mathrm{sys}} = 1 - 0.00076864 = 0.99923136$, or 99.9231 %. The transient approach to this value is the dashed curve of Figure 19.3: with $q_{01} = 0.05\,\mathrm{h}^{-1}$, equation (19.33) is $s^2 + 0.103\,s + 0.002602 = 0$, with roots -0.058589 and $-0.044411\,\mathrm{h}^{-1}$, time constants of 17.1 and 22.5 h.

Table 19.3 compares the two repair arrangements. Its rows are the two models, its columns the steady-state availability and unavailability. The shared crew approximately doubles the steady-state unavailability, by the factor $2(1+\rho)^2/(1+2\rho+2\rho^2) = 1.9992$ for $\rho = 0.02$.

Table 19.3: Steady-state availability of two repairable pumps with two independent repair processes and with one shared crew ($\lambda = 0.001\,\mathrm{h}^{-1}$, $\mu = 0.05\,\mathrm{h}^{-1}$; illustrative).

Repair capacity	Steady availability	Steady unavailability
Two independent repair processes	99.9616 %	0.03845 %
One shared repair crew	99.9231 %	0.07686 %

The pumps have not become more failure-prone: their failure rates are unchanged. When both pumps are down, however, the system takes longer to leave the failed state, because only

one repair can progress at a time; the mean time to leave state 0 is $1/\mu = 20\,\text{h}$ instead of $1/(2\mu) = 10\,\text{h}$. The difference matters for remote transmission assets, for data centers with limited maintenance staffing and for facilities that require specialized repair personnel.

Principle 19.2. *Availability depends on the complete failure-and-restoration architecture, not only on the failure rates of the equipment. Repair crews, spare parts and access are transitions of the model, and changing one of them changes the availability.*

19.5.4 Expected annual downtime

If the system has reached steady state and the model remains stationary, the expected number of hours per year in which the system is unavailable is

$$\mathbb{E}[D_{\text{yr}}] = 8760\,\text{h} \times (1 - A_{\text{sys}}), \tag{19.42}$$

as in equation (3.21). With two independent repair processes,

$$\mathbb{E}[D_{\text{yr}}] = 8760 \times 0.00038447 = 3.37\,\text{h per year},$$

and with one shared crew

$$\mathbb{E}[D_{\text{yr}}] = 8760 \times 0.00076864 = 6.73\,\text{h per year},$$

against 171.8 h per year for a single pump. These are expected values under steady-state assumptions. They do not say how many separate outages occur, nor how long any particular outage lasts; a year can contain no outage at all, or one long outage. The number of outages requires the transitions into the unavailable state, which the next subsection counts.

19.5.5 Failure frequency and mean down time

In the pump system the only transition into state 0 is from state 1, and it occurs at rate λ while state 1 is occupied. In steady state the expected number of entries into state 0 per unit time, the failure frequency of the system, is therefore

$$f_{\text{F}} = \pi_1\,\lambda. \tag{19.43}$$

For the two-crew model, $\pi_1 = 0.03844675$, so

$$f_{\text{F}} = 0.03844675 \times 0.001\,\text{h}^{-1} = 3.8447 \times 10^{-5}\,\text{h}^{-1},$$

and on an annual basis

$$8760\,\text{h} \times 3.8447 \times 10^{-5}\,\text{h}^{-1} = 0.337\ \text{outages per year},$$

about one complete outage every $1/0.337 = 2.97$ years in reciprocal-frequency terms. The number of outages in any actual year remains random.

When the system enters state 0 with two crews, both repairs are in progress and either may finish first, so the system leaves state 0 at rate 2μ and the mean down time, the mean duration of a complete outage, is

$$\text{MDT} = \frac{1}{q_{01}} = \frac{1}{2\mu} = \frac{1}{0.1\,\text{h}^{-1}} = 10\,\text{h}. \tag{19.44}$$

The long-run fraction of time in the unavailable state is the frequency of entering it multiplied by the mean duration of each visit,

$$1 - A_{\text{sys}} = f_{\text{F}} \cdot \text{MDT}, \tag{19.45}$$

a relation that holds for any stationary regenerative process, and that for the Markov model follows from equation (19.43) and the balance $\lambda\pi_1 = q_{01}\pi_0$: $f_{\mathrm{F}} \cdot \mathrm{MDT} = \lambda\pi_1/q_{01} = \pi_0$. Numerically,

$$1 - A_{\mathrm{sys}} = (3.8447 \times 10^{-5}\,\mathrm{h}^{-1})(10\,\mathrm{h}) = 0.00038447,$$

exactly the value obtained from the stationary distribution. For one crew, $\mathrm{MDT} = 1/\mu = 20\,\mathrm{h}$, and $f_{\mathrm{F}} = 0.038432 \times 0.001 = 3.8432 \times 10^{-5}\,\mathrm{h}^{-1}$, or 0.3367 per year: the frequency of entering the failed state is almost unchanged, but the mean outage doubles, which is why the unavailability doubles.

For a system whose failed set contains several states, the same argument applies to the set: f_{F} is the total steady-state probability flow from the available states into the failed states, $\sum_{i\in\mathrm{up}} \sum_{j\in\mathrm{down}} \pi_i\, q_{ij}$, and the MDT is the probability of the failed set divided by f_{F}. This is the frequency and duration method of power-system reliability evaluation. Dividing the probability of the available set by the same frequency gives the mean up time between outages, $A_{\mathrm{sys}}/f_{\mathrm{F}} = 0.999616/(3.8447 \times 10^{-5}) = 26\,000\,\mathrm{h}$, and the mean cycle from one outage to the next is $1/f_{\mathrm{F}} = 26\,010\,\mathrm{h}$. The mean up time is the T_1 of equation (19.35), because after every outage the system restarts in state 1 with one pump repaired, and it is 500 h shorter than the mean time to the first outage from state 2.

Figure 19.4 shows the three quantities on one illustrative history. The horizontal axis is time; the blue line is the state of the system, at the level of state 2, 1 or 0; the shaded bands are the periods in state 0. The lower brace marks one outage, whose mean is the MDT; the long brace marks the interval from one entry into state 0 to the next, whose mean is $1/f_{\mathrm{F}}$; the upper brace marks the up period between them. Most excursions to state 1 end with a repair and a return to state 2; only a fraction $\lambda/(\lambda+\mu) = 0.0196$ of them end in state 0. Table 19.4 states the question each metric answers.

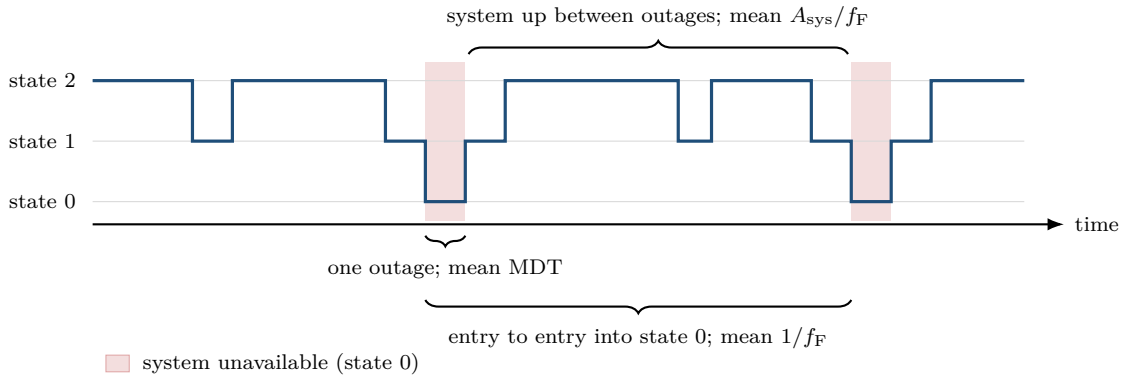


Figure 19.4: One illustrative history of the two-pump system, schematic and not to scale. Excursions to state 1 usually end in a repair; a few end in state 0, an outage. The failure frequency f_{F} counts the outages, the mean down time measures their duration, and the unavailability is their product, equation (19.45).

Table 19.4: Three distinct steady-state metrics of a repairable system.

Metric	Expression	Question answered
Failure frequency	$f_{\mathrm{F}} = \pi_1\lambda$	How often does the system enter a failed state?
Mean down time	$\mathrm{MDT} = 1/q_{01}$	How long does a failure episode last?
Unavailability	$1 - A_{\mathrm{sys}} = \pi_0$	What fraction of time is the system failed?

The three metrics are connected by equation (19.45), but none should be confused with another. Two systems with the same unavailability can differ by a factor of ten in failure frequency, as the two pumps of Chapter 3 did, and a requirement stated in one of the three is not met by a design justified in another.

19.5.6 Unavailability, failure frequency and mean down time: a worked table

The model of this section can be evaluated for any failure rate, repair rate and number of crews. The model is the three-state chain of Figure 19.2 with identical pumps, independent exponential failures, exponential repair completions, no common-cause failures and no change in the hazard of the surviving pump after its partner fails. With $\rho = \lambda/\mu$, the stationary distribution is equation (19.38) for two crews and equation (19.41) for one crew; then

$$1 - A_{\text{sys}} = \pi_0, \qquad f_{\text{F}} = 8760\,\pi_1\lambda \text{ per year}, \qquad \text{MDT} = \frac{1}{c\,\mu}, \qquad \mathbb{E}[D_{\text{yr}}] = 8760\,\pi_0,$$

with $c = 2$ or 1 the number of crews.

Table 19.5 evaluates these quantities. The first two rows are the example of this chapter; the next two reduce the repair rate to $0.02\,\text{h}^{-1}$, a mean repair duration of 50 h; the rest raise the failure rate to 0.002 and $0.005\,\text{h}^{-1}$, the last pair combining the higher failure rate with the slower repair. The row $\lambda = 0.001\,\text{h}^{-1}$, $\mu = 0.02\,\text{h}^{-1}$, one crew shows the arithmetic: $\rho = 0.05$, $1 + 2\rho + 2\rho^2 = 1.105$, $\pi_1 = 0.1/1.105 = 0.090498$, $\pi_0 = 0.005/1.105 = 0.0045249$, $f_{\text{F}} = 8760 \times 0.090498 \times 0.001 = 0.7928$ per year, MDT $= 1/0.02 = 50$ h, and $\mathbb{E}[D_{\text{yr}}] = 8760 \times 0.0045249 = 39.64$ h per year, which equals 0.7928×50 h as equation (19.45) requires.

Table 19.5: Steady-state unavailability, failure frequency, mean down time and expected annual downtime of two redundant pumps, for several failure rates, repair rates and numbers of repair crews; illustrative values.

λ (h^{-1})	μ (h^{-1})	Crews	π_1	$1 - A_{\text{sys}}$	f_{F} (yr^{-1})	MDT (h)	$\mathbb{E}[D_{\text{yr}}]$ (h/yr)
0.001	0.05	2	0.038447	0.0003845	0.3368	10	3.37
0.001	0.05	1	0.038432	0.0007686	0.3367	20	6.73
0.001	0.02	2	0.090703	0.0022676	0.7946	25	19.86
0.001	0.02	1	0.090498	0.0045249	0.7928	50	39.64
0.002	0.05	2	0.073964	0.0014793	1.2959	10	12.96
0.002	0.05	1	0.073855	0.0029542	1.2939	20	25.88
0.005	0.05	2	0.165289	0.0082645	7.2397	10	72.40
0.005	0.05	1	0.163934	0.0163934	7.1803	20	143.61
0.005	0.02	2	0.320000	0.0400000	14.016	25	350.40
0.005	0.02	1	0.307692	0.0769231	13.477	50	673.85

Three trends follow from the table. First, the number of crews barely changes the failure frequency, which is set by the occupation of state 1 and the failure rate of the remaining pump, but it doubles the mean down time and therefore nearly doubles the unavailability; the ratio $2(1+\rho)^2/(1+2\rho+2\rho^2)$ falls from 1.9992 at $\rho = 0.02$ to 1.92 at $\rho = 0.25$. Second, the unavailability of the redundant pair grows roughly with ρ^2: doubling the failure rate from 0.001 to $0.002\,\text{h}^{-1}$ multiplies it by 3.85, and lowering the repair rate from 0.05 to $0.02\,\text{h}^{-1}$, which multiplies ρ by 2.5, multiplies it by 5.9. Third, a slower repair affects the system twice, by raising the frequency of entering the failed state (more time in state 1 exposes the remaining pump) and by lengthening each outage, whereas a higher failure rate affects only the frequency.

Repair-resource constraints thus influence the unavailability strongly even when the failure rates of the components do not change.

19.6 Extensions, applications and validation of Markov models

The three-state model rests on assumptions that real systems often violate. This section shows how the model is extended to load sharing, common-cause failures and age-dependent failure rates, applies the same construction to data center power continuity and to transmission networks, and lists the common mistakes and the checks that guard against them.

19.6.1 Load sharing and state-dependent failure rates

The model so far assumes that each pump fails at the same rate whether or not its partner is operating. In some systems the loss of one component increases the load on the remaining one, the load-sharing dependence of Section 13.5.2. Let λ_2 be the failure rate of each pump while two operate and λ_1 the failure rate of the remaining pump while it operates alone, the indices naming the state in which each rate applies (Chapter 13 wrote the shared-load rate λ_0). The transitions become $q_{21} = 2\lambda_2$ and $q_{10} = \lambda_1$, and with two crews

$$\mathbf{Q} = \begin{bmatrix} -2\lambda_2 & 2\lambda_2 & 0 \\ \mu & -(\mu + \lambda_1) & \lambda_1 \\ 0 & 2\mu & -2\mu \end{bmatrix}. \tag{19.46}$$

The model is still a CTMC, because each transition rate is determined by the current state. Figure 19.5(a) draws it. The boundary balances give $\pi_1 = (2\lambda_2/\mu)\,\pi_2$ and $\pi_0 = [\lambda_1/(2\mu)]\,\pi_1$; with the two ratios $\rho_{21} = 2\lambda_2/\mu$ and $\rho_{10} = \lambda_1/(2\mu)$ and the normalization $\pi_2(1 + \rho_{21} + \rho_{21}\rho_{10}) = 1$,

$$\pi_0 = \frac{\rho_{21}\rho_{10}}{1 + \rho_{21} + \rho_{21}\rho_{10}}. \tag{19.47}$$

Take $\lambda_2 = 0.001\,\mathrm{h}^{-1}$ and, after one pump fails, $\lambda_1 = 0.005\,\mathrm{h}^{-1}$, so that the remaining pump is five times more failure-prone, with $\mu = 0.05\,\mathrm{h}^{-1}$ (illustrative). Then $\rho_{21} = 0.002/0.05 = 0.04$, $\rho_{10} = 0.005/0.1 = 0.05$, $\rho_{21}\rho_{10} = 0.002$, and

$$\pi_0 = \frac{0.002}{1 + 0.04 + 0.002} = \frac{0.002}{1.042} = 0.0019194, \qquad A_{\mathrm{sys}} = 0.998081,$$

or 99.8081 %, against 99.9616 % for the model without load sharing. State-dependent loading multiplies the steady-state unavailability by $0.0019194/0.00038447 = 4.99$, close to the factor $\lambda_1/\lambda_2 = 5$ by which it raises the rate of the only transition into state 0. The failure frequency is $f_{\mathrm{F}} = \pi_1\lambda_1 = 0.038388 \times 0.005\,\mathrm{h}^{-1}$, or 1.68 per year, and the MDT is unchanged at 10 h. With $\mu = 0$ the model reduces to the nonrepairable load-sharing model of Figure 13.7, which is a check on both.

19.6.2 Common-cause failures as a direct transition

Suppose a shared electrical event can disable both pumps at once, with occurrence rate λ_C. The model then gains a direct transition from state 2 to state 0 at rate λ_C, which the independent model excluded in Section 19.4.1. The model must also say what the event does in state 1: if it disables the remaining operating pump, it adds λ_C to the transition from 1 to 0. In general the model must define which mechanisms cause each transition and how the repairs restore service, and the previous generator cannot be reused without modification. The failure logic of a fault tree, with its common-cause basic events, thus becomes a set of transitions of the state model.

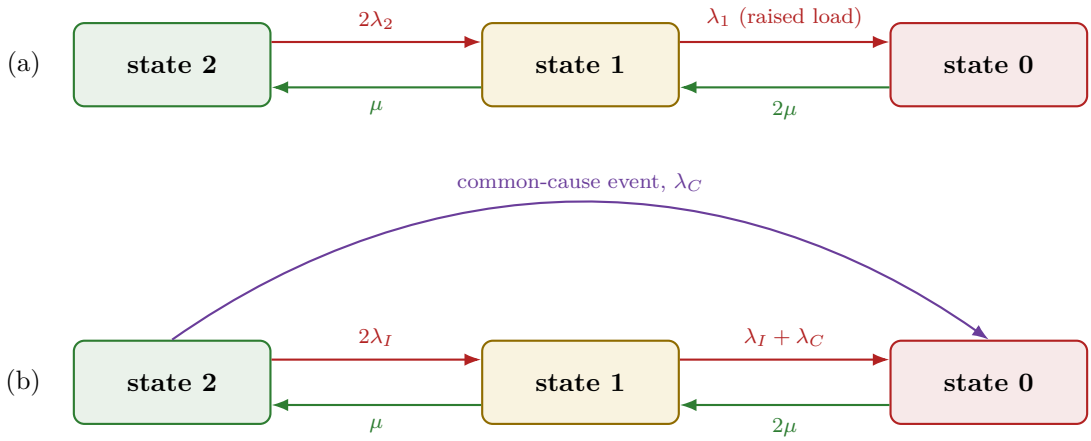


Figure 19.5: Extensions of the three-state model of Figure 19.2. (a) Load sharing: each pump fails at λ_2 while two operate and the survivor at λ_1. (b) Common-cause failure: an independent rate λ_I per pump, and a shared event at rate λ_C that fails both pumps in state 2 (direct transition to state 0) and the remaining pump in state 1. Two repair crews in both.

Figure 19.5(b) draws the model with the beta-factor split of Chapter 14: each pump keeps a total failure rate $\lambda = 0.001\,\mathrm{h}^{-1}$, of which a fraction $\beta = 0.1$ is common cause, so $\lambda_C = 0.0001\,\mathrm{h}^{-1}$ and the independent rate is $\lambda_I = 0.0009\,\mathrm{h}^{-1}$ (illustrative). The generator, with two crews, is

$$\mathbf{Q} = \begin{bmatrix} -(2\lambda_I + \lambda_C) & 2\lambda_I & \lambda_C \\ \mu & -(\mu + \lambda_I + \lambda_C) & \lambda_I + \lambda_C \\ 0 & 2\mu & -2\mu \end{bmatrix} = \begin{bmatrix} -0.0019 & 0.0018 & 0.0001 \\ 0.05 & -0.051 & 0.001 \\ 0 & 0.10 & -0.10 \end{bmatrix} \mathrm{h}^{-1}.$$

The chain is no longer a birth–death chain, but the flow balance still holds across any boundary that splits the states in two. Across the boundary between state 2 and the set $\{1, 0\}$, $(2\lambda_I + \lambda_C)\,\pi_2 = \mu\,\pi_1$, so $\pi_1 = 0.038\,\pi_2$; across the boundary around state 0, $2\mu\,\pi_0 = \lambda_C\pi_2 + (\lambda_I + \lambda_C)\pi_1 = (0.0001 + 0.001 \times 0.038)\,\pi_2$, so $\pi_0 = 0.00138\,\pi_2$. Normalizing, $\pi_2 = 1/1.03938 = 0.962112$, $\pi_1 = 0.036560$ and $\pi_0 = 0.0013277$, so that $A_{\mathrm{sys}} = 0.998672$. A common-cause fraction of 10 % multiplies the unavailability by $0.0013277/0.00038447 = 3.45$, although each pump fails no more often than before. The failure frequency is $f_{\mathrm{F}} = \lambda_C\pi_2 + (\lambda_I + \lambda_C)\pi_1 = 9.62 \times 10^{-5} + 3.66 \times 10^{-5} = 1.328 \times 10^{-4}\,\mathrm{h}^{-1}$, or 1.16 per year, of which 72 % comes from the direct transition; the MDT remains 10 h, and $f_{\mathrm{F}} \cdot \mathrm{MDT} = 0.0013277 = \pi_0$, as equation (19.45) requires.

19.6.3 Age dependence and the phase-type approximation

Suppose the pumps have Weibull lifetimes with increasing hazards. Their propensity to fail then depends on accumulated operating age, and a three-state model that records only the number of functioning pumps is generally insufficient. State 1 might mean that one newly installed pump remains, or that one heavily worn pump remains, and the two conditions have different future hazards; the Markov property of equation (19.6) fails for that state description. The state can be enlarged with the virtual ages of the pumps,

$$X(t) = (N_{\mathrm{op}}(t),\ V_{\mathrm{A}}(t),\ V_{\mathrm{B}}(t)),$$

where $N_{\mathrm{op}}(t)$ is the number of operating pumps and V_{A}, V_{B} are the virtual ages of Chapter 18. The enlarged process can be Markovian, but it now contains continuous variables, and its analysis requires a piecewise-deterministic Markov process, the supplementary-variable method, a semi-Markov model or Monte Carlo simulation. A finite-state CTMC is convenient precisely when the transition times are memoryless within each modeled state.

A practical alternative approximates a nonexponential lifetime by several exponential phases. Instead of representing a pump as simply operational or failed, the model passes it through

$$\text{new} \longrightarrow \text{aging phase 1} \longrightarrow \text{aging phase 2} \longrightarrow \text{failed},$$

each transition with an exponential rate. The expanded model is still a CTMC, but the total time to failure, a sum of exponential phases, is not exponential; such a lifetime has a phase-type distribution. With k phases of equal rate $k\lambda$ the lifetime is Erlang distributed with mean $1/\lambda$ and coefficient of variation $1/\sqrt{k}$, and its hazard increases with age from zero toward $k\lambda$. The approach represents more realistic lifetime distributions at the cost of additional states and parameters.

Figure 19.6 shows the construction and its effect. The upper panel is the chain of phases for one pump, with three transitions at rate 3λ and, for a repairable pump, a restoration from the failed state back to new. The lower panel plots the hazard functions of one, two and three phases with the same mean of 1000 h; the horizontal axis is the age of the pump and the vertical axis the hazard in units of $10^{-3}\,\mathrm{h}^{-1}$. One phase is the exponential model with constant hazard $0.001\,\mathrm{h}^{-1}$; two and three phases have hazards that start at zero and rise past $0.001\,\mathrm{h}^{-1}$, reaching 0.00171 and $0.00241\,\mathrm{h}^{-1}$ at 3000 h. The three-phase chain thus approximates a wear-out lifetime with increasing hazard while keeping every holding time exponential. Because the steady-state availability depends only on the mean up and down times, Section 3.3.2, the phases leave it unchanged; they change the transient, as Exercise 19.3 shows.

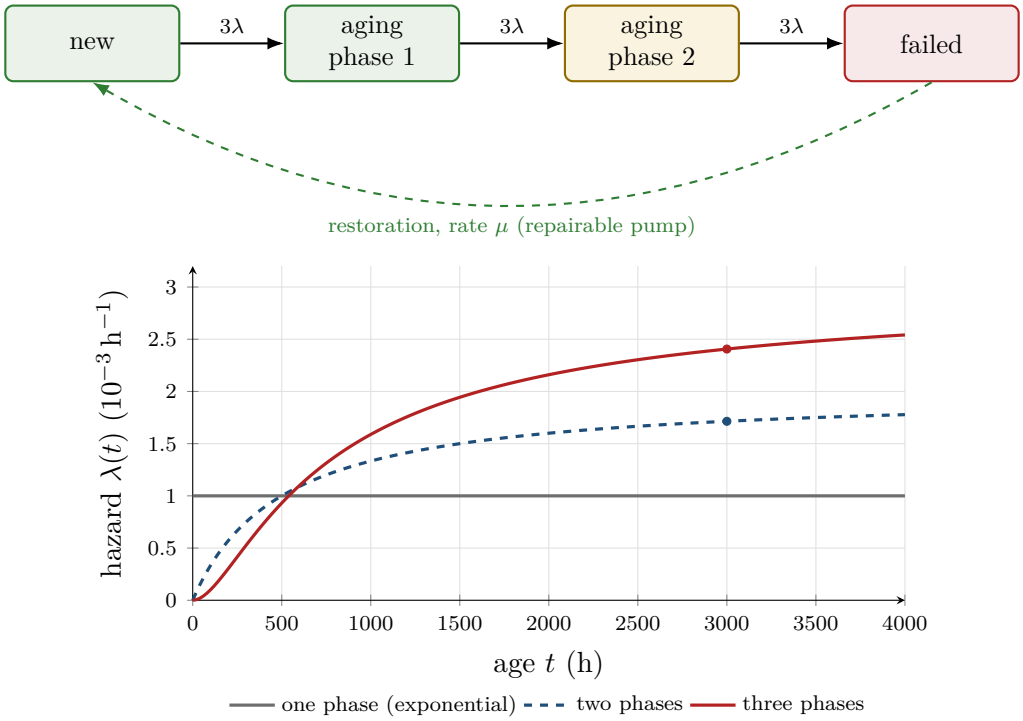


Figure 19.6: Phase-type approximation of a lifetime with increasing hazard. Upper panel: three exponential aging phases of rate 3λ in series, with an optional restoration to the new state. Lower panel: hazard of an Erlang lifetime with k phases of rate $k\lambda$ and mean $1/\lambda = 1000\,\mathrm{h}$: $\lambda(t) = \lambda$ for $k = 1$; $\lambda(t) = \theta^2 t/(1+\theta t)$ with $\theta = 0.002\,\mathrm{h}^{-1}$ for $k = 2$; $\lambda(t) = (\theta^3 t^2/2)/(1+\theta t+\theta^2 t^2/2)$ with $\theta = 0.003\,\mathrm{h}^{-1}$ for $k = 3$; illustrative values. Dots at 3000 h.

19.6.4 Data center power continuity

The same construction applies to the electrical supply of a data center. Table 19.6 lists one illustrative set of states, numbered here as names, not as counts of operating units: utility power available; utility unavailable with the uninterruptible power supply (UPS) carrying the load; generator supplying power; power interrupted; and restoration in progress. The transitions represent utility loss, generator startup, UPS energy depletion, generator failure, utility restoration and transfer-switch operation. The system is available in states 0, 1 and 2.

Table 19.6: Illustrative states of a data center power supply and the transitions that leave them.

State	Meaning	Transitions out of the state
0	utility power available	utility loss, to 1
1	utility unavailable, UPS supplying power	generator startup and transfer, to 2; UPS energy depletion, to 3; utility restoration, to 0
2	generator supplying power	generator failure, to 1 or 3; utility restoration and transfer, to 0
3	power interrupted	start of restoration, to 4
4	restoration in progress	completion, to 0

Not every transition has a constant rate. UPS battery depletion is time dependent: a UPS that has carried the load for eight minutes of a ten-minute autonomy is much closer to depletion than one that has just started, so the time spent in state 1 is not memoryless, and a simple CTMC with a constant depletion rate does not describe a finite battery autonomy accurately. The state must then be expanded with the stored energy of the UPS, $E_{\mathrm{UPS}}(t)$, which evolves by the energy balance of the battery while the UPS carries the load and resets as it recharges, or the autonomy must be represented by a chain of phases as in Section 19.6.3. Either way the reliability model is linked to the physical energy balance: whether state 3 is reached depends on whether the generator starts before the stored energy is exhausted.

19.6.5 Transmission-system availability

A transmission network can be represented by the availability states of its components. For two lines, each available or not,

$$X(t) \in \{(1,1),\ (1,0),\ (0,1),\ (0,0)\}, \tag{19.48}$$

where each binary entry indicates whether the corresponding line is available. Transition rates represent line outages and restorations, and a Markov model gives the probability of each network configuration. Figure 19.7 draws the four states for two lines with the same outage rate λ_ℓ and restoration rate μ_ℓ; each arrow changes the status of one line.

The probability of a configuration is not yet an availability. Whether a configuration can serve the load requires an electrical analysis. Let

$$\phi(\mathbf{x},\mathbf{u}) = \begin{cases} 1, & \text{configuration } \mathbf{x} \text{ meets the operating requirements under conditions } \mathbf{u}, \\ 0, & \text{the requirements are violated}, \end{cases} \tag{19.49}$$

a structure function in the sense of Chapter 9 that depends on the operating conditions $\mathbf{u}$ (demand, dispatch, the weather-dependent ratings of the lines) as well as on the configuration $\mathbf{x}$. Then

$$A_{\mathrm{network}}(t) = \sum_{\mathbf{x}} \Pr\big(X(t)=\mathbf{x}\big)\,\phi(\mathbf{x},\mathbf{u}). \tag{19.50}$$

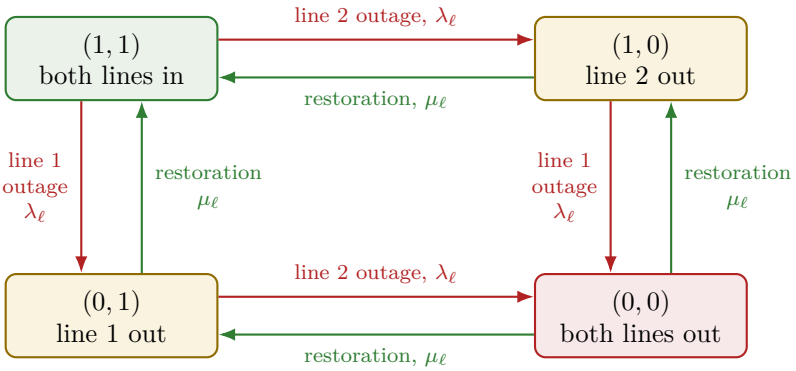


Figure 19.7: Markov model of the configurations of two transmission lines, each with outage rate λ_ℓ and restoration rate μ_ℓ (independent restorations). Whether a configuration serves the load is decided by the network calculation $\phi(\mathbf{x}, \mathbf{u})$ of equation (19.50), not by the diagram.

This connects probabilistic component configurations with power-flow or other physical feasibility calculations. As an illustrative example, take $\lambda_\ell = 0.0005\,\mathrm{h}^{-1}$ and $\mu_\ell = 0.1\,\mathrm{h}^{-1}$ for each line, with independent restorations. Each line is out with steady-state probability $q_\ell = 0.0005/0.1005 = 0.004975$, and the configurations have probabilities 0.990075 for $(1,1)$, 0.004950 each for $(1,0)$ and $(0,1)$, and 0.0000248 for $(0,0)$. If either line can carry the demand, $\phi = 1$ in the first three configurations and $A_{\text{network}} = 0.999975$; if the demand exceeds the rating of one line, as can happen when a hot, still day lowers the ratings, $\phi = 1$ only in $(1,1)$ and $A_{\text{network}} = 0.990075$, an unavailability four hundred times larger from the same component model. This is more informative than labeling the network available whenever at least one line remains in service: actual network adequacy depends on capacity, topology, dispatch, voltage constraints and demand, which a later part of the book on power-system adequacy develops.

19.6.6 Common mistakes and practical checks

Table 19.7 lists the common mistakes in building and reading Markov models, each with the reason it is incorrect. Most of them treat a quantity of the model as a different quantity (a rate as a probability, a point availability as a reliability) or carry an assumption into a system that violates it (unlimited repair, independence, memorylessness, a stationary regime).

A CTMC model is verified by a set of practical checks:

1. all off-diagonal generator entries are nonnegative;
2. every generator row sums to zero;
3. the state probabilities sum to one at all times;
4. all transition rates are in consistent time units;
5. every modeled transition has a defensible physical explanation;
6. the initial conditions represent the actual starting configuration;
7. the limiting behavior agrees with engineering expectation;
8. the results are tested against alternative models, simulation or operational data where possible.

Limiting cases are the most useful tests of the model structure. If $\lambda = 0$ and the system starts fully operational, it must remain in state 2 forever: the first row of $\mathbf{Q}$ is then zero, and $\mathbf{P}(t) = [1,\ 0,\ 0]$ solves the forward equations. If $\mu = 0$ while failures remain possible, the completely failed state becomes absorbing in the two-pump model ($\nu_0 = 0$), and the system eventually becomes unavailable with probability one; the stationary "distribution" is then $[0,\ 0,\ 1]$, which describes certain eventual failure rather than a long-run availability. A second kind of check computes one quantity in two ways. The failure frequency of the example,

Table 19.7: Common mistakes in Markov availability models.

Mistake	Why it is incorrect
Treating transition rates as finite-time probabilities.	A rate has inverse-time units and must be integrated through the state process (Section 19.2.2).
Confusing transient and steady-state availability.	State probabilities evolve from the initial conditions before approaching equilibrium (Section 19.4.3).
Treating availability as first-failure reliability.	A repaired system may be available after earlier failures (Sections 19.3.3 and 19.4.4).
Assuming unlimited repair capacity.	Shared crews or spare-part constraints change the restoration transitions (Section 19.5.3).
Ignoring common-cause failures.	Redundant components may fail together, through a direct transition (Section 19.6.2).
Using exponential rates for aging equipment without justification.	The chosen state may not contain sufficient age information (Section 19.6.3).
Assuming that every available configuration performs the physical function.	Equipment availability does not imply adequate capacity (Section 19.6.5).
Treating the stationary distribution as always existing.	Some models have absorbing states, time-varying rates or several classes of states that do not communicate (Section 19.5.1).

$\pi_1\lambda = 3.8447 \times 10^{-5}\,\mathrm{h}^{-1}$, equals the rate of entries into state 1, $2\lambda\pi_2 + 2\mu\pi_0 = 0.0019608\,\mathrm{h}^{-1}$, multiplied by the probability $\lambda/(\lambda+\mu) = 0.019608$ that an excursion to state 1 ends in state 0; and the unavailability equals $f_\mathrm{F} \cdot \mathrm{MDT}$.

The Markov models of this chapter give availability, failure frequency and down time as functions of the failure rates, the repair rates and the repair resources; Chapter 20 uses them to choose maintenance schedules, crews, spare parts and redundancy.

19.7 What to retain

Definition 19.1 (Markov availability model)**.** A continuous-time Markov model of a repairable system is a finite set of states that contains every variable on which the future transitions depend, and a generator $\mathbf{Q}$ of constant transition rates whose rows sum to zero; the state probabilities solve the Kolmogorov forward equations $\mathrm{d}\mathbf{P}/\mathrm{d}t = \mathbf{P}\mathbf{Q}$, so that $\mathbf{P}(t) = \mathbf{P}(0)\,\mathrm{e}^{\mathbf{Q}t}$, and in steady state $\boldsymbol{\pi}\mathbf{Q} = \mathbf{0}$ with $\sum_i \pi_i = 1$. Availability, failure frequency and mean down time are three distinct quantities obtained from them, related by $1 - A = f_\mathrm{F} \cdot \mathrm{MDT}$.

Markov models represent the evolution of system configurations and account for failures and restorations explicitly. Transient availability depends on the starting state; steady-state availability describes the long-run state distribution when a stationary regime exists. Transition rates are not probabilities. Repair capacity, load sharing, common causes, component age and stored energy are all parts of the model, and each can change the answer as much as the failure rates do. Table 19.8 lists the chapter's terms, their meaning and units, and the subsection that defines each.

Table 19.8: The terms of this chapter and where each is defined.

Term	Meaning	Expression	Units	Defined in
State probability	system in state i at t	$P_i(t) = \Pr(X(t) = i)$	–	19.1.2
Markov property	future depends on present state only	equation (19.6)	–	19.2.1
Transition rate	instantaneous rate from i to j	q_{ij}	h^{-1}	19.2.2
Holding time	time spent in state i per visit	exponential, mean $1/\nu_i$	h	19.2.2
Transient availability	available at t, from a given start	equation (19.15)	–	19.3.2
Steady-state availability	long-run probability of being available	$\mu/(\lambda+\mu)$	–	19.3.4
Generator matrix	matrix of transition rates	$\mathbf{Q}$, rows sum to zero	h^{-1}	19.3.5
Kolmogorov forward equations	state equations	$\mathrm{d}\mathbf{P}/\mathrm{d}t = \mathbf{P}\mathbf{Q}$	h^{-1}	19.3.6
Matrix exponential	transient solution	$\mathbf{P}(0)\,\mathrm{e}^{\mathbf{Q}t}$	–	19.3.6
Birth–death process	transitions between neighboring states only	Figure 19.2	–	19.4.1
Absorbing state	state that cannot be left	$\nu_i = 0$	–	19.4.4
Stationary distribution	state probabilities left unchanged	$\boldsymbol{\pi}\mathbf{Q} = \mathbf{0}$, $\sum_i \pi_i = 1$	–	19.5.1
Redundant pair, two crews	steady-state availability	$1 - [\lambda/(\lambda+\mu)]^2$	–	19.5.1
Expected annual downtime	hours unavailable per year	$8760\,(1 - A_{\mathrm{sys}})$	h/yr	19.5.4
Failure frequency	rate of entering the failed states	$f_{\mathrm{F}} = \pi_1 \lambda$	h^{-1}, yr^{-1}	19.5.5
Mean down time	mean duration of one outage	$\mathrm{MDT} = 1/q_{01}$	h	19.5.5
Frequency–duration relation	unavailability from both	$1 - A = f_{\mathrm{F}} \cdot \mathrm{MDT}$	–	19.5.5
Phase-type distribution	lifetime as a chain of exponential phases	Figure 19.6	h	19.6.3
Network availability	configuration probability times feasibility	equation (19.50)	–	19.6.5

Further reading. Rausand, Barros and Høyland, *System Reliability Theory*, third edition, for Markov models of repairable and redundant systems; Ross, *Introduction to Probability Models*, for continuous-time Markov chains and stationary distributions; Modarres, Kaminskiy and Krivtsov, *Reliability Engineering and Risk Analysis*, for Markov reliability and availability models.

19.8 Exercises

Exercise 19.1. (a) State what the Markov property requires. (b) State the two properties that the entries of a generator matrix must satisfy, and decide whether $\left[\begin{smallmatrix} -0.01 & 0.01 \\ 0.2 & -0.1 \end{smallmatrix}\right] \mathrm{h}^{-1}$ is a valid generator. (c) State what transient availability is and how it differs from reliability. (d) For one pump with $\lambda = 0.001\,\mathrm{h}^{-1}$ and $\mu = 0.05\,\mathrm{h}^{-1}$, operating at $t = 0$, compute $A(50\,\mathrm{h})$ and $R(50\,\mathrm{h})$. (Answers: (a) that the future be conditionally independent of the earlier history given the current state, so that the chosen state contains all the information needed to predict the future under the model. (b) Off-diagonal entries nonnegative and every row summing to zero, the diagonal entry being minus the total departure rate; the matrix is not valid, because its second row sums to 0.1, and its diagonal entry should be -0.2. (c) The probability that the system is operational at a specified time, which permits failures and restorations before that time; reliability is the probability that no failure has occurred at all. (d) $A(50\,\mathrm{h}) = 0.980392 + 0.019608\,\mathrm{e}^{-2.55} = 0.981923$ and $R(50\,\mathrm{h}) = \mathrm{e}^{-0.05} = 0.951229$.)

Exercise 19.2. Two identical redundant pumps have $\lambda = 0.0003\,\mathrm{h}^{-1}$ and $\mu = 0.01\,\mathrm{h}^{-1}$ (illustrative). (a) For two independent repair processes and for one shared crew, compute the stationary distribution, the steady-state availability, the failure frequency per year, the mean down time and the expected annual downtime. (b) Explain why one repair crew reduces the availability of the pair. (c) State which quantity gives the steady-state frequency of transitions into the fully failed state, and why. (Answers: (a) $\rho = 0.03$. Two crews: $\pi_2 = 1/1.0609 = 0.942596$, $\pi_1 = 0.056556$, $\pi_0 = 0.00084834$, $A_{\mathrm{sys}} = 0.999152$, $f_{\mathrm{F}} = 0.1486$ per year, MDT $= 50\,\mathrm{h}$, 7.43 h per year. One crew: $\pi_2 = 1/1.0618 = 0.941797$, $\pi_1 = 0.056508$, $\pi_0 = 0.0016952$, $A_{\mathrm{sys}} = 0.998305$, $f_{\mathrm{F}} = 0.1485$ per year, MDT $= 100\,\mathrm{h}$, 14.85 h per year. (b) Both failed pumps cannot be repaired at the same time, so restoration out of the fully failed state is slower; the failure rates are unchanged and the redundancy remains. (c) $\pi_1\lambda$, because the model enters state 0 only from state 1, at rate λ.)

Exercise 19.3. A pump's lifetime is represented by two exponential phases, new and aged, each left at rate $0.002\,\mathrm{h}^{-1}$, so that the mean lifetime is 1000 h; a failed pump is repaired at $\mu = 0.05\,\mathrm{h}^{-1}$ and returns to the new phase. (a) Explain why a model that records only whether the pump is operating would not be Markovian if the lifetime had an increasing hazard, and state what should be done when future failure rates depend on accumulated age. (b) Write the generator with the states ordered (new, aged, failed). (c) Compute the stationary distribution and the steady-state availability, and compare with the exponential pump of Section 19.3.4. (d) The transient availability of this model at 24 h, starting new, is 0.999226; compare it with the exponential pump and explain the difference. (Answers: (a) the operating state does not record the age on which the hazard depends, so it does not contain the information needed to predict the future; expand the state, for example with phases or virtual ages, or use a more general stochastic model, rather than assume memorylessness. (b) Rows $[-0.002,\ 0.002,\ 0]$, $[0,\ -0.002,\ 0.002]$, $[0.05,\ 0,\ -0.05]$. (c) Balance gives $\pi_{\mathrm{new}} = \pi_{\mathrm{aged}}$ and $0.002\,\pi_{\mathrm{aged}} = 0.05\,\pi_{\mathrm{failed}}$, so $\boldsymbol{\pi} = [0.490196,\ 0.490196,\ 0.019608]$ and $A = 0.980392$, identical to the exponential pump, because the steady-state availability depends only on the mean up and down times. (d) 0.999226 against 0.986158: a new pump must pass through two phases before it can fail, so early failures are rarer than under the exponential model.)

Exercise 19.4. The cooling plant has three identical chillers, of which two are required. Each chiller has MTTF $= 2000\,\mathrm{h}$ and MTTR $= 24\,\mathrm{h}$ (illustrative). (a) With the states counting the operating chillers, 3, 2, 1, 0, draw the birth–death model and give its rates for three independent crews and for one crew. (b) For three crews, compute the stationary distribution and the plant availability, and verify the availability with the binomial formula for a 2-out-of-3 system of independent components. (c) For both crew arrangements compute the failure frequency of the plant per year, its mean down time and its expected annual downtime. (d) Explain why the

mean down time with three crews is not exactly $1/(2\mu) = 12\,\mathrm{h}$. (Answers: (a) failures 3λ, 2λ, λ from 3, 2, 1 with $\lambda = 0.0005\,\mathrm{h}^{-1}$; repairs from 2, 1, 0 at μ, 2μ, 3μ with three crews and at μ from every state with one crew, $\mu = 1/24\,\mathrm{h}^{-1}$. (b) Boundary ratios $3\lambda/\mu = 0.036$, $\lambda/\mu = 0.012$ and $\lambda/(3\mu) = 0.004$ give $\boldsymbol{\pi} = [0.964847,\ 0.034734,\ 0.00041681,\ 0.0000016673]$ for states 3, 2, 1, 0, and $A_{\mathrm{plant}} = 0.999582$; with $q = \lambda/(\lambda+\mu) = 0.011858$, $(1-q)^3 + 3q(1-q)^2 = 0.999582$. (c) The plant fails only by the transition from 2 to 1, so $f_{\mathrm{F}} = 2\lambda\,\pi_2$: three crews, 0.3043 per year, MDT $= (\pi_1 + \pi_0)/f_{\mathrm{F}} = 12.05\,\mathrm{h}$, 3.67 h per year; one crew (ratios 0.036, 0.024, 0.012), $A_{\mathrm{plant}} = 0.999157$, 0.3041 per year, 24.29 h, 7.39 h per year. (d) The failed set contains two states; a few outages pass from state 1 to state 0 before a repair is completed, which lengthens the mean.)

Chapter 20

Availability and Maintenance Optimization

20.1 Availability as an engineering objective

Chapters 17, 18 and 19 developed the mathematics of repairable systems: counting processes for recurrent failures, renewal processes and imperfect repair, and continuous-time Markov models of the transitions between operating and failed states. This chapter uses those models to make engineering decisions. A data center has two redundant cooling-water pumps, either of which can supply the required flow, and the operator is considering four improvements: hiring a second maintenance crew, keeping a replacement pump or its critical spare parts on site, increasing preventive maintenance, and installing additional redundant pumping capacity. The question of this chapter is

how redundancy, maintenance crews, spare parts and preventive maintenance should be allocated to achieve a required availability of a system function at an acceptable cost.

Which improvement gives the largest gain in availability, and which one is economically justified, depends on more than the equipment failure rates. It depends on the repair time, the repair resources, the maintenance schedule, the spare-parts logistics, the redundancy, the operational consequences of an outage, and the uncertainty in every one of these inputs. The chapter defines the availability measures that a decision can target, builds the availability model of the cooling-water pumps, adds repair crews and spare parts to it, compares five maintenance strategies by availability and by expected annual cost, treats maintenance scheduling and spare-parts inventory as reliability problems, and ends with the effect of uncertainty and model adequacy on the decision.

Jargon: point, average and steady-state availability, expected downtime, corrective and preventive maintenance, repair crew, logistic delay, spare part, stockout, service level

Point availability $A(t)$ is the probability that the required function is available at time t. *Average availability* $\bar{A}(H)$ is the expected fraction of an assessment period $[0, H]$ during which it is available. *Steady-state availability* A is the limit of $A(t)$ as $t \to \infty$, when the limit exists. All three are dimensionless. *Expected downtime* is the expected total time, in hours, during which the function is unavailable over a stated period. *Corrective maintenance* is performed after a failure to restore the item; *preventive maintenance* is performed before failure to reduce the probability or the effects of failure. A *repair crew* is the set of people and equipment that can carry out one repair at a time. A *logistic delay* is the part of a restoration spent waiting for resources (people, parts, permits, tools) rather than repairing. A *spare part* is a replacement item held in stock; a *stockout* occurs when a demand for a spare finds none in stock; the *service level* of a stock is the probability that demand during a replenishment period does not exceed the stock.

Key idea

Availability is a property of the whole failure-and-restoration architecture, not of the equipment alone, and the design that maximizes availability is in general not the design that minimizes expected cost. A maintenance decision needs an availability model, a cost of downtime, the costs of the candidate programs, and an explicit statement of the objective and of any availability requirement.

This section defines the four measures that a maintenance decision can target and explains why the operating objective of a continuously running plant is usually stated as an availability.

20.1.1 Four measures of a repairable system

Chapter 3 defined reliability, point availability and steady-state availability, and introduced the average availability in a remark. Table 20.1 collects the four measures with the engineering question each answers. They are restated here for a system rather than a component, because the decisions of this chapter concern a system function.

Let the system start in a state in which its required function is available, at $t = 0$, and let T be the first time at which the function is lost. The *first-failure reliability* is

$$R(t) = \Pr(T > t), \tag{20.1}$$

the probability that the function survives $[0, t]$ without any loss. A restoration after the first loss does not change this probability: a mission that was interrupted was not survived. To describe a system that is restored, define the binary system state

$$Y(t) = \begin{cases} 1, & \text{the required function is available at } t, \\ 0, & \text{the required function is unavailable at } t. \end{cases} \tag{20.2}$$

The *point availability* is the probability that the function is available at the instant t,

$$A(t) = \Pr[Y(t) = 1] = \mathbb{E}[Y(t)], \tag{20.3}$$

whatever happened before t; the system may have failed and been restored any number of times. Since every history without a loss in $[0, t]$ ends with $Y(t) = 1$, $A(t) \geq R(t)$ for every t, as in equation (3.9). Over an assessment horizon H, the *average availability* is the time average of the point availability,

$$\bar{A}(H) = \frac{1}{H} \int_0^H A(t)\, \mathrm{d}t. \tag{20.4}$$

It is the expected fraction of $[0, H]$ spent available. The total downtime in the horizon is the integral of the indicator of unavailability, $D_H = \int_0^H [1 - Y(t)]\,\mathrm{d}t$, and by linearity of the expectation

$$\mathbb{E}[D_H] = \int_0^H [1 - A(t)]\,\mathrm{d}t = H\,[1 - \bar{A}(H)], \tag{20.5}$$

in hours. Finally, when the limit exists, the *steady-state availability* is

$$A = \lim_{t \to \infty} A(t). \tag{20.6}$$

For a system whose up and down periods form an alternating renewal process, or more generally a regenerative process with cycles of mean up time $\mathbb{E}[U]$ and mean down time $\mathbb{E}[D]$, the limit is the standard result of Section 3.3.2,

$$A = \frac{\mathbb{E}[U]}{\mathbb{E}[U] + \mathbb{E}[D]}. \tag{20.7}$$

This formula requires a stationary or regenerative model; it is not a universal expression for every complex system, and a system with an absorbing failed state or with time-varying rates has no such limit.

Table 20.1: Four measures of a repairable system and the questions they answer. The last column evaluates each for one repairable pump of the two-state model with $\lambda = 0.001\,\mathrm{h}^{-1}$ and $\mu = 0.05\,\mathrm{h}^{-1}$, up at $t = 0$, at $t = H = 100\,\mathrm{h}$ (illustrative values).

Measure	Symbol	Engineering question	One pump, 100 h
First-failure reliability	$R(t)$	Will the function be maintained without any loss throughout the mission?	0.904837
Point availability	$A(t)$	Is the function available at time t?	0.980512
Average availability	$\bar{A}(H)$	What fraction of the assessment period is the function expected to be available?	0.984213
Steady-state availability	A	What is the long-run probability that the function is available?	0.980392

For the two-state model of one repairable item (Section 3.3.1), with failure rate λ and repair rate μ, the four measures have closed forms. The first-failure reliability is $R(t) = \mathrm{e}^{-\lambda t}$; the point availability of an item up at $t = 0$ is equation (3.11),

$$A(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu}\,\mathrm{e}^{-(\lambda+\mu)t}; \tag{20.8}$$

integrating it over $[0, H]$ and dividing by H,

$$\bar{A}(H) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{(\lambda + \mu)^2\,H}\left[1 - \mathrm{e}^{-(\lambda+\mu)H}\right]; \tag{20.9}$$

and the steady-state availability is $A = \mu/(\lambda+\mu)$. For one cooling-water pump with $\lambda = 0.001\,\mathrm{h}^{-1}$ (MTTF $= 1000\,\mathrm{h}$) and $\mu = 0.05\,\mathrm{h}^{-1}$ (MTTR $= 20\,\mathrm{h}$), $A = 0.05/0.051 = 0.980392$ and the time constant of the transient is $1/(\lambda + \mu) = 19.61\,\mathrm{h}$. At $H = 100\,\mathrm{h}$,

$$R(100\,\mathrm{h}) = \mathrm{e}^{-0.1} = 0.904837,$$

$$A(100\,\mathrm{h}) = 0.980392 + 0.019608\,\mathrm{e}^{-5.1} = 0.980392 + 0.000120 = 0.980512,$$

$$\bar{A}(100\,\mathrm{h}) = 0.980392 + \frac{0.001}{0.051^2 \times 100}\,(1 - \mathrm{e}^{-5.1}) = 0.980392 + 0.003845 \times 0.993903 = 0.984213,$$

which are the values of the last column of Table 20.1.

Figure 20.1 plots the four measures of this pump against time. The horizontal axis of each panel is the time since the pump was put into service in the up state, in hours; the vertical axis is a probability. Panel (a) spans 1000 h and the full range in which the reliability falls; panel (b) magnifies the first 150 h and the band between 0.975 and 1, where the three availability measures differ. The solid red curve is the first-failure reliability, the solid blue curve the point availability, the dash-dotted green curve the average availability, and the dashed gray line the steady-state availability. The reliability falls without limit, to 0.368 at one MTTF; the point availability falls within a few time constants to the steady-state value and stays there; the average availability lies above the point availability, because it averages the early hours in which the pump is almost certainly up, and approaches A only slowly, as $1/H$. In panel (b) the reliability leaves the magnified band after about 25 h, while the three availability measures remain within two percentage points of 1.

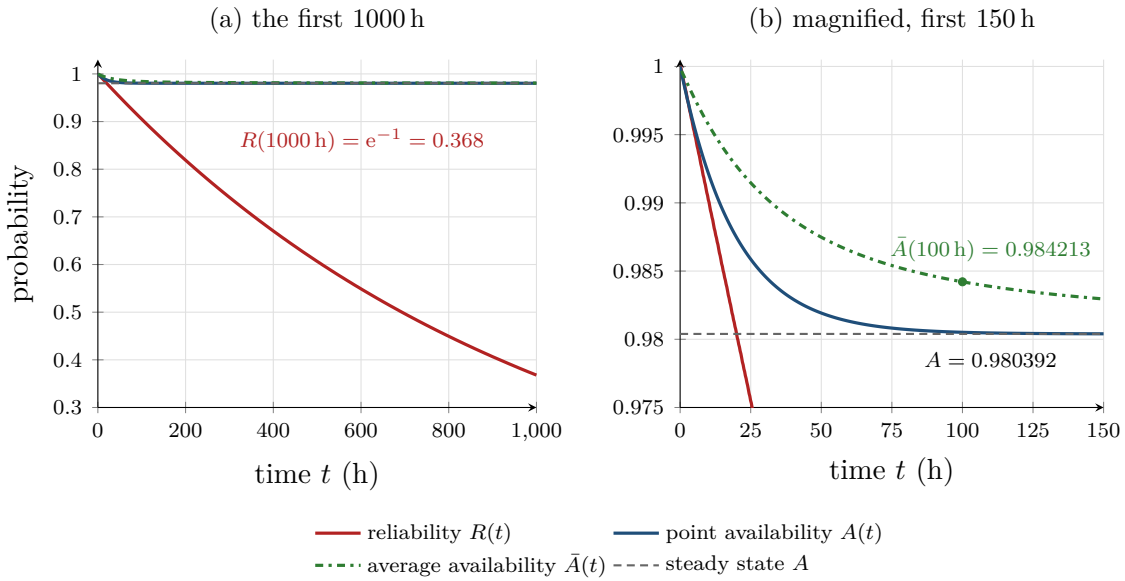


Figure 20.1: The four measures of one repairable pump of the two-state model, up at $t = 0$, with $\lambda = 0.001\,\mathrm{h}^{-1}$ and $\mu = 0.05\,\mathrm{h}^{-1}$ (illustrative values): first-failure reliability $R(t) = \mathrm{e}^{-\lambda t}$; point availability $A(t)$ of equation (20.8); average availability $\bar{A}(t)$ of equation (20.9) over the horizon $[0, t]$; and steady-state availability $A = \mu/(\lambda + \mu) = 0.980392$. Panel (b) magnifies the band near 1.

20.1.2 Availability as the operational objective

A data center cooling plant operates continuously. An operator can tolerate occasional equipment faults, provided redundancy and repair procedures prevent a loss of the required cooling. The engineering objective is therefore usually not to eliminate every component failure, which neither redundancy nor maintenance can do, but to keep the probability

$$\Pr\big(C_{\text{available}}(t) \geq L(t)\big) \tag{20.10}$$

acceptably high, where $C_{\text{available}}(t)$ is the available cooling capacity and $L(t)$ the required cooling load at time t. This is an availability of the cooling function: it allows failures, provided the capacity that remains, or that is restored in time, still meets the load. First-failure reliability remains the relevant measure only when the first interruption itself defeats the purpose, as in a mission without access for repair.

The probability (20.10) depends on two distinct properties of the design. *Reliability* governs how frequently equipment fails. *Maintainability* governs how quickly the system recovers when it does. Redundancy maintains the function during component outages; maintenance and repair logistics determine how long equipment remains degraded or unavailable, and therefore how long the system depends on the redundancy that remains. The decisions of this chapter act on both properties: preventive maintenance acts on the failure rate, while crews and spare parts act on the restoration rate.

20.2 The cooling-plant availability model and repair capacity

This section defines the pumping plant, restates the three-state Markov model of Chapter 19 with its numerical results, interprets the availability in hours and outages per year, and evaluates the first candidate improvement, a second repair crew.

20.2.1 The pumping plant

The system is the cooling-water pumping branch of a data center cooling plant, with two identical pumps in parallel. Either pump alone can supply the flow required for a cooling duty of 900 kW under the selected operating conditions, so the pumping function is available when at least one pump operates. The chillers, cooling towers, power supply and controls of the plant are outside the system boundary in this chapter and are assumed adequate; Section 20.6.4 returns to that assumption.

Figure 20.2 draws the reliability block diagram. The block on the left is the supply manifold, the two blocks in the middle are the pumps, each labeled with the duty it can carry alone, and the block on the right is the cooling service that receives the flow. The two parallel paths mean that the function survives the loss of either pump (Chapter 9); it is lost only when both pumps are down at the same time. The manifold and the service are drawn for orientation and are not modeled as failing.

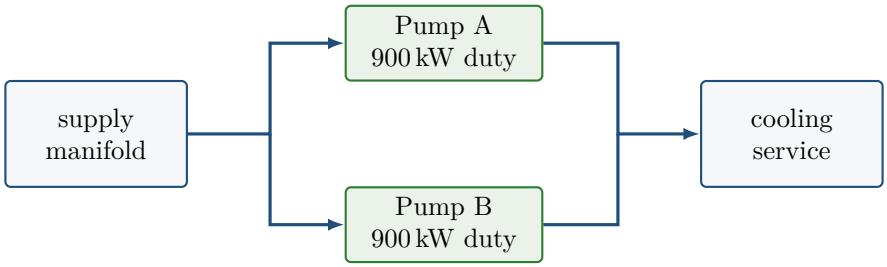


Figure 20.2: Reliability block diagram of the pumping plant: two identical pumps in parallel between the supply manifold and the cooling service. The pumping function is available when at least one pump operates.

Each operating pump fails at the constant rate

$$\lambda = 0.001\,\mathrm{h}^{-1}, \qquad \mathrm{MTTF} = \frac{1}{\lambda} = 1000\,\mathrm{h},$$

and the mean time to restore a failed pump, from failure to return to service, is $\mathrm{MTTR} = 20\,\mathrm{h}$. Under an exponential repair-time assumption the repair rate is

$$\mu = \frac{1}{\mathrm{MTTR}} = \frac{1}{20\,\mathrm{h}} = 0.05\,\mathrm{h}^{-1}.$$

Initially the plant has one repair crew, which can work on one pump at a time. These failure rates are deliberately high, so that the differences among maintenance strategies are visible in a few decimal places; they are illustrative and are not representative of any particular commercial pump. The failures of the two pumps are assumed independent, which excludes the common-cause failures of Chapter 13, and a pump's failure rate is assumed not to change when its partner fails.

20.2.2 The three-state Markov model

Chapter 19 modeled this plant as a continuous-time Markov model with three states, indexed by the number of operating pumps: state 2, both pumps operating; state 1, one operating and one failed; state 0, both failed. The pumping function is available in states 2 and 1 and unavailable in state 0. Figure 20.3 draws the model. Each box is a state, colored green when the function is available with redundancy, gold when it is available through one pump, and red when it is not. The straight arrows between the boxes are the failure transitions and the curved arrows below the boxes are the repair transitions, each labeled with its rate. From state 2 either of two operating pumps can fail, so the rate to state 1 is 2λ; from state 1 the remaining pump fails at rate λ and the failed pump is repaired at rate μ. The rate from state 0 to state 1 depends on the repair resources: μ when one crew must repair the two pumps one after the other, 2μ when two repairs can progress at the same time.

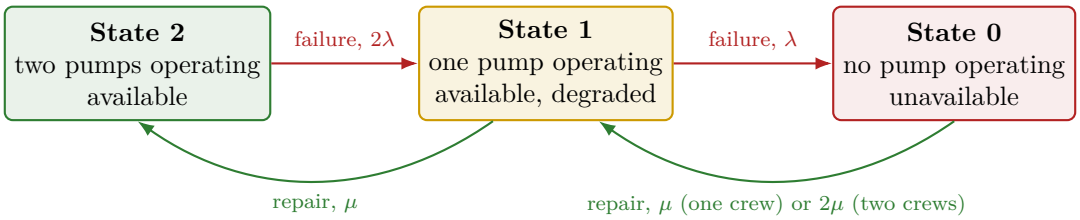


Figure 20.3: Three-state Markov model of the two-pump plant. Failure transitions (straight arrows) occur at 2λ and λ; repair transitions (curved arrows) at μ, and from the fully failed state at μ with one repair crew or 2μ with two independent repairs.

With one crew and the states ordered $(2, 1, 0)$, the generator matrix is

$$\mathbf{Q} = \begin{bmatrix} -2\lambda & 2\lambda & 0 \\ \mu & -(\lambda+\mu) & \lambda \\ 0 & \mu & -\mu \end{bmatrix}. \tag{20.11}$$

Each off-diagonal entry is a transition rate and each row sums to zero. The stationary distribution $\boldsymbol{\pi} = (\pi_2, \pi_1, \pi_0)$ solves $\boldsymbol{\pi}\mathbf{Q} = \mathbf{0}$ with $\pi_2 + \pi_1 + \pi_0 = 1$. Because the chain moves only between neighboring states, the stationary equations reduce to a balance of probability flow across each cut between neighbors: the flow from state 2 to state 1 equals the flow back, and likewise between states 1 and 0,

$$2\lambda\,\pi_2 = \mu\,\pi_1, \qquad \lambda\,\pi_1 = \mu\,\pi_0.$$

Writing $\rho = \lambda/\mu$ for the ratio of the failure rate to the repair rate, the first gives $\pi_1 = 2\rho\,\pi_2$ and the second $\pi_0 = \rho\,\pi_1 = 2\rho^2\pi_2$. Normalization gives $\pi_2\,(1 + 2\rho + 2\rho^2) = 1$, so

$$\pi_2 = \frac{1}{1+2\rho+2\rho^2}, \qquad \pi_1 = \frac{2\rho}{1+2\rho+2\rho^2}, \qquad \pi_0 = \frac{2\rho^2}{1+2\rho+2\rho^2}, \tag{20.12}$$

and since the plant is unavailable only in state 0,

$$A = 1 - \pi_0. \tag{20.13}$$

For the plant, $\rho = 0.001/0.05 = 0.02$ and $1 + 2\rho + 2\rho^2 = 1 + 0.04 + 0.0008 = 1.0408$, so

$$\pi_2 = 0.960799, \qquad \pi_1 = \frac{0.04}{1.0408} = 0.038432, \qquad \pi_0 = \frac{0.0008}{1.0408} = 0.00076864,$$

and the steady-state availability is $A = 0.99923136$, or $99.92314\,\%$. These are the one-crew results of Chapter 19.

The number of crews does not affect the first-failure reliability of the plant, because the transition out of state 0 plays no role until the function has already been lost once. With state 0 made absorbing, let m_2 and m_1 be the mean times to reach state 0 from states 2 and 1. One step of the chain from each state gives

$$m_2 = \frac{1}{2\lambda} + m_1, \qquad m_1 = \frac{1}{\lambda + \mu} + \frac{\mu}{\lambda + \mu}\, m_2,$$

since the chain leaves state 2 after a mean $1/(2\lambda)$ and always goes to state 1, and leaves state 1 after a mean $1/(\lambda + \mu)$, returning to state 2 with probability $\mu/(\lambda + \mu)$. Substituting the first into the second and solving,

$$m_2 = \frac{3\lambda + \mu}{2\lambda^2}, \qquad m_1 = m_2 - \frac{1}{2\lambda}. \tag{20.14}$$

For the plant, the mean time to the first loss of the pumping function from the new condition is $m_2 = 0.053/(2 \times 10^{-6}) = 26\,500\,\mathrm{h}$, and the probability of a year of 8760 h without any loss, from the matrix exponential of the absorbing chain, is 0.7189. A year without any outage is therefore far from certain, although the plant is available $99.92\,\%$ of the time.

20.2.3 Operational meaning of the availability

An availability of $99.92314\,\%$ is easier to judge as hours and outages per year. Assume 8760 h per year and stationary operation. By equation (20.5) with $A(t) = A$, the expected downtime per year is

$$\mathbb{E}[D_{\mathrm{yr}}] = 8760\,\mathrm{h} \times (1 - A) = 8760 \times 0.00076864 = 6.733\,\mathrm{h}. \tag{20.15}$$

The expected number of complete pumping outages per year is the failure frequency of the system, f_{F}. The plant enters state 0 only from state 1, at rate λ while state 1 is occupied; in the stationary regime, the expected number of such transitions per hour is $\pi_1\lambda$, and per year

$$f_{\mathrm{F}} = 8760\,\lambda\,\pi_1 = 8760 \times 0.001 \times 0.038432 = 0.3367\,\mathrm{yr}^{-1}, \tag{20.16}$$

about one complete outage every three years. With one crew, the plant leaves state 0 at rate μ, so the mean down time per outage is

$$\mathrm{MDT} = \frac{1}{\mu} = 20\,\mathrm{h}.$$

The three quantities are linked. The second balance equation, $\lambda\pi_1 = \mu\pi_0$, says that the stationary frequency of entries into state 0 equals the frequency of exits from it, and dividing by μ gives $\pi_0 = \lambda\pi_1 \cdot (1/\mu)$; multiplying by 8760 h,

$$\mathbb{E}[D_{\mathrm{yr}}] = f_{\mathrm{F}} \cdot \mathrm{MDT} = 0.3367 \times 20\,\mathrm{h} = 6.733\,\mathrm{h}. \tag{20.17}$$

The relation holds for any stationary regenerative system: the long-run fraction of time in a set of states equals the frequency of entering it times the mean duration of a visit. The same numbers satisfy equation (20.7) at the system level. After each outage the plant restarts in state 1, so its mean up time per cycle is m_1 of equation (20.14), $26\,500 - 500 = 26\,000\,\mathrm{h}$, its mean down time is 20 h, and $A = 26\,000/(26\,000 + 20) = 0.99923136$; the cycle length 26 020 h is $8760/f_{\mathrm{F}}$.

The plant therefore experiences complete pumping outages rarely, but each lasts 20 h on average. The duration, not the frequency, dominates the expected downtime, which suggests that the restoration process is where an improvement can be bought.

20.2.4 A second repair crew

Suppose a second crew is hired, so that two failed pumps can be repaired at the same time, independently, each at rate μ. When both pumps are down, either repair may finish first, and the rate from state 0 to state 1 changes from μ to 2μ; no other rate changes. The generator becomes

$$\mathbf{Q}_2 = \begin{bmatrix} -2\lambda & 2\lambda & 0 \\ \mu & -(\lambda+\mu) & \lambda \\ 0 & 2\mu & -2\mu \end{bmatrix}. \tag{20.18}$$

The first balance equation is unchanged, $\pi_1 = 2\rho\,\pi_2$; the second becomes $\lambda\pi_1 = 2\mu\pi_0$, so $\pi_0 = (\rho/2)\pi_1 = \rho^2\pi_2$, and normalization gives $\pi_2(1+\rho)^2 = 1$. The stationary distribution is binomial,

$$\pi_2 = \Big(\frac{\mu}{\lambda+\mu}\Big)^2, \qquad \pi_1 = \frac{2\lambda\mu}{(\lambda+\mu)^2}, \qquad \pi_0 = \Big(\frac{\lambda}{\lambda+\mu}\Big)^2,$$

which is the result for two independent repairable pumps, each with availability $\mu/(\lambda+\mu)$, in parallel. The steady-state unavailability is

$$1 - A = \Big(\frac{\lambda}{\lambda+\mu}\Big)^2 = \Big(\frac{0.001}{0.051}\Big)^2 = 0.00038447, \tag{20.19}$$

so $A = 99.96155\,\%$, and the expected downtime is $8760 \times 0.00038447 = 3.368\,\mathrm{h}$ per year. The second crew reduces the expected downtime by $6.733 - 3.368 = 3.365\,\mathrm{h}$ per year, a reduction of $49.98\,\%$, close to one half.

The reason is in the restoration of the fully failed state. With one crew, the first repair completion from state 0 takes a mean of $1/\mu = 20\,\mathrm{h}$; with two independent repairs in progress, the first of the two completions takes a mean of $1/(2\mu) = 10\,\mathrm{h}$. The mean down time per outage halves. The failure frequency hardly changes: $\pi_1 = 0.038447$ gives $f_\mathrm{F} = 8760 \times 0.001 \times 0.038447 = 0.3368\,\mathrm{yr}^{-1}$, against $0.3367\,\mathrm{yr}^{-1}$ with one crew, and the first-failure reliability of equation (20.14) does not change at all. The second crew does not reduce the failure rate of either pump; the whole improvement comes from faster restoration of the system function.

> **Principle 20.1.** *Availability depends on the complete failure-and-restoration architecture. Two systems with identical equipment and identical failure rates can differ by a factor of two in unavailability because they differ in the repair resources that act on the fully failed state.*

20.2.5 Repair capacity as a constrained resource

The rate 2μ in equation (20.18) assumes that two repairs can genuinely progress independently. A real maintenance operation has a limited set of shared resources: one electrical technician and one mechanical technician, a limited set of tools, one crane or lifting platform, one replacement motor, and specialist contractors shared with other sites. The transition rate from the fully failed state to the partially restored state then depends on which resources are available, not on the number of people alone. If both repairs need the same crane, a second technician adds little: the repairs still proceed one at a time through the crane, and the rate out of state 0 stays close to μ.

Repair capacity is therefore a resource-constrained system in its own right. A maintenance model that credits each additional technician with a proportional increase in the repair rate overstates the benefit of staffing whenever another resource is the bottleneck. The model must represent the actual resource constraints, either by the rate assigned to each transition or, when the constraints interact, by additional states that record which resources are busy.

20.3 Spare parts and repair logistics

The mean repair time of 20 h used so far is an aggregate of several delays of different kinds. This section decomposes it, derives the expected repair duration as a function of the availability of spare parts, states the approximation needed to keep the three-state model, and recomputes the plant availability with stocked spares.

20.3.1 Decomposition of the repair time

A measured mean time to repair usually includes several distinct delays. The restoration duration T_{R} of one failure can be written as the sum

$$T_{\mathrm{R}} = D_{\mathrm{diagnosis}} + D_{\mathrm{waiting}} + D_{\mathrm{repair}} + D_{\mathrm{testing}}, \tag{20.20}$$

in hours, of the time to detect and diagnose the failure, the time waiting for resources, the time of hands-on repair and the time to test and restart, as in the restoration phases of Section 3.4.2. The waiting time includes the availability of a technician, the delivery of spare parts, the issue of required permits, the availability of specialized equipment and the mobilization of contractors. Its mean adds to the MTTR in the same way as the repair itself.

The useful engineering distinction is between *active repair time*, the terms D_{repair} and, by most definitions, D_{testing}, and *logistic delay*, the term D_{waiting}. An asset can be mechanically repairable within four hours and remain unavailable for several days because a replacement part is not on site. The two kinds of delay are reduced by different investments: active repair time by design for maintainability and by training, logistic delay by staffing, stocking and contracting.

20.3.2 Expected repair duration with spare-parts availability

Suppose a pump repair needs a spare part, and that the repair duration depends on whether the part is in stock when the repair demand arises. Let p_{S} be the probability that the spare is in stock at a repair demand. If it is, the mean repair duration is $\mathbb{E}[T_{\mathrm{R}} \mid \text{in stock}] = 8\,\mathrm{h}$; if it is not, the repair waits for delivery and the mean is $\mathbb{E}[T_{\mathrm{R}} \mid \text{not in stock}] = 48\,\mathrm{h}$. The two events partition the repair demands, so the law of total expectation gives the mean of the mixture,

$$\mathbb{E}[T_{\mathrm{R}}] = p_{\mathrm{S}}\,\mathbb{E}[T_{\mathrm{R}} \mid \text{in stock}] + (1 - p_{\mathrm{S}})\,\mathbb{E}[T_{\mathrm{R}} \mid \text{not in stock}]. \tag{20.21}$$

With $p_{\mathrm{S}} = 0.70$,

$$\mathbb{E}[T_{\mathrm{R}}] = 0.70 \times 8\,\mathrm{h} + 0.30 \times 48\,\mathrm{h} = 5.6\,\mathrm{h} + 14.4\,\mathrm{h} = 20\,\mathrm{h},$$

which is the plant's present MTTR. Of these 20 h, $0.30 \times (48 - 8) = 12\,\mathrm{h}$, or 60 %, is waiting for parts in the 30 % of repairs that find none in stock. A stocking program that raises p_{S} to 1 makes the mean repair time 8 h, a reduction of $(20 - 8)/20 = 60\,\%$. Equation (20.21) is linear in p_{S}, so every percentage point of spare availability removes $(48 - 8)/100 = 0.4\,\mathrm{h}$ from the mean repair time.

20.3.3 The exponential approximation of the repair time

A mean repair time of 8 h does not imply that the repair duration is exponentially distributed. A repair that involves diagnosis, delivery, installation and testing is a sum of phases, and the mixture of equation (20.21) has two modes, so its distribution is in general far from exponential. To keep the three-state Markov model, the repair duration is replaced by an exponential duration with the same mean, which is the *exponential repair-time approximation*,

$$\mu_{\mathrm{eff}} = \frac{1}{\mathbb{E}[T_{\mathrm{R}}]}. \tag{20.22}$$

The approximation matches the mean repair time but not the variance or the higher moments. As an illustration, if the durations within each branch of the mixture were themselves exponential, with means 8 and 48 h, the second moment would be $\mathbb{E}[T_{\mathrm{R}}^2] = 0.70 \times 2 \times 8^2 + 0.30 \times 2 \times 48^2 = 89.6 + 1382.4 = 1472\,\mathrm{h}^2$, the variance $1472 - 20^2 = 1072\,\mathrm{h}^2$, and the standard deviation 32.7 h, against 20 h for the exponential duration of the same mean. The approximation understates the probability of a long outage. For the steady-state availability of the plant the error is limited, because a long-run fraction of time depends mainly on means; for the probability of an outage longer than a stated duration, it is not.

A more accurate model introduces explicit repair phases (diagnosis, waiting, repair, testing, each exponential, so that their sum is a phase-type distribution) or explicit states of the spare-parts inventory, as Section 20.5.6 does. The comparison of strategies below uses the effective-rate approximation, with its limitations kept in view.

20.3.4 Availability with stocked spare parts

With spares always in stock, the mean repair time is 8 h and $\mu_{\mathrm{eff}} = 1/8 = 0.125\,\mathrm{h}^{-1}$. With one repair crew, the ratio is $\rho = 0.001/0.125 = 0.008$, and by equation (20.12)

$$1 - A = \pi_0 = \frac{2(0.008)^2}{1 + 2(0.008) + 2(0.008)^2} = \frac{0.000128}{1.016128} = 0.00012597,$$

so $A = 99.98740\,\%$ and the expected downtime is $8760 \times 0.00012597 = 1.103$ h per year, against 6.733 h for the baseline and 3.368 h with a second crew.

Faster repair improves both factors of equation (20.17). The mean down time per outage falls from 20 to 8 h, and the failure frequency also falls, because the plant spends less time in state 1, where the second failure occurs: $\pi_1 = 0.016/1.016128 = 0.015746$ gives $f_{\mathrm{F}} = 8760 \times 0.001 \times 0.015746 = 0.1379\,\mathrm{yr}^{-1}$, against $0.3367\,\mathrm{yr}^{-1}$. The second crew, by contrast, shortened only the outage itself. Improving spare-parts logistics can therefore give a larger benefit than adding a repair crew, depending on the costs and on where the present delays are.

20.4 Comparison of maintenance strategies

The models of the two previous sections are enough to compare the candidate improvements on one footing. This section defines five designs, compares their availabilities, attaches costs, and computes the expected annual cost of each.

20.4.1 Five candidate designs

Table 20.2 lists the designs. The baseline is the present plant: two pumps, one repair crew and a mean repair time of 20 h. Option A adds a second repair crew, so that two repairs can progress at the same time, each with mean 20 h. Option B stocks critical spare parts, reducing the mean repair time to 8 h with one crew. Option C is a preventive-maintenance program that reduces the pump failure rate from 0.001 to $0.0007\,\mathrm{h}^{-1}$ with one crew and a mean repair time of 20 h. Option D combines the spares of Option B with the second crew of Option A. The columns give the failure rate, the mean repair time, the number of crews, the rate of the transition from state 0 to state 1, and the incremental annualized program cost of each design, which is used in Section 20.4.3.

Option C carries an assumption that the others do not. It assumes that the preventive program genuinely reduces the failure rate by 30 %. That benefit is not achieved by scheduling maintenance; it must be demonstrated from operating data or supported by an engineering model of the degradation that the maintenance arrests, as in the age replacement of Chapter 18. Under a constant hazard, restoring a working pump to an as-good-as-new condition does not change its failure rate at all.

Table 20.2: Five candidate designs of the pumping plant. Program costs are incremental, annualized and illustrative; they include the relevant equipment, inventory holding, staffing and program expenditures.

Design	λ (h^{-1})	MTTR (h)	Crews	Rate $0 \to 1$ (h^{-1})	Program cost (R$/yr)
Baseline	0.001	20	1	0.05	0
A, second repair crew	0.001	20	2	0.10	45 000
B, critical spare inventory	0.001	8	1	0.125	18 000
C, preventive maintenance	0.0007	20	1	0.05	25 000
D, spares and second crew	0.001	8	2	0.25	63 000

20.4.2 Availability comparison

Each design is evaluated with the one-crew model, equation (20.12), or the two-crew model, equation (20.19), with its own λ and μ. For Option C, $\rho = 0.0007/0.05 = 0.014$ and $\pi_0 = 2(0.014)^2/(1 + 0.028 + 0.000392) = 0.000392/1.028392 = 0.00038118$; for Option D, $1 - A = (0.001/0.126)^2 = 0.00006299$. Table 20.3 gives, for each design, the steady-state availability, the expected downtime per year, the failure frequency of equation (20.16) and the mean down time per outage. The last two columns multiply to the expected downtime, by equation (20.17).

Table 20.3: Steady-state availability, expected downtime, failure frequency and mean down time per complete pumping outage of the five designs (three-state Markov models, exponential repair-time approximation; illustrative values).

Design	A (%)	$\mathbb{E}[D_{\text{yr}}]$ (h/yr)	f_{F} (yr^{-1})	MDT (h)
Baseline	99.92314	6.733	0.3367	20
A, second crew	99.96155	3.368	0.3368	10
B, stock critical spares	99.98740	1.103	0.1379	8
C, preventive maintenance	99.96188	3.339	0.1670	20
D, spares and second crew	99.99370	0.552	0.1379	4

Figure 20.4 draws the expected downtime column. The horizontal axis lists the five designs; the height of each bar is the expected downtime in hours per year, with its value printed above it. By availability alone, Option D performs best, with about half an hour of expected downtime per year, followed by Option B. Options A and C give almost the same downtime by different routes: Option A halves the mean down time of each outage and leaves the frequency unchanged, while Option C reduces the frequency, by $1 - 0.1670/0.3367 = 50.4\,\%$ because π_1 and λ both fall with the failure rate, and leaves the duration unchanged. But Option D combines two investments, and it may not be economically justified relative to the simpler spare-parts program. That question needs costs.

20.4.3 Annual costs and the decision objective

Assume that a complete loss of the pumping function costs

$$c_{\text{D}} = \text{R\$ } 10\,000 \text{ per hour of downtime,}$$

an illustrative financial consequence for this plant, not a general estimate of data center outage costs. The incremental annualized program costs $C_{\text{prog}}(a)$ of the designs are the last column of Table 20.2: R$ 45 000 per year for a second crew, R$ 18 000 for critical spares, R$ 25 000 for the

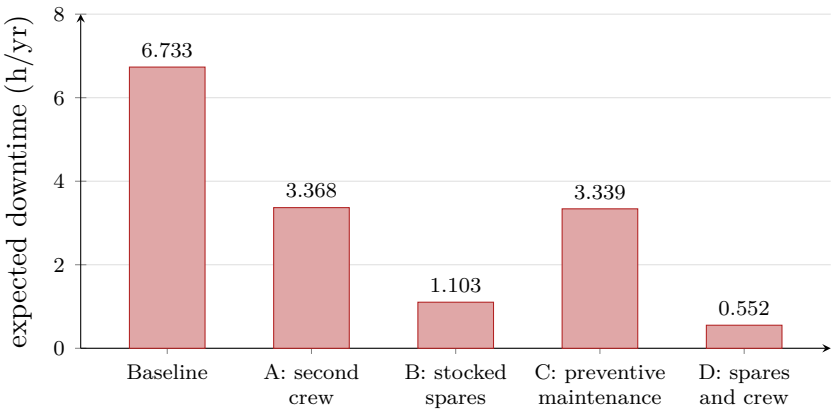


Figure 20.4: Expected downtime per year of the five designs of Table 20.2, $\mathbb{E}[D_{\text{yr}}] = 8760\,\text{h} \times (1 - A)$ from the three-state Markov models (illustrative values).

preventive program and R$ 63 000 for spares and a second crew together. All are expressed on a consistent annual basis. The decision objective is the expected total annual cost of design a,

$$J(a) = c_{\text{D}}\,\mathbb{E}[D_{\text{yr}} \mid a] + C_{\text{prog}}(a), \tag{20.23}$$

in R$/yr, the expected cost of downtime plus the cost of the program.

Equation (20.23) is the cost form of the net-benefit formula of Section 12.5.2. With the expected downtime cost as the consequence and the annualized program cost as the implementation cost, equation (12.31) reads

$$B_{\text{net}}(a) = c_{\text{D}}\,\mathbb{E}[D_{\text{yr}} \mid \text{baseline}] - c_{\text{D}}\,\mathbb{E}[D_{\text{yr}} \mid a] - C_{\text{prog}}(a) = J(\text{baseline}) - J(a),$$

because the baseline has no program cost. Minimizing $J(a)$ over the designs is therefore the same as maximizing the net benefit.

20.4.4 Expected annual cost

Table 20.4 evaluates equation (20.23). For Option B, the downtime cost is R$ 10 000 × 1.1035 h = R$ 11 035 per year and the total is R$ 11 035 + R$ 18 000 = R$ 29 035 per year.

Table 20.4: Expected annual cost of the five designs, equation (20.23), with c_{D} = R$ 10 000 per hour (illustrative values). The last column states whether the design meets an availability requirement $A \geq 99.99\,\%$.

Design	Downtime cost (R$/yr)	Program cost (R$/yr)	$J(a)$ (R$/yr)	$A \geq$ 99.99 %
Baseline	67 333	0	67 333	no
A, second crew	33 679	45 000	78 679	no
B, critical spares	11 035	18 000	**29 035**	no
C, preventive maintenance	33 391	25 000	58 391	no
D, spares and second crew	5 518	63 000	68 518	yes

Figure 20.5 draws the same results. The horizontal axis lists the designs; each bar is the expected annual cost in thousands of reais, split into the expected downtime cost (red, lower part) and the program cost (blue, upper part), with the total printed above the bar. Under these assumptions, Option B, stocking critical spare parts, has the lowest expected annual cost,

R$ 29 035 per year, which is R$ 67 333 − R$ 29 035 = R$ 38 298 per year less than the baseline. Option A costs more than doing nothing: it avoids R$ 33 654 of expected downtime cost per year for a program of R$ 45 000. Option D has the highest availability, but the staffing it adds to Option B costs R$ 45 000 per year and avoids only R$ 11 035 − R$ 5 518 = R$ 5 517 more of downtime cost.

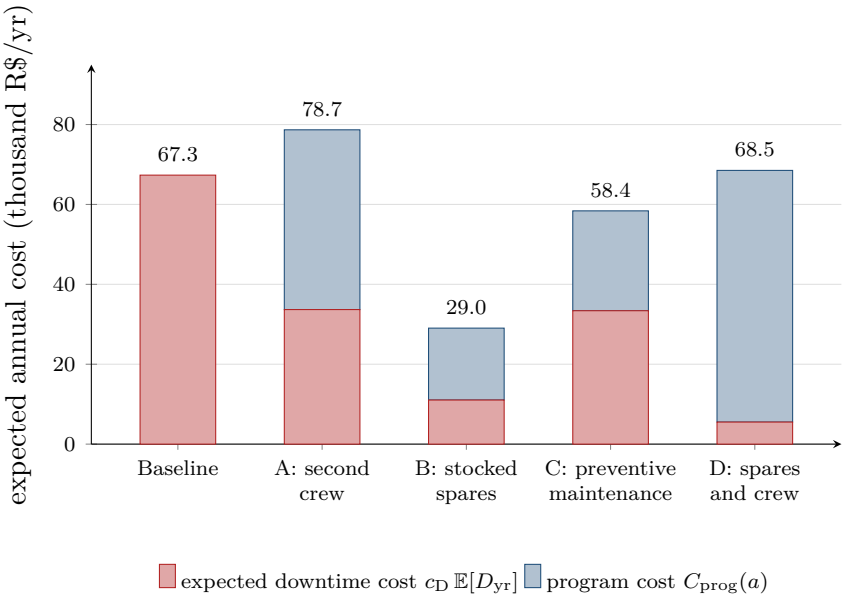


Figure 20.5: Expected annual cost $J(a)$ of the five designs, equation (20.23), in thousands of reais per year: expected downtime cost at c_D = R$ 10 000 per hour plus the annualized program cost (illustrative values). Totals above the bars.

The comparison illustrates a general distinction: the design of maximum availability is not the design of minimum expected cost. Which design is preferred depends on the decision objective. For a highly critical facility where even brief outages are unacceptable, management may impose an availability requirement, for example

$$A \geq A_{\min} = 99.99\,\%, \tag{20.24}$$

and among the five designs only Option D satisfies it (Table 20.4). The design that minimizes cost without a requirement (Option B) then differs from the design that minimizes cost subject to the requirement (Option D), and the difference, R$ 68 518 − R$ 29 035 = R$ 39 483 per year, is the expected annual cost of meeting the requirement with these five designs. Section 20.6.1 shows that a combination outside the five meets it more cheaply.

Principle 20.2. *Maximum availability and minimum expected cost are different objectives. A maintenance decision states which one it optimizes, and an availability requirement enters as a constraint on the cost minimization, not as a replacement for it.*

20.5 Maintenance scheduling, redundancy and spare-parts inventory

Preventive maintenance and spare-parts stocks were treated so far as changes of one rate in the Markov model. This section treats each as a reliability problem of its own: preventive maintenance as a source of planned unavailability, maintenance scheduling as an optimization, and spare-parts demand as a stochastic process whose stock level is chosen for a service level and coupled to the availability model.

20.5.1 Planned unavailability from preventive maintenance

Preventive maintenance usually requires taking equipment out of service. Suppose Pump A is taken offline for a maintenance window of $w = 4\,\mathrm{h}$. During the window the plant depends entirely on Pump B: it is a single-pump system, and a failure of Pump B ends the pumping function. If Pump B is operating at the start of the window, has the constant hazard λ, and cannot be repaired within the window, the probability that it survives the window is its reliability over w,

$$R_B(4\,\mathrm{h}) = \mathrm{e}^{-4\lambda} = \mathrm{e}^{-0.004} = 0.996008,$$

and the probability that it fails during the maintenance is

$$\Pr(\text{Pump B fails during the window}) = 1 - 0.996008 = 0.003992,$$

about $0.399\,\%$. Under these assumptions, a four-hour planned outage of Pump A creates a $0.399\,\%$ probability of complete pumping loss during the window. For comparison, with both pumps operating at the start, the probability of a complete loss within four hours, from the absorbing version of the one-crew model of Section 20.2.2, is 1.49×10^{-5}, close to the $(1 - \mathrm{e}^{-0.004})^2 = 1.59 \times 10^{-5}$ of two independent failures without repair. The maintenance window raises the probability of a loss over those four hours by a factor of about 270.

The $0.399\,\%$ is conditional on Pump B being operational when the window opens. If its state at that moment is uncertain, the result must include the probability p_{start} that it is operating. Under the stated exponential model, with failures during the window independent of the starting state,

$$\Pr(\text{continuous service through the window}) = p_{\mathrm{start}}\,\mathrm{e}^{-\lambda w}, \tag{20.25}$$

since service is continuous only if Pump B is up at the start and then survives the window. If the window were opened at a random time without checking Pump B, p_{start} would be the stationary probability that a given pump is up in the one-crew model, $\pi_2 + \pi_1/2 = 0.960799 + 0.019216 = 0.980015$, and the probability of a loss during the window would be $1 - 0.980015 \times 0.996008 = 0.0239$, six times the conditional value. Confirming that the remaining pump is operating and healthy before the window opens sets $p_{\mathrm{start}} = 1$; this is the reason to schedule maintenance only after such a confirmation.

The expected downtime caused by one window follows from the same model. If Pump B fails at time $\tau \in [0, w]$ after the window opens, the function is lost until Pump A returns at w, so the downtime is $w - \tau$; integrating against the failure density $\lambda \mathrm{e}^{-\lambda\tau}$, by parts,

$$\mathbb{E}[D_{\mathrm{window}}] = \int_0^w (w - \tau)\,\lambda \mathrm{e}^{-\lambda\tau}\,\mathrm{d}\tau = w - \frac{1 - \mathrm{e}^{-\lambda w}}{\lambda} \approx \frac{\lambda w^2}{2}, \tag{20.26}$$

which for $w = 4\,\mathrm{h}$ is $4 - 0.003992/0.001 = 0.007989\,\mathrm{h}$ per window, given Pump B up at the start.

20.5.2 Staggered and simultaneous maintenance

Suppose both pumps require one maintenance task of 4 h per year. Two schedules are possible. Under *simultaneous maintenance* both pumps are taken offline in one window; under *staggered maintenance* one pump is maintained while the other remains in service, and the roles are then exchanged. Figure 20.6 draws both. The horizontal axis is time in hours from an arbitrary origin; each panel has one row for each pump and one for the pumping function. A green bar is a pump in service, a gold bar a pump under maintenance; in the function row, green is available with redundancy, light gold is available through one pump only, and red is the function lost.

If both pumps are unavailable at the same time and no alternate flow source exists, the required cooling-water service is lost by design. The simultaneous schedule of panel (a) therefore

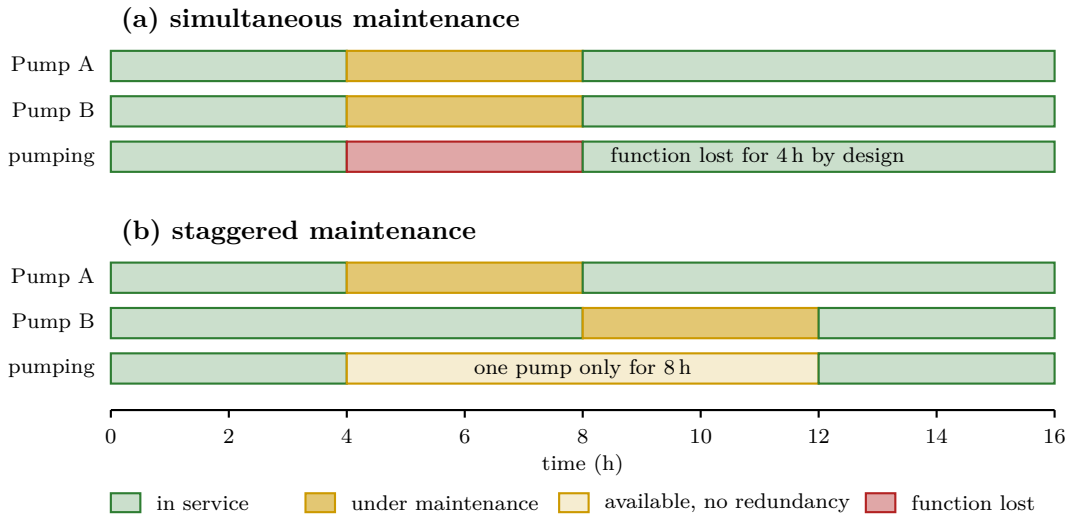


Figure 20.6: Two schedules for one 4 h maintenance task on each pump. (a) Simultaneous: the pumping function is lost for the whole window. (b) Staggered: the function is kept through one pump for 8 h, during which a failure of the pump in service would end it. Schematic.

adds 4 h of certain downtime per year, which at $c_{\mathrm{D}} = \mathrm{R\$}\,10\,000$ per hour costs R$ 40 000 per year, more than half the expected downtime cost of the whole baseline plant. Staggering, panel (b), preserves one operating path, at the price of 8 h per year of reduced redundancy. With the remaining pump confirmed operational at the start of each window, the probability of a loss in the two windows together is $1 - \mathrm{e}^{-0.008} = 0.00797$, and by equation (20.26) the expected downtime is $2 \times 0.007989 = 0.0160$ h per year, an expected cost of R$ 160 per year. Staggering reduces the expected planned downtime by a factor of about 250.

The same calculation settles the hidden cost of Option C. If its preventive program takes each pump out of service for 4 h every quarter, eight staggered windows per year, each window with $\lambda = 0.0007\,\mathrm{h}^{-1}$ adds $4 - (1 - \mathrm{e}^{-0.0028})/0.0007 = 0.005595$ h of expected downtime, 0.0448 h per year in total, or R$ 448 per year; $J(\mathrm{C})$ rises from R$ 58 391 to R$ 58 839 and the ranking of Table 20.4 does not change. Performed simultaneously, the same program would add $4 \times 4 = 16$ h of certain downtime, R$ 160 000 per year, and make Option C the most expensive design. The illustrative schedule is labeled as such; the point is that the scheduling of a preventive program can matter more than its effect on the failure rate.

The best schedule depends on the required load, the available redundant capacity, the condition of the remaining pump, the maintenance duration, the operational risk during the window, the efficiency of the maintenance labor, and the demand and ambient conditions. An operator may schedule maintenance during a period of low IT load or favorable ambient conditions, when the remaining equipment has a larger capacity margin and a single path can carry the load with less risk. The objective is not to minimize the planned maintenance hours; it is to minimize the expected operational consequence while satisfying the maintenance and safety requirements.

20.5.3 Maintenance scheduling as an optimization problem

The scheduling choice can be stated as an optimization. Divide the planning period into intervals $t = 1, \ldots, N$ and define the decision variables

$$u_{i,t} = \begin{cases} 1, & \text{component } i \text{ is scheduled for maintenance in interval } t, \\ 0, & \text{otherwise,} \end{cases} \tag{20.27}$$

which determine which assets are intentionally unavailable; $\mathbf{u}_t$ is the vector of the $u_{i,t}$ in interval t. A simple scheduling problem is

$$\min_{\mathbf{u}} \sum_{t=1}^{N} \Big[C_{\text{maint}}(\mathbf{u}_t) + \mathbb{E}\big[C_{\text{outage}}(\mathbf{u}_t)\big] \Big], \tag{20.28}$$

where $C_{\text{maint}}(\mathbf{u}_t)$ is the cost of the maintenance performed in interval t and $\mathbb{E}[C_{\text{outage}}(\mathbf{u}_t)]$ the expected outage cost in that interval given the planned configuration, computed from the availability model with the scheduled components removed. It is subject to a resource constraint,

$$\sum_{i} u_{i,t} \le N_{\text{crews}} \quad \text{for every } t,$$

where N_{crews} is the number of crews, and to an adequacy constraint. A deterministic adequacy constraint requires that the capacity of the components not under maintenance meets the demand, $\sum_i (1 - u_{i,t})\, C_i \ge L_t$ for every t. A probabilistic one, which accounts for forced outages of the components still in service, requires

$$\Pr\big(C_{\text{available}}(t) < L_t \mid \mathbf{u}_t\big) \le \varepsilon \quad \text{for every } t, \tag{20.29}$$

where L_t is the required load in interval t and ε the acceptable probability of insufficient capacity. For the two-pump plant in a staggered window, the left side is the 0.003992 of Section 20.5.1, so a requirement $\varepsilon = 0.01$ admits a 4 h window and $\varepsilon = 0.001$ does not, unless the window is shortened below $-\ln(0.999)/\lambda = 1.0$ h.

Equations (20.28)–(20.29) define a risk-constrained maintenance scheduling problem, which combines reliability engineering with operations research. A realistic formulation also imposes maintenance deadlines, task precedence, technician qualifications, spare-part inventories and minimum outage durations for each task.

20.5.4 Stochastic demand for spare parts

Consider a facility that keeps replacement motors for its cooling pumps. Let $K(\tau)$ be the number of spare-motor demands in an interval of length τ. If replacement demands follow a homogeneous Poisson process with rate ν (Chapter 17), the count is Poisson distributed,

$$K(\tau) \sim \text{Poisson}(\nu\tau), \qquad \Pr[K(\tau) = k] = \frac{\mathrm{e}^{-\nu\tau}(\nu\tau)^k}{k!}, \quad k = 0, 1, 2, \ldots \tag{20.30}$$

This is a simplifying model, suitable when the demands are approximately independent and their rate is stable; a fleet whose failure rate increases with age, or whose failures share common causes, violates it. As an illustration, let $\nu = 2$ motors per year and let the replenishment lead time, the time from placing an order to receiving the motor, be $\tau = 0.25$ yr. The expected demand during the lead time is

$$\mathbb{E}[K(\tau)] = \nu\tau = 2 \times 0.25 = 0.5 \text{ motors}.$$

With s motors in stock at the start of the lead time and no other supply, the demands during the lead time exceed the stock with probability

$$\Pr[K(\tau) > s] = 1 - \sum_{k=0}^{s} \frac{\mathrm{e}^{-\nu\tau}(\nu\tau)^k}{k!}, \tag{20.31}$$

the probability that the stock level s is insufficient to cover the demand during one lead time, under a policy that reorders when the stock reaches that level.

20.5.5 The stock level for a service level

For $K(\tau) \sim \text{Poisson}(0.5)$, Table 20.5 lists the probability of each number of demands and the probability that the demand exceeds each stock level. Its rows are the number of demands k; the second column is $\Pr[K = k]$, the third $\Pr[K > k]$, the stockout probability with $s = k$, and the fourth its complement, the service level.

Table 20.5: Spare-motor demand during a lead time, $K(\tau) \sim \text{Poisson}(0.5)$, with $\nu = 2\,\text{yr}^{-1}$ and $\tau = 0.25\,\text{yr}$ (illustrative values).

Demands k (or stock s)	$\Pr[K = k]$	$\Pr[K > s]$, stockout	$\Pr[K \le s]$, service level
0	0.606531	0.393469	0.606531
1	0.303265	0.090204	0.909796
2	0.075816	0.014388	0.985612
3	0.012636	0.001752	0.998248
4	0.001580	0.000172	0.999828

The probabilities of 0, 1 and 2 demands are 60.65 %, 30.33 % and 7.58 %, and three or more demands occur with probability 1.44 %. For one spare in stock,

$$\Pr[K(\tau) > 1] = 1 - \mathrm{e}^{-0.5}(1 + 0.5) = 1 - 0.606531 \times 1.5 = 0.0902,$$

and for two,

$$\Pr[K(\tau) > 2] = 1 - \mathrm{e}^{-0.5}\Big(1 + 0.5 + \frac{0.5^2}{2}\Big) = 1 - 0.606531 \times 1.625 = 0.0144.$$

A second spare reduces the probability of exhausting the stock during a lead time from 9.02 % to 1.44 %. Figure 20.7 draws the distribution. The horizontal axis is the number of demands during the lead time; each bar is its probability, printed above it. The dashed vertical lines separate the demands covered by a stock of one and of two motors from those that exceed it, and the label beside each line gives the stockout probability, the total height of the bars to its right.

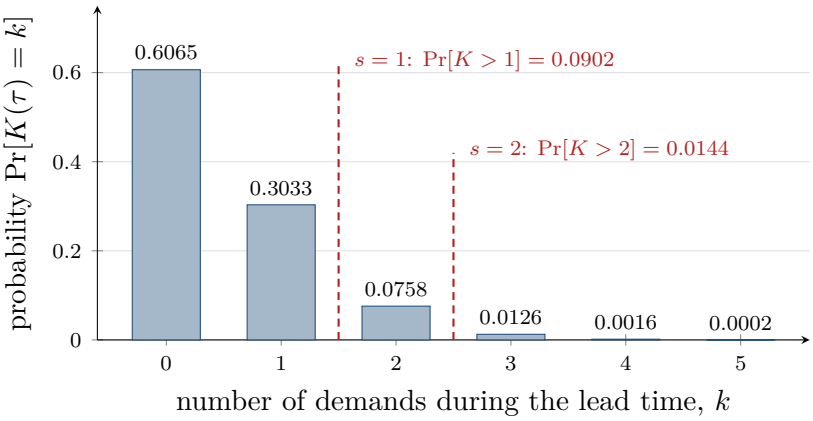


Figure 20.7: Distribution of the spare-motor demand during a lead time, $\Pr[K(\tau) = k] = \mathrm{e}^{-0.5}\,0.5^k/k!$ for $\nu = 2\,\text{yr}^{-1}$ and $\tau = 0.25\,\text{yr}$ (illustrative values), with the stockout probabilities of stock levels $s = 1$ and $s = 2$, equation (20.31).

The stock level that meets a required service level follows directly. If the probability of a stockout during a lead time must not exceed ε, the stock level is the smallest s that satisfies

$\Pr[K(\tau) > s] \le \varepsilon$,

$$s^*(\varepsilon) = \min\Big\{ s \in \{0,1,2,\ldots\} : \sum_{k=0}^{s} \frac{\mathrm{e}^{-\nu\tau}(\nu\tau)^k}{k!} \ge 1-\varepsilon \Big\}. \tag{20.32}$$

The minimum exists for every $\varepsilon > 0$, because the partial sums increase with s toward 1. From Table 20.5, $s^* = 1$ for $\varepsilon = 0.10$, $s^* = 2$ for $\varepsilon = 0.05$, and $s^* = 3$ for $\varepsilon = 0.01$. Each additional spare incurs inventory costs: purchase capital, storage, inspection and preservation, periodic testing, and the management of obsolescence. The appropriate stock level therefore depends on the consequences of a stockout, which the availability model of the plant prices, and on the replenishment lead time, which sets $\nu\tau$.

The scope of the stockout probability. A stockout probability over one replenishment lead time is not automatically the long-run probability that no spare is available when a repair needs one, which is the $1 - p_{\mathrm{S}}$ of equation (20.21). The second depends on the replenishment policy, the stock position at the start of each lead time, the outstanding orders, the treatment of backorders, and the dynamics of the inventory over many lead times. A sound spare-parts model couples an inventory-state process with the repair-state process of the equipment.

20.5.6 Coupling the inventory to the Markov model

Suppose every pump repair consumes one replacement motor. If a motor is in stock, the repair proceeds at the rate $\mu_{\mathrm{s}} = 1/8\,\mathrm{h}^{-1}$; if none is available, the repair waits for replenishment. Instead of assigning one average repair rate through equation (20.21), the state of the Markov model is enlarged to

$$X(t) = \big(n(t), s(t)\big), \tag{20.33}$$

where $n(t)$ is the number of operating pumps and $s(t)$ the number of spare motors in stock. The state space now distinguishes conditions such as two pumps operating with one spare in stock, one pump operating with no spare, both pumps failed with a spare on order, and one pump under repair with a replenishment in progress. The transitions include component failures, repair completions, which consume a spare, and replenishments, which add one.

Figure 20.8 draws the coupled model for an illustrative policy that keeps at most one motor: every motor consumed is reordered at once, and each order arrives after an exponential lead time with mean $1/\gamma$. The columns are the number of operating pumps, 2, 1 and 0 from left to right; the upper row has one motor in stock and the lower row none. Horizontal arrows are pump failures, at 2λ and λ; diagonal arrows are repairs with a motor in stock, at μ_{s} with one crew, which add an operating pump and empty the stock; vertical arrows are replenishments, at γ. In the lower row a failed pump cannot be repaired: the plant must wait for a replenishment before any repair transition is possible. The two states of the right column are the unavailable states.

The stationary distribution of the coupled model is found as in Section 20.2.2, from $\boldsymbol{\pi}\mathbf{Q} = \mathbf{0}$ with the six states, now numerically, since the chain is no longer a simple chain of neighbors. Table 20.6 gives the results for stocks of up to one, two and three motors, each with a replenishment lead time of mean 40 h (an expedited delivery, the difference between the 48 and 8 h of Section 20.3.2) and of mean one week, 168 h; with $\lambda = 0.001\,\mathrm{h}^{-1}$, $\mu_{\mathrm{s}} = 0.125\,\mathrm{h}^{-1}$ and one crew. The models with larger stocks have the same structure with more rows. The last row is Option B, which assumes a spare is always in stock.

The table shows what the effective-rate model of Option B hides. With one motor and a one-week lead time, the plant is available only 98.62 % of the time, worse than the baseline, because both pumps consume motors at a combined rate near 2λ, about 17 per year, and the single motor is often on order when it is needed. With two motors and an expedited 40 h lead

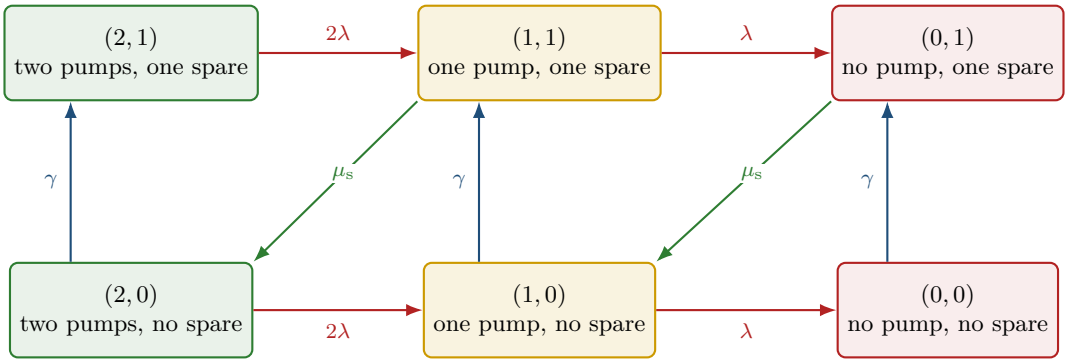


Figure 20.8: Coupled availability and inventory model, state (n, s) with n operating pumps and s spare motors in stock, for an illustrative policy that keeps at most one motor. Failures at 2λ and λ (horizontal), repairs consuming the motor at μ_{s} with one crew (diagonal), replenishment at γ (vertical). The right column is unavailable.

Table 20.6: Steady-state availability and expected downtime of the coupled model of Figure 20.8, extended to larger stocks, with $\lambda = 0.001\,\mathrm{h}^{-1}$, $\mu_{\mathrm{s}} = 0.125\,\mathrm{h}^{-1}$, one crew, every repair consuming one motor, each consumed motor reordered at once and each order delivered after an exponential lead time (illustrative values).

	Lead time mean 40 h		Lead time mean 168 h	
Maximum stock	A (%)	$\mathbb{E}[D_{\mathrm{yr}}]$ (h/yr)	A (%)	$\mathbb{E}[D_{\mathrm{yr}}]$ (h/yr)
One motor	99.95668	3.795	98.61678	121.2
Two motors	99.98706	1.134	99.92513	6.559
Three motors	99.98740	1.104	99.98426	1.379
Spare always in stock (Option B)	99.98740	1.103	99.98740	1.103

time, the availability is within 0.0004 percentage points of Option B. The assumption $p_{\mathrm{S}} = 1$ is therefore a claim about the stock level and the replenishment time, which the coupled model can check. The enlarged state space is a more realistic availability model where spare-part constraints dominate the restoration time, and it shows the general cost of fidelity: every resource made explicit multiplies the number of states.

20.6 Investment choices and risk-informed optimization

The comparison of Section 20.4 used one set of inputs and five designs. This section extends it to combinations of the investments and to other input values in a worked table, then treats the inputs as uncertain, states the availability-constrained decision, and ends with the adequacy of the model itself.

20.6.1 A worked table of investment choices

The decision model is the one of Section 20.4: each design fixes the pump failure rate λ (multiplied by 0.7 when the preventive program is included), the mean repair time (20 h, or 8 h with stocked spares) and the number of crews; the steady-state availability comes from equation (20.12) with one crew or equation (20.19) with two; the expected downtime is $8760\,\mathrm{h} \times (1 - A)$; and the expected annual cost is equation (20.23), with the program costs of Table 20.2 added when programs are combined. Table 20.7 evaluates the five designs and three further combinations of the three investments at the base inputs: $\lambda = 0.001\,\mathrm{h}^{-1}$, a 30 % reduction from the preventive program, and $c_{\mathrm{D}} = \mathrm{R\$}\,10\,000$ per hour.

Table 20.7: Availability, expected downtime, failure frequency and expected annual cost of eight designs at the base inputs. Crew A is the second crew, spares B the critical spare inventory, program C the preventive program; combined program costs are summed (illustrative values).

Design	λ (h^{-1})	MTTR (h)	Crews	A (%)	$\mathbb{E}[D_{\mathrm{yr}}]$ (h/yr)	f_{F} (yr^{-1})	$J(a)$ (R$/yr)
Baseline	0.001	20	1	99.92314	6.733	0.3367	67 333
A	0.001	20	2	99.96155	3.368	0.3368	78 679
B	0.001	8	1	99.98740	1.103	0.1379	29 035
C	0.0007	20	1	99.96188	3.339	0.1670	58 391
D = A + B	0.001	8	2	99.99370	0.552	0.1379	68 518
B + C	0.0007	8	1	99.99380	0.543	0.0679	48 433
A + C	0.0007	20	2	99.98094	1.670	0.1670	86 699
A + B + C	0.0007	8	2	99.99690	0.272	0.0679	90 717

One row of the arithmetic, B + C: with $\lambda = 0.0007\,\mathrm{h}^{-1}$ and $\mu = 0.125\,\mathrm{h}^{-1}$, $\rho = 0.0056$ and

$$1 - A = \frac{2(0.0056)^2}{1 + 2(0.0056) + 2(0.0056)^2} = \frac{0.00006272}{1.01126272} = 0.00006202,$$

so $A = 99.99380\,\%$, the expected downtime is $8760 \times 0.00006202 = 0.5433\,\mathrm{h}$ per year, and $J = \mathrm{R\$}\,10\,000 \times 0.5433 + \mathrm{R\$}\,18\,000 + \mathrm{R\$}\,25\,000 = \mathrm{R\$}\,48\,433$ per year.

Three trends follow from the table. First, the investments do not add: the second crew is worth less once spares are stocked, because it shortens an outage that is already short, and so A + B + C costs more than B + C while gaining only 0.27 h per year. Second, preventive maintenance and spares act on different factors, the failure frequency and the outage duration, and therefore combine well: B + C has almost the availability of D at R$ 20 085 per year less.

Third, with the requirement (20.24), the cheapest design that meets $A \geq 99.99\,\%$ among the eight is B + C, not D; enlarging the set of candidate designs changed the constrained decision, while B remains the unconstrained minimum.

Table 20.8 varies the inputs. Each row is one set of inputs, changed one at a time from the base: the downtime cost raised to R$ 50 000 per hour, the failure rate halved, the baseline repair time raised to 50 h with stocked-spare repairs still 8 h, and, as additional cases, the downtime cost lowered to R$ 1000 per hour and the failure rate doubled. The columns are the expected annual costs of the five designs, in thousands of reais, the design that minimizes expected cost, and the design that minimizes it subject to $A \geq 99.99\,\%$.

Table 20.8: Expected annual cost $J(a)$ of the five designs, in thousands of R$ per year, for inputs changed one at a time from the base ($\lambda = 0.001\,\mathrm{h}^{-1}$, baseline MTTR 20 h, spare MTTR 8 h, 30 % preventive reduction, c_D = R$ 10 000 per hour); illustrative values.

Inputs	Baseline	A	B	C	D	Minimum cost	Minimum with $A \geq 99.99\,\%$
Base	67.3	78.7	29.0	58.4	68.5	B	D
c_D = R$ 50 000/h	336.7	213.4	73.2	192.0	90.6	B	D
$\lambda = 0.0005\,\mathrm{h}^{-1}$	17.2	53.6	20.8	33.5	64.4	Baseline	B
Baseline MTTR 50 h	396.4	243.6	29.0	225.1	68.5	B	D
c_D = R$ 1000/h	6.7	48.4	19.1	28.3	63.6	Baseline	D
$\lambda = 0.002\,\mathrm{h}^{-1}$	258.8	174.6	61.4	154.9	84.7	B	none

A higher downtime cost raises the value of every reduction in unavailability: at R$ 50 000 per hour the second crew becomes cheaper than the baseline, and Option D comes within R$ 17 400 per year of Option B, so a still higher cost would make the more expensive programs worth their price. When failures become rarer ($\lambda = 0.0005\,\mathrm{h}^{-1}$) or downtime cheaper ($c_\mathrm{D}$ = R$ 1000 per hour), repair capacity and spare inventory buy little, and doing nothing is the minimum-cost choice. When the baseline repair takes 50 h while a stocked-spare repair still takes 8 h, the value of spare-parts availability grows, because waiting for parts becomes the larger contributor to unavailability: the baseline cost rises to R$ 396 380 per year while Option B is unchanged. With the failure rate doubled, no design meets the availability requirement; even A + B + C gives $1 - A = (0.0014/0.1264)^2 = 0.00012268$, or 99.98773 %. The maintenance strategy depends on both the reliability physics, through λ and the repair times, and the economic consequences, through c_D and the program costs.

20.6.2 Uncertainty in availability calculations

The analysis so far treats λ, μ and c_D as known. In practice each must be estimated. The failure rate may depend on the operating load and the equipment age; the repair duration on spare availability, technician staffing, environmental conditions and the severity of the failure; the downtime cost on the IT workload or the production requirements, which can vary widely. Let the uncertain parameters be collected in the vector $\boldsymbol{\theta} = (\lambda, \mu, c_\mathrm{D}, \ldots)$, described by a probability distribution that represents the state of knowledge about them, for example a posterior from Chapter 8. The annual cost of design a is then a function $J(a, \boldsymbol{\theta})$. A risk-neutral expected-cost decision is

$$a^* = \arg\min_a \mathbb{E}_{\boldsymbol{\theta}}\big[J(a, \boldsymbol{\theta})\big], \tag{20.34}$$

the expectation being taken over the parameter uncertainty, on the assumption that its distribution is appropriate for decision analysis. Other criteria impose a chance constraint on the

availability,

$$\Pr_{\boldsymbol{\theta}}\big(A(a,\boldsymbol{\theta}) < A_{\min}\big) \le \varepsilon, \tag{20.35}$$

which expresses the required confidence that an availability target is met under the uncertain parameters.

As an illustration, suppose the pump failure rate is uncertain and takes the values 0.0005, 0.001 and $0.002\,\mathrm{h}^{-1}$ with probabilities 0.25, 0.5 and 0.25, all other inputs at their base values. Table 20.9 gives the annual cost of each design for each value, its expectation, and the probability that the availability falls below 99.99 %.

Table 20.9: Annual cost of the five designs under an uncertain failure rate $\lambda \in \{0.0005, 0.001, 0.002\}\,\mathrm{h}^{-1}$ with probabilities $\{0.25, 0.5, 0.25\}$, its expectation, and the probability that $A < 99.99\,\%$ (illustrative values; costs in R$ per year).

Design	$\lambda = 0.0005$	$\lambda = 0.001$	$\lambda = 0.002$	$\mathbb{E}_{\boldsymbol{\theta}}[J]$	$\Pr(A < 99.99\,\%)$
Baseline	17 173	67 333	258 789	102 657	1
A	53 587	78 679	174 586	96 383	0.75
B	20 781	29 035	61 439	35 072	0.75
C	33 465	58 391	154 880	76 282	0.75
D	64 390	68 518	84 725	71 538	0.25

Option B remains the risk-neutral choice, with an expected cost of R$ 35 072 per year. That expectation is higher than the cost at the base failure rate, and higher than the cost at the mean failure rate, $0.001125\,\mathrm{h}^{-1}$, which is R$ 31 938: the unavailability grows as ρ^2, a convex function of λ, so the expected cost under uncertainty exceeds the cost at the expected parameter. A decision computed at a point estimate understates the expected cost whenever the cost is convex in the uncertain parameter. Under the chance constraint (20.35) with $A_{\min} = 99.99\,\%$, Option D is feasible for $\varepsilon \ge 0.25$, and no design of the five is feasible for a smaller ε; a stricter confidence requirement calls for a design outside the set.

Uncertainty in a model parameter must be distinguished from the random operating state of the system. The availability $A(a,\boldsymbol{\theta})$ already averages over the aleatory randomness of failures and repairs for fixed parameters; the distribution of $\boldsymbol{\theta}$ is epistemic, a statement of what is known about the parameters. A complete probabilistic decision model may need to represent both, and must not count either twice.

20.6.3 The cheapest expected solution and an availability requirement

Suppose the operator must meet $A \ge 99.99\,\%$. Among the five designs of Section 20.4, only Option D satisfies it. The unconstrained cost-minimizing design, Option B, does not: its 99.98740 % is below the target. The constrained problem

$$\min_{a} J(a) \quad \text{subject to} \quad A(a) \ge A_{\min} \tag{20.36}$$

therefore produces a different decision from the minimization of cost alone. The distinction is common in safety-critical and high-reliability engineering. An operator may reasonably require an availability constraint even when accepting more outages would minimize the expected financial expenditure, because the cost c_{D} does not capture every consequence of an outage (contractual penalties, reputation, the safety of people working in an overheated hall), and because the expected value treats a rare long outage and many short ones alike, as Chapter 3 observed for risk. The value of the constraint is visible in the decision: the difference between the constrained and the unconstrained optimum, R$ 19 398 per year for B + C against B in Table 20.7, is the expected annual price of the requirement.

Common mistake

Treating the design of maximum availability as the economically optimal one, or the design of minimum expected cost as acceptable. The first can buy reductions in downtime that cost more than they are worth; the second can violate an availability requirement that expected cost does not represent. The objective and the constraints are stated before the designs are ranked.

20.6.4 Model adequacy

The most important caveat concerns the model itself. The analysis treats a complete loss of pumping as equivalent to a complete loss of cooling, which is a simplification. The cooling availability of a data center depends on the chiller capacity, the cooling towers and heat rejection, the electrical infrastructure, valve positions, the control systems, the distribution of flow and pressure, any thermal storage, the IT load, the ambient conditions, and common-cause failures. Moreover, a pump outage does not necessarily interrupt the IT equipment at once: the thermal capacity of the room and the coolant provides time for recovery.

A more credible model defines the top event by the physical requirement. Let $T_{\mathrm{IT}}(t)$ be the temperature of the air supplied to the IT equipment, a temperature and not the time to failure T, and $T_{\max}$ its upper limit. The top event over an assessment period $[0, H]$ is

$$\max_{t\in[0,H]} T_{\mathrm{IT}}(t) > T_{\max}. \tag{20.37}$$

The heat-removal rate of the cooling plant depends on the states of the components, the control actions and the external conditions,

$$\dot{Q}_{\mathrm{cool}}(t) = \psi\big(\mathbf{x}(t), \mathbf{u}(t), \mathbf{z}(t)\big),$$

where $\mathbf{x}(t)$ collects the component states (pumps, chillers, towers), $\mathbf{u}(t)$ the control and maintenance decisions, and $\mathbf{z}(t)$ the ambient conditions, and ψ is the plant's capacity model. The temperature then follows an energy balance,

$$C_{\mathrm{th}} \frac{\mathrm{d}T_{\mathrm{IT}}}{\mathrm{d}t} = \dot{Q}_{\mathrm{IT}}(t) - \dot{Q}_{\mathrm{cool}}(t), \tag{20.38}$$

where C_{th} is the effective heat capacity of the room and the coolant, in J/K, and $\dot{Q}_{\mathrm{IT}}$ the heat dissipated by the IT load, in W. As an order of magnitude, with $\dot{Q}_{\mathrm{cool}} = 0$ and the IT load at the plant's 900 kW duty, an illustrative $C_{\mathrm{th}} = 50\,\mathrm{MJ/K}$ gives a temperature rise of $900\,000/(50 \times 10^6) = 0.018\,\mathrm{K/s}$, about 1.1 K per minute, so a margin of 10 K lasts about nine minutes; outages shorter than that would cause no violation of equation (20.37).

The two models have separate roles. The availability and maintenance model determines which cooling configurations occur and how long they persist; the physical model determines whether those configurations actually violate the required function. The integration matters more as the consequences become more dependent on the state, the load and the duration of each outage. The models of this chapter describe equipment that is either operating or failed; a later chapter on degradation models the physical deterioration that precedes failure, and the condition measurements that can inform its prediction.

20.7 What to retain

Definition 20.1 (Availability optimization)**.** Availability optimization chooses redundancy, repair capacity, spare parts and preventive maintenance to minimize the expected total annual cost $J(a) = c_{\mathrm{D}}\,\mathbb{E}[D_{\mathrm{yr}} \mid a] + C_{\mathrm{prog}}(a)$, subject where required to an availability constraint

$A(a) \geq A_{\min}$, using an availability model that represents the complete failure-and-restoration architecture. It joins reliability, maintainability, resource allocation and operational constraints; these interact, and the result is not their sum.

Availability is a system-level property that depends on component failures, redundancy, repair resources, logistics and operating requirements. Repair capacity matters: a second crew halves the unavailability of the two-pump plant without changing any failure rate, but only if two repairs can truly progress at once. Spare-parts availability can dominate the restoration time, and its benefit depends on the stock level and the lead time, which a coupled model can check. Preventive maintenance trades a lower failure rate against planned unavailability, and its schedule can matter more than its effect on the failure rate. Maximum availability and minimum expected cost are different objectives, and uncertainty in the inputs changes both the expected costs and the confidence that a requirement is met. Equipment availability does not imply an adequate physical function, which must be modeled when the consequences depend on it. Table 20.10 lists the chapter's terms, their meaning, expression and units, and the subsection that defines each.

Table 20.10: The terms of this chapter and where each is defined.

Term	Meaning	Expression	Units	Defined in
First-failure reliability	no loss of the function in $[0,t]$	$R(t)=\Pr(T>t)$	–	20.1.1
Point availability	function available at t	$A(t)=\Pr[Y(t)=1]$	–	20.1.1
Average availability	expected fraction of $[0,H]$ available	$\bar{A}(H)=H^{-1}\int_0^H A\,\mathrm{d}t$	–	20.1.1
Expected downtime	expected hours unavailable	$H\,[1-\bar{A}(H)]$	h	20.1.1
Steady-state availability	long-run probability available	$\lim A(t)=\mathbb{E}[U]/(\mathbb{E}[U]+\mathbb{E}[D])$	–	20.1.1
One-crew unavailability	two pumps, one repair at a time	$2\rho^2/(1+2\rho+2\rho^2)$, $\rho=\lambda/\mu$	–	20.2.2
Failure frequency	complete outages per year	$f_{\mathrm{F}}=8760\,\lambda\,\pi_1$	yr^{-1}	20.2.3
Mean down time	mean duration of one outage	MDT; $\mathbb{E}[D_{\mathrm{yr}}]=f_{\mathrm{F}}\,\mathrm{MDT}$	h	20.2.3
Two-crew unavailability	two independent repairs	$[\lambda/(\lambda+\mu)]^2$	–	20.2.4
Logistic delay	waiting for resources	D_{waiting} in T_{R}	h	20.3.1
Spare-dependent repair time	mixture over spare availability	equation (20.21)	h	20.3.2
Effective repair rate	exponential approximation	$\mu_{\mathrm{eff}}=1/\mathbb{E}[T_{\mathrm{R}}]$	h^{-1}	20.3.3
Expected annual cost	downtime cost plus program cost	$J(a)$, equation (20.23)	R$/yr	20.4.3
Window loss probability	loss during a maintenance window	$1-p_{\mathrm{start}}\,\mathrm{e}^{-\lambda w}$	–	20.5.1
Scheduling problem	maintenance plan with risk constraint	equations (20.28)–(20.29)	R$	20.5.3
Lead-time demand	spare demands in a lead time	$K(\tau)\sim\mathrm{Poisson}(\nu\tau)$	–	20.5.4
Stockout probability, service level	demand exceeds stock; its complement	$\Pr[K(\tau)>s]$, $\Pr[K(\tau)\le s]$	–	20.5.5
Stock level for a service level	smallest adequate stock	$s^*(\varepsilon)$, equation (20.32)	–	20.5.5
Coupled model	equipment and inventory states	$X(t)=(n(t),s(t))$	–	20.5.6
Risk-neutral decision	minimum expected cost under uncertainty	$\arg\min_a \mathbb{E}_{\boldsymbol{\theta}}[J(a,\boldsymbol{\theta})]$	R$/yr	20.6.2
Chance constraint	confidence in an availability target	equation (20.35)	–	20.6.2
Constrained optimum	minimum cost meeting $A_{\min}$	equation (20.36)	R$/yr	20.6.3

Further reading. Rausand, Barros and Høyland, *System Reliability Theory*, third edition, for availability, maintainability, repairable systems and optimization; Modarres, Kaminskiy and Krivtsov, *Reliability Engineering and Risk Analysis*, for availability and maintenance

modeling; David J. Smith, *Reliability, Maintainability and Risk*, for practical RAM analysis and maintenance strategies; Jardine and Tsang, *Maintenance, Replacement, and Reliability*, for maintenance optimization and replacement decisions.

20.8 Exercises

Exercise 20.1. (a) State what distinguishes point availability from first-failure reliability. (b) State why adding a second repair crew can improve the availability of a redundant pair. (c) State why keeping critical spares on site can improve availability. (d) Name the measure that reflects both outage frequency and outage duration under stationary conditions. (e) State whether the strategy of highest availability necessarily minimizes the expected annual cost. (f) State why planned maintenance is sometimes scheduled during low-load periods. (Answers: (a) point availability admits earlier failures followed by repairs, asking only whether the function is available at t, while reliability asks whether it has survived without any loss; (b) it changes the restoration transition rate out of the fully failed state, since two repairs progress at once, shortening the outage without changing any failure rate; (c) it reduces the logistic delay in the restoration time; (d) the steady-state unavailability, equal to the outage frequency times the mean down time per outage; (e) no, because implementation costs and downtime consequences set the economic tradeoff, and higher availability can cost more than the outages it avoids; (f) the remaining equipment then has a larger capacity margin, so the temporary loss of redundancy carries less operational risk.)

Exercise 20.2. The pumps of the plant age and their failure rate doubles to $\lambda = 0.002\,\mathrm{h}^{-1}$, with MTTR $= 20\,\mathrm{h}$. (a) With one repair crew, compute ρ, the stationary probabilities, the steady-state availability, the expected downtime per year, the failure frequency and the mean down time per outage, and verify equation (20.17). (b) Repeat with two crews. (c) Compute the mean time to the first complete outage from the new condition, and state whether the number of crews affects it. (d) At $c_{\mathrm{D}} = \mathrm{R\$}\,10\,000$ per hour and a second-crew cost of R\$ 45 000 per year, state whether the second crew is now justified. (Answers: (a) $\rho = 0.04$, $\pi_2 = 0.923191$, $\pi_1 = 0.073855$, $\pi_0 = 0.0032/1.0832 = 0.0029542$, $A = 99.70458\,\%$, $25.879\,\mathrm{h}$ per year, $f_{\mathrm{F}} = 8760 \times 0.002 \times 0.073855 = 1.2939\,\mathrm{yr}^{-1}$, MDT $= 20\,\mathrm{h}$, and $1.2939 \times 20 = 25.88\,\mathrm{h}$; (b) $1 - A = (0.04/1.04)^2 = 0.0014793$, $A = 99.85207\,\%$, $12.959\,\mathrm{h}$ per year, $f_{\mathrm{F}} = 1.2959\,\mathrm{yr}^{-1}$, MDT $= 10\,\mathrm{h}$; (c) $m_2 = (0.006 + 0.05)/(2 \times 0.002^2) = 7000\,\mathrm{h}$, the same for one or two crews; (d) yes: it avoids $12.920\,\mathrm{h}$, worth R\$ 129 203 per year, for R\$ 45 000, a net saving of R\$ 84 203 per year.)

Exercise 20.3. A partial stocking program raises the probability that a spare is in stock at a repair demand from 0.70 to $p_{\mathrm{S}} = 0.90$, at an illustrative cost of R\$ 10 000 per year; repairs take a mean of 8 h with the spare and 48 h without. (a) Compute the mean repair time. (b) With one crew, $\lambda = 0.001\,\mathrm{h}^{-1}$ and the exponential repair-time approximation, compute the steady-state availability and the expected downtime per year. (c) Compute J at $c_{\mathrm{D}} = \mathrm{R\$}\,10\,000$ per hour and compare it with Option B. (d) If the durations within each branch were exponential, compute the standard deviation of the repair time and compare it with that of the exponential approximation. (Answers: (a) $0.9 \times 8 + 0.1 \times 48 = 12\,\mathrm{h}$; (b) $\rho = 0.012$, $1 - A = 0.000288/1.024288 = 0.00028117$, $A = 99.97188\,\%$, $2.463\,\mathrm{h}$ per year; (c) R\$ 24 631 + R\$ 10 000 = R\$ 34 631 per year, more than Option B's R\$ 29 035, so full stocking remains preferable; (d) $\mathbb{E}[T_{\mathrm{R}}^2] = 0.9 \times 2 \times 64 + 0.1 \times 2 \times 2304 = 576\,\mathrm{h}^2$, variance $576 - 144 = 432\,\mathrm{h}^2$, standard deviation $20.8\,\mathrm{h}$, against $12\,\mathrm{h}$ for the exponential approximation.)

Exercise 20.4. (a) A larger facility consumes spare motors at $\nu = 4$ per year, with a replenishment lead time of $\tau = 0.25\,\mathrm{yr}$. Compute the stockout probability $\Pr[K(\tau) > s]$ for $s = 0, 1, 2, 3, 4$, and the stock levels s^* that meet service levels of 95 % and 99 %. (b) Pump A

of the plant is taken out of service for a maintenance window of $w = 8\,\mathrm{h}$, with $\lambda = 0.001\,\mathrm{h}^{-1}$ and no repair within the window. Compute the probability of a complete pumping loss during the window given that Pump B is operating at the start, the expected downtime caused by the window, and the probability of a loss when Pump B is operating at the start with probability 0.98. (Answers: (a) $\nu\tau = 1$; 0.6321, 0.2642, 0.0803, 0.0190, 0.0037; $s^{*} = 3$ for 95 % and $s^{*} = 4$ for 99 %; (b) $1 - \mathrm{e}^{-0.008} = 0.007968$; $8 - 0.007968/0.001 = 0.0319\,\mathrm{h}$, close to $\lambda w^{2}/2 = 0.032\,\mathrm{h}$; $1 - 0.98\,\mathrm{e}^{-0.008} = 0.0278$.)