Probability Models and Statistical Methods in Reliability

Larry Leemis
Department of Mathematics
College of William and Mary
Williamsburg, VA 23187-8795
leemis@math.wm.edu 757-221-2034

Undergraduate Simulation, Modeling and Analysis
February 14, 2000

Outline

1. Introduction
2. Coherent Systems Analysis
3. Lifetime Distributions
4. Parametric Lifetime Models
5. Specialized Models
6. Repairable Systems
7. Lifetime Data Analysis
8. Fitting Parametric Models to Data
9. Parametric Estimation for Models with Covariates
10. Nonparametric Methods
11. Assessing Model Adequacy

1. Introduction

Motivation

- Space Shuttle *Challenger* accident
- Chernobyl and Three Mile Island accidents
- product liability
- customer goodwill
- corporate reputation

Three closely-related fields of study

- Actuarial science
- Biostatistics
- Reliability engineering

Terminology

The event at the end of a lifetime is called

- a *failure* by reliability engineers
- *death* by actuaries and biostatisticians
- an *epoch* by point process researchers

The object of a study is called

- a *system*, *component* or *item* by reliability engineers
- an *individual* by actuaries
- an *organism* by biostatisticians

To avoid switching terms, *failure* of an *item* will be used here.

1.1 A definition of reliability

Definition 1.1 The *reliability* of an *item* is the *probability* that it will *adequately perform* its specified *purpose* for a specified *period of time* under specified *environmental conditions.*

Item

- Resolution

  an item may be an interacting arrangement of components or the component level of detail in the model may not be of interest

- Level of detail

  determine the level of detail to be modeled

- External boundary for the item

  what is to be considered part of the item and what is to be considered part of the environment around the item

Probability

- Range

  all reliabilities must be between 0 and 1 inclusive

- Spinoffs from the probability axioms

  statistical independence

Adequate performance

- Must be stated unambiguously
- Standards

  Example: a ball bearing has failed when its diameter falls outside of 3 $\pm$ 0.05 mm
- Binary models

  the item is in either the functioning or failed state (e.g., a fuse)

Purpose

- Intended use

  Example: a drill may have one grade for a handyman and another for a contractor

Time

- Units

  must be specified (e.g., hours, years)
- Notation

  many lifetime models use the random variable T
- Time need not be taken literally

  consider an automobile tire, light switch
- Time duration must be specified

  Example: 1000 hour reliability is 0.8
- Continuous operation vs. on/off cycling

  time alone may not be the only consideration (e.g., motors, computers)

Environmental conditions

- Factors

  temperature, humidity, and turning speed all affect the lifetime of a machine tool

- Preventive maintenance

  usually effective in prolonging the lifetime of the item and hence increasing the reliability

Reliability vs. quality

- reliability incorporates the passage of time
- quality is a static descriptor of an item

Example 1: Two transistors of equal quality. One used in a television set, the other in a cannon launch environment. Identical quality, different reliabilities.

Example 2: Two automobile tires, each of high quality. One was produced in 1957, the other in 1994. Same purpose, different reliabilities due to improved design (e.g., tread or steel belts), components (e.g., rubber) or processes (e.g., manufacturing advances). Some quality improvements (e.g., improved tread design) improve the reliability of the tire, while others (e.g., improved white wall design) will not.

1.2 Case study

Item under consideration: the O-rings on the solid rocket motors on the Space Shuttle

Subsystems

- orbiter
- external liquid-fuel tank
- two solid rocket motors

Each assembled solid rocket motor contains three field joints that must be sealed.

O-rings

- 37.5 feet in diameter
- 0.28 inches thick
- all six O-rings must operate to avoid having the propellant escape causing potential failure, so the O-rings form a six-component series system

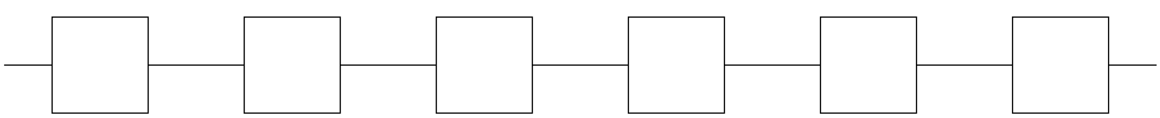

Figure 1.1 A six-component series arrangement of O-rings.

Redundancy: a technique to increase reliability

- Redundancy is highly effective if the components are independent.
- In 1977, NASA discovered *field joint rotation* indicating that the failure of the primary and secondary O-rings may not be independent.
- Prior to the *Challenger* accident, the solid rocket motors were recovered in 23 of the 24 shuttle flights.
- There was concern that an environmental variable, temperature at launch, might influence the reliability of the field joints.
- There was a forecast of 31°F for the morning of the launch of the *Challenger*, the coldest launch temperature to date.

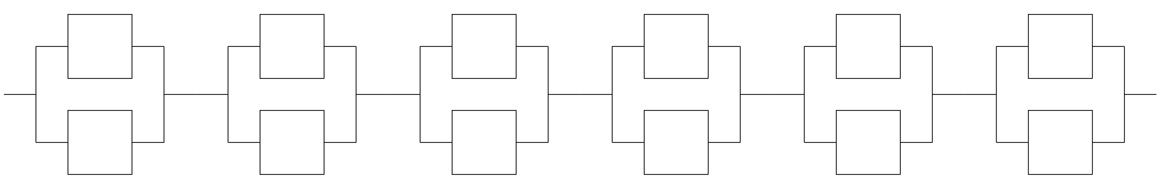

Figure 1.2 A 12-component arrangement of O-rings.

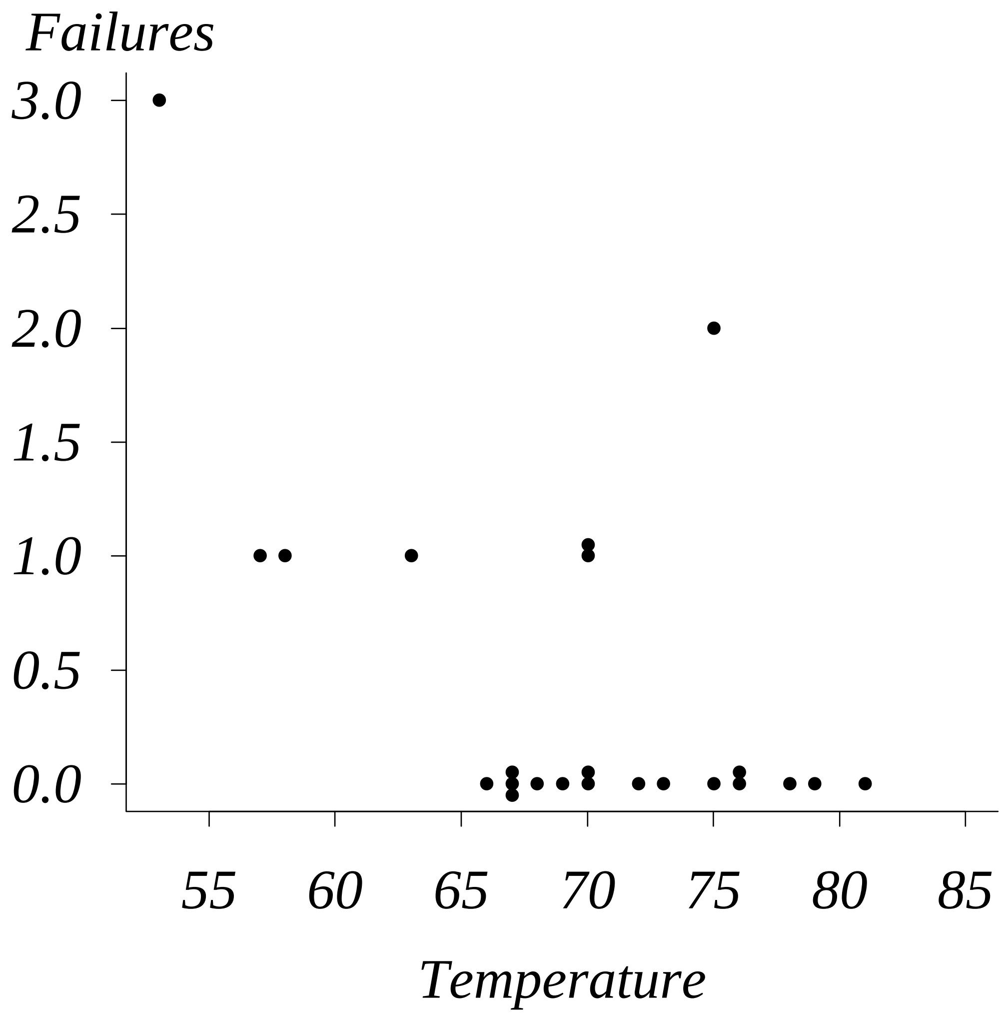

Figure 1.3 Launch temperature versus number of field joint failures.

Conclusion: temperature was indeed significant

2. Coherent Systems Analysis

- assume that an item (system) consists of *n* components

- two key modeling decisions
  which elements of the system are included
  the level of detail
- the first two sections: *structural properties*
- the next two sections: *probabilistic properties*
- outline
  structure functions
  minimal path and cut sets
  reliability functions
  reliability bounds

2.1 Structure functions

Definition 2.1 The state of component i, x_i, is

$$x_i = \begin{cases} 0 & \text{if component } i \text{ has failed} \\ 1 & \text{if component } i \text{ is functioning} \end{cases}$$

for $i = 1, 2, \ldots, n$.

The *binary model*

- n components form a system
- system state vector, $\mathbf{x} = (x_1, x_2, \ldots, x_n)$
- the system state vector can assume 2^n different values
- $\binom{n}{j}$ of these vectors correspond to exactly j functioning components, $j = 0, 1, \ldots, n$

- the structure function, $\phi(\mathbf{x})$, maps the system state vector $\mathbf{x}$ to 0 or 1, the system state

Definition 2.2 The *structure function* ϕ is

$$\phi(\mathbf{x}) = \begin{cases} 0 & \textit{if the system has failed under } \mathbf{x} \\ 1 & \textit{if the system is functioning under } \mathbf{x}. \end{cases}$$

Example 2.1 A *series system* functions when all of its components function. Thus $\phi(\mathbf{x})$ assumes the value 1 when $x_1 = x_2 = \ldots = x_n = 1$, and 0 otherwise.

$$\phi(\mathbf{x}) = \begin{cases} 0 & \textit{if there exists an } i \textit{ such that } x_i = 0 \\ 1 & \textit{if } x_i = 1 \;\; \textit{for all } i = 1, 2, \ldots, n \end{cases}$$
$$= \min\{x_1, x_2, \ldots, x_n\}$$
$$= \prod_{i=1}^{n} x_i.$$

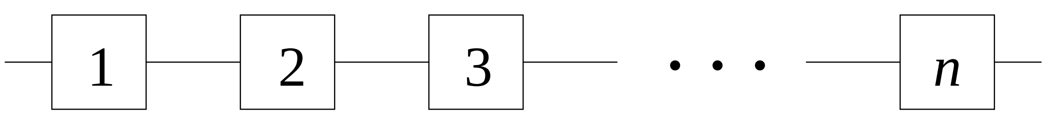

Figure 2.1 A series system block diagram.

Example 2.2 A *parallel system* functions when one or more of the components function. Thus $\phi(\mathbf{x})$ assumes the value 0 when $x_1 = x_2 = \ldots = x_n = 0$, and 1 otherwise.

$$\phi(\mathbf{x}) = \begin{cases} 0 & \textit{if } x_i = 0 \textit{ for all } i = 1, 2, \ldots, n \\ 1 & \textit{if there exists an } i \textit{ such that } x_i = 1 \end{cases}$$
$$= \max\{x_1, x_2, \ldots, x_n\}$$
$$= 1 - \prod_{i=1}^{n} (1 - x_i).$$

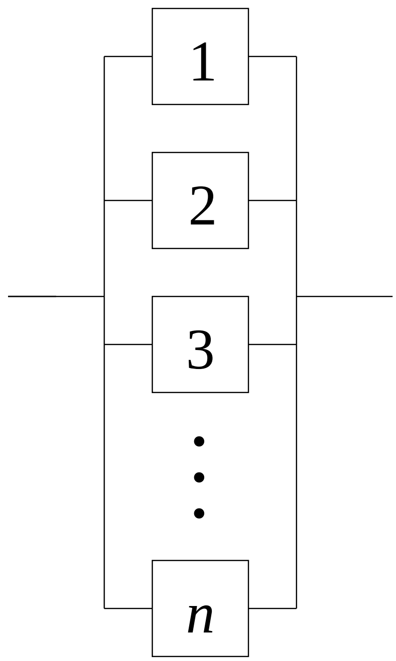

Figure 2.2 A parallel system block diagram.

Applications

- kidneys
- brake system on an automobile with two brake fluid reservoirs

Series and parallel systems are special cases of *k-out-of-n systems*, where the system functions if k or more of the n components function.

Applications

• suspension bridge (components: cables)
• an automobile engine (components: cylinders)
• a bicycle wheel (components: spokes)

Example 2.3 The structure function for a *k-out-of-n* system is

$$\phi(\mathbf{x}) = \begin{cases} 0 & \text{if } \sum_{i=1}^{n} x_i < k \\ 1 & \text{if } \sum_{i=1}^{n} x_i \geq k. \end{cases}$$

The block diagram for a *k-out-of-n* system is difficult to draw in general.

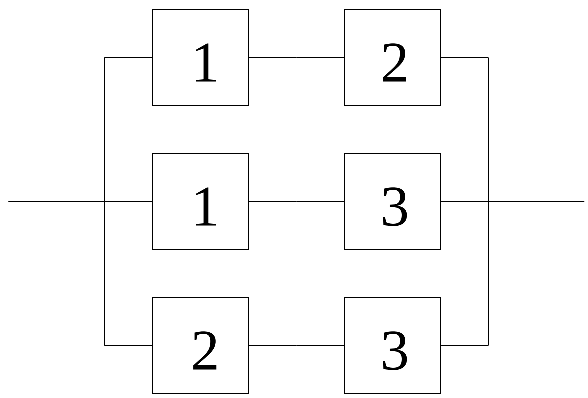

Figure 2.3 A 2-out-of-3 system block diagram.

$$\phi(\mathbf{x}) = 1 - (1 - x_1 x_2)(1 - x_1 x_3)(1 - x_2 x_3)$$

2.3 Reliability functions

Assumptions

- the binary model applies to components and systems
- the n components must be *nonrepairable*
- the components are independent

Definition 2.9 The random variable denoting the state of component i, X_i, is

$$X_i = \begin{cases} 0 & \text{if component } i \text{ has failed} \\ 1 & \text{if component } i \text{ is functioning} \end{cases}$$

for $i = 1, 2, \ldots, n$.

Random component states

- these n values can be written as a random system state vector $\mathbf{X}$
- $p_i = P[X_i = 1]$ is the *reliability* of the i^{th} component, $i = 1, 2, \ldots, n$
- reliability vector $\mathbf{p} = (p_1, p_2, \ldots, p_n)$
- must specify the time to which the reliability applies (e.g., 5000-hour reliability is 0.83)
- the *system reliability*, r, is defined by

$$r(\mathbf{p}) = P[\phi(\mathbf{X}) = 1]$$

- $r(p)$ used when all component reliabilities are equal

Technique 1: Definition of $r(\mathbf{p})$

Example 2.12 *Series* system of n independent components

$$r(\mathbf{p}) = P[\phi(\mathbf{X}) = 1] = P[\prod_{i=1}^{n} X_i = 1] = \prod_{i=1}^{n} P[X_i = 1] = \prod_{i=1}^{n} p_i$$

"Weakest link" for series systems

- system reliability less than smallest component reliability
- improvement of weakest component most effective

Special case: identical components $r(p) = p^n$

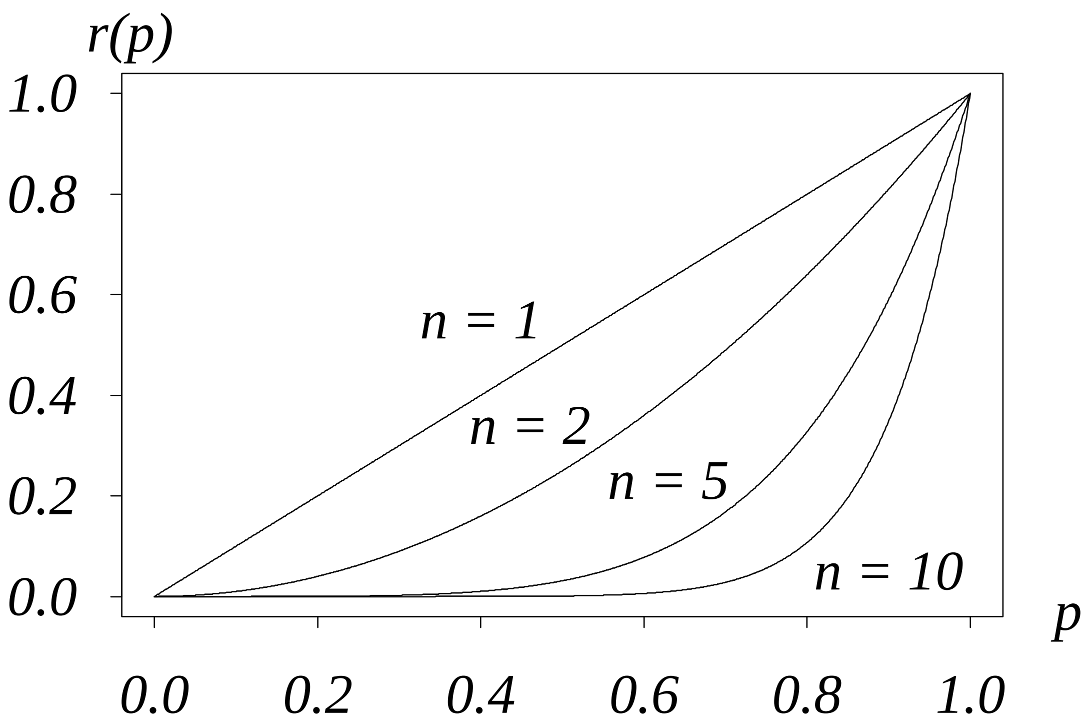

Figure 2.12 Reliability of n-component series systems.

Technique 2: Expected value of $\phi(\mathbf{X})$

$$P[\phi(\mathbf{X}) = 1] = E[\phi(\mathbf{X})]$$

since $\phi(\mathbf{X})$ is a Bernoulli random variable.

Example 2.13 Parallel system of n independent components.

$$r(\mathbf{p}) = E[\phi(\mathbf{X})] = E[1 - \prod_{i=1}^{n} (1 - X_i)]$$
$$= 1 - \prod_{i=1}^{n} E[1 - X_i] = 1 - \prod_{i=1}^{n} (1 - p_i)$$

Special case: identical components $r(p) = 1 - (1 - p)^n$

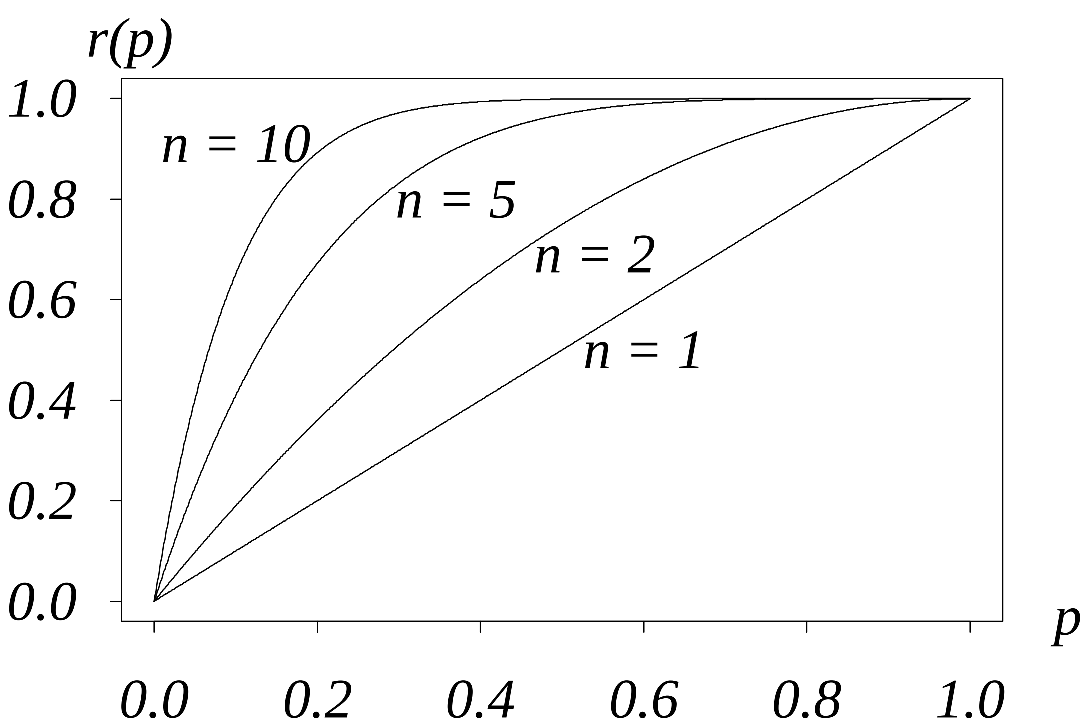

Figure 2.13 Reliability of n-component parallel systems.

"Law of diminishing returns" for parallel systems

- marginal gain in reliability decreases dramatically as more components are added
- improvement of the strongest component is the most effective

Notes on parallel systems

- *standby* system
- *shared-parallel* system

3. Lifetime Distributions

Motivation

Up to this point, reliability has only been considered at one particular instance of time.

Outline

- lifetime distribution representations
- discrete distributions
- moments and fractiles
- system lifetime distributions
- distribution classes

3.1 Distribution representations

Five functions that define the distribution of *T*

- survivor function
- probability density function
- hazard function

- cumulative hazard function
- mean residual lifetime function

Survivor function (reliability function)

$$S(t) = P[T \geq t] \qquad t \geq 0$$

All survivor functions satisfy three conditions

$$S(0) = 1 \qquad \lim_{t \to \infty} S(t) = 0 \qquad S(t) \textit{ is nonincreasing}$$

Interpretations

- $S(t)$ is the probability that an individual item is functioning at time t
- $S(t)$ is the expected fraction of items surviving to time t

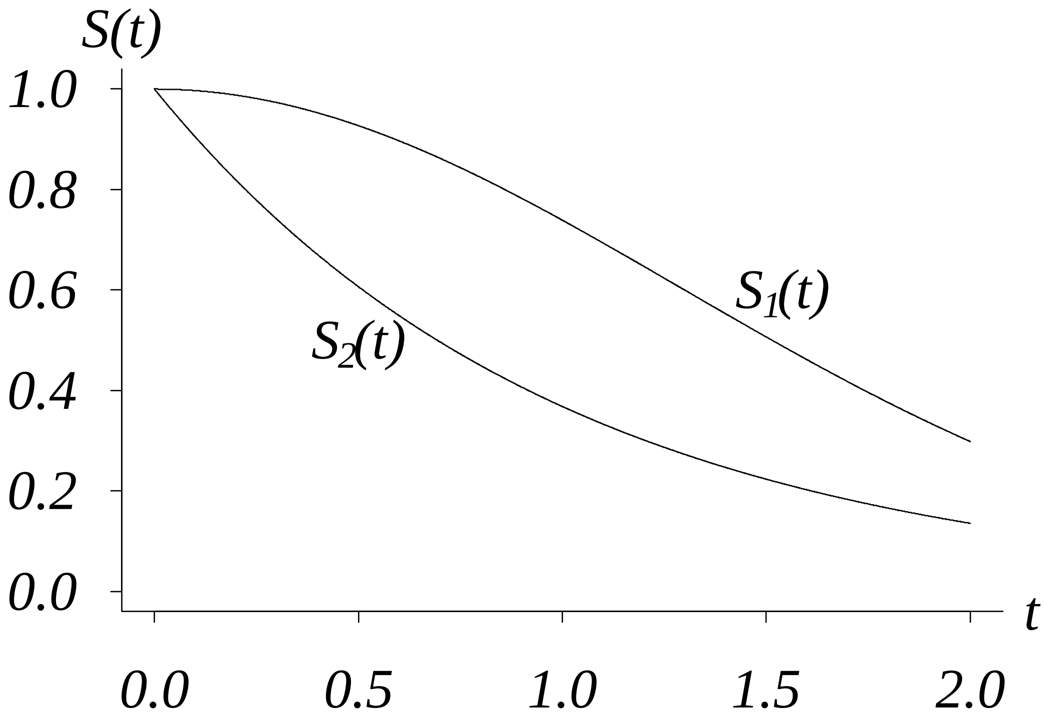

Figure 3.1 Two survivor functions.

Conditional survivor functions

$$S_{T \mid T \geq a}(t) = \frac{P[T \geq t \text{ and } T \geq a]}{P[T \geq a]} = \frac{P[T \geq t]}{P[T \geq a]} = \frac{S(t)}{S(a)}$$

for all $t \geq a$.

Probability density function

$$f(t) = -S'(t)$$
$$f(t)\Delta t = P[t \le T \le t + \Delta t]$$

for small Δt values.

$$P[a \le T \le b] = \int_a^b f(t)\,dt = S(a) - S(b)$$

All probability density functions satisfy

$$\int_0^\infty f(t)\,dt = 1 \qquad\qquad f(t) \ge 0 \; for\ all\ t \ge 0$$

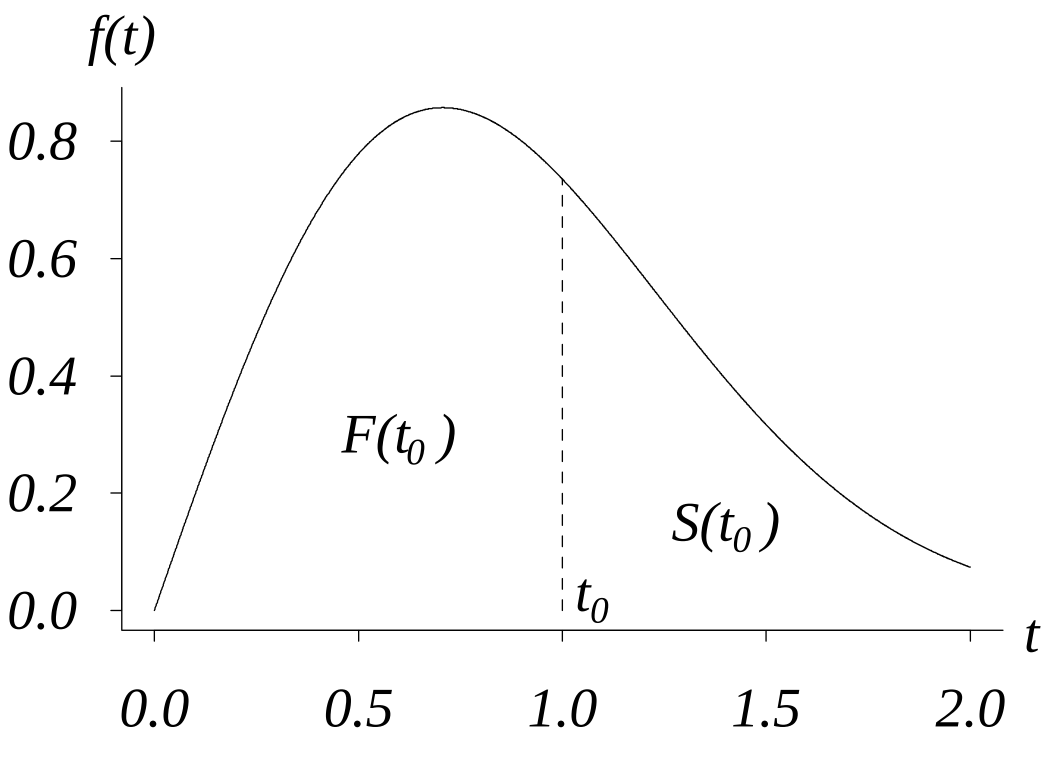

Figure 3.3 The relationship between survivor and cumulative distribution functions.

Hazard function (failure rate, force of mortality)

$$h(t) = f(t)\,/\,S(t) \qquad\qquad t \ge 0$$

$$h(t)\Delta t = P[t \le T \le t + \Delta t \mid T \ge t]$$

for small Δt values. Units: failures per unit time.

Interpretations

- $h(t)$ is the amount of *risk* an item is under at t
- $h(t)$ is a special case of the intensity function for a nonhomogeneous Poisson process

All hazard functions must satisfy

$$\int_0^\infty h(t)\,dt = \infty \qquad\qquad h(t) \ge 0 \textit{ for all } t \ge 0$$

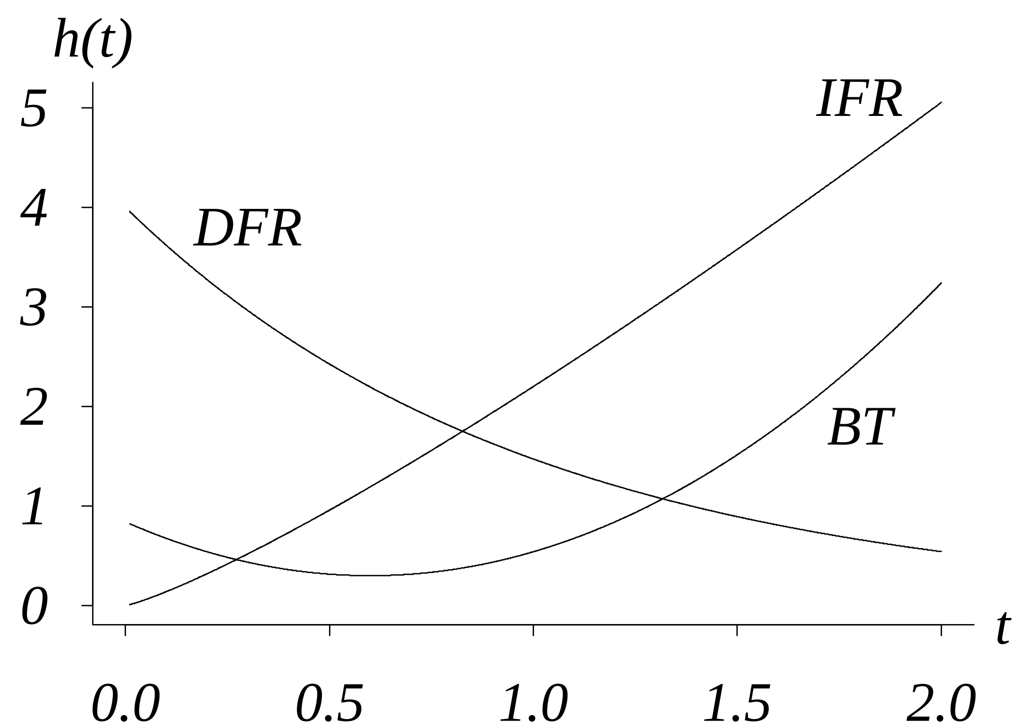

Figure 3.5 Common hazard function shapes.

Cumulative hazard function (integrated hazard function and the renewal function)

$$H(t) = \int_0^t h(\tau)d\tau \qquad t \ge 0$$

All cumulative hazard functions satisfy

$$H(0) = 0 \qquad \lim_{t \to \infty} H(t) = \infty \qquad H(t) \textit{ is nondecreasing}$$

Applications

- variate generation in Monte Carlo simulation
- implementing certain procedures in statistical inference
- defining certain distribution classes

Mean residual life function

$$L(t) = E[T - t \mid T \ge t] = \frac{1}{S(t)} \int_t^\infty \tau\, f(\tau)d\tau - t \qquad t \ge 0$$

All mean residual life functions satisfy

$$L(t) \ge 0 \qquad L'(t) \ge -1 \qquad \int_0^\infty \frac{dt}{L(t)} = \infty$$

Example 3.2 Consider the exponential distribution defined by the survivor function

$$S(t) = e^{-\lambda t} \qquad t \geq 0$$

with positive scale parameter λ.

$$f(t) = \lambda e^{-\lambda t} \qquad t \geq 0$$

The mean residual life function is

$$L(t) = e^{\lambda t} \int_t^\infty \tau \lambda e^{-\lambda \tau} d\tau - t = \frac{1}{\lambda} \qquad t \geq 0$$

by using integration by parts.

Knowing one of the five lifetime distribution representations implies knowledge of the other four.

If the survivor function is known, for example, the cumulative hazard function can be determined by

$$H(t) = \int_0^t h(\tau) d\tau = \int_0^t \frac{f(\tau)}{S(\tau)} d\tau = -\log S(t)$$

3.3 Moments and fractiles

Motivation

Moments and fractiles contain less information than a lifetime distribution representation, but they are often useful ways to summarize the distribution of a random lifetime.

Examples

- the mean time to failure, $E(T)$
- the median, $t_{0.50}$
- the 95^{th} percentile of a distribution, $t_{0.95}$

Assumption: random lifetime T is continuous

$$E[u(T)] = \int_0^\infty u(t)\ f(t)\,dt$$

Mean (abbreviated by MTTF or MTBF)

$$\mu = E[T] = \int_0^\infty t\ f(t)\,dt = \int_0^\infty S(t)\,dt$$

Variance

$$\sigma^2 = V[T] = E[(T-\mu)^2] = E[T^2] - (E[T])^2$$

Coefficient of variation

$$\gamma = \frac{\sigma}{\mu}$$

Skewness

$$\gamma_3 = E\left[\left(\frac{T-\mu}{\sigma}\right)^3\right]$$

Kurtosis

$$\gamma_4 = E\left[\left(\frac{T-\mu}{\sigma}\right)^4\right]$$

Fractiles: t_p satisfies

$$F(t_p) = P[T \le t_p] = p \qquad or \qquad t_p = F^{-1}(p)$$

Example 3.5 The exponential distribution has survivor function

$$S(t) = e^{-\lambda t} \qquad\qquad t \ge 0$$

$$\mu = E[T] = \int_0^\infty S(t)\,dt = \int_0^\infty e^{-\lambda t}\,dt = \frac{1}{\lambda}$$

$$E[T^2] = \int_0^\infty t^2\, f(t)\,dt = \int_0^\infty t^2\,\lambda\, e^{-\lambda t}\,dt = \frac{2}{\lambda^2}$$

$$\sigma^2 = E[T^2] - (E[T])^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

$$\gamma_3 = \lambda^3\left[6\lambda^{-3} - 6\lambda^{-3} + 3\lambda^{-3} - \lambda^{-3}\right] = 2$$

$$t_p = -\frac{1}{\lambda}\log(1-p)$$

4. Parametric Lifetime Models

Motivation
Survival patterns of a drill bit, a fuse, and an automobile are vastly different.

Outline

- parameters
- exponential
- Weibull
- gamma
- other distributions

4.1 Parameters

Three types of parameters:

- location
- scale
- shape

Location (or shift) parameters
Shift a distribution along the time axis. If c_1 and c_2 are two values of a location parameter for a lifetime distribution with survivor function $S(t; c)$, then there exists a constant α such that $S(t; c_1) = S(\alpha + t; c_2)$.

Example Mean μ in the normal distribution.

Scale parameters

Used to expand or contract the time axis by a factor of α. If λ_1 and λ_2 are two values for a scale parameter for a lifetime distribution with survivor function $S(t;\lambda)$, then there exists a constant α such that $S(\alpha t;\lambda_1) = S(t;\lambda_2)$.

Example Exponential scale parameter λ.

Shape parameters

Affect the shape of the probability density function.

Example Weibull shape parameter κ.

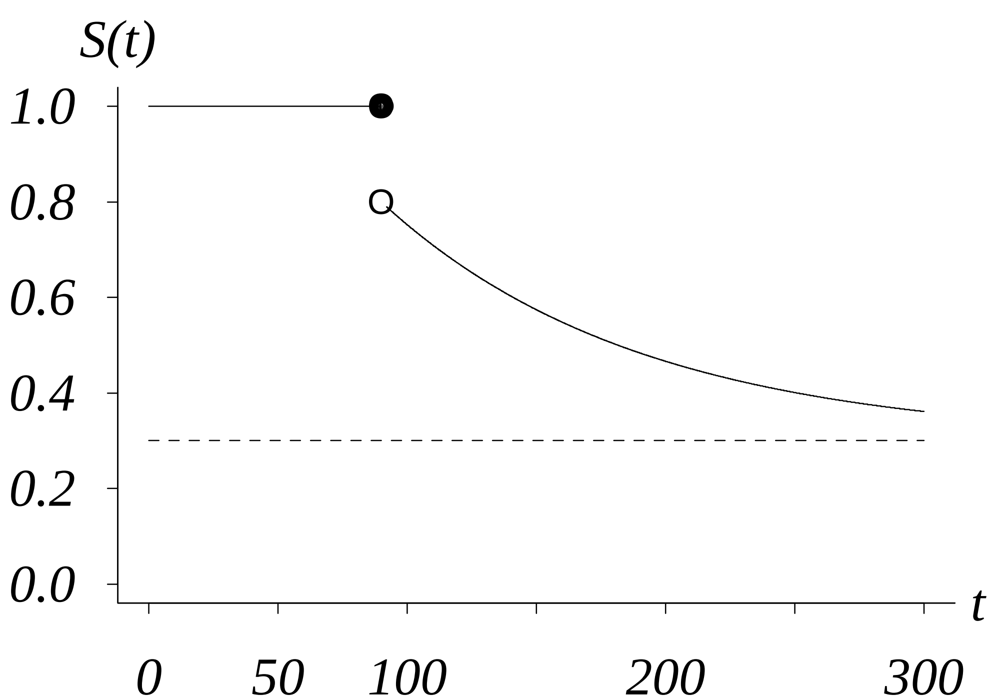

Figure 4.2 A mixed discrete-continuous

survivor function.

4.2 The exponential distribution

Motivation

The exponential distribution plays a central role in reliability modeling since it is the only continuous distribution with a constant hazard function.

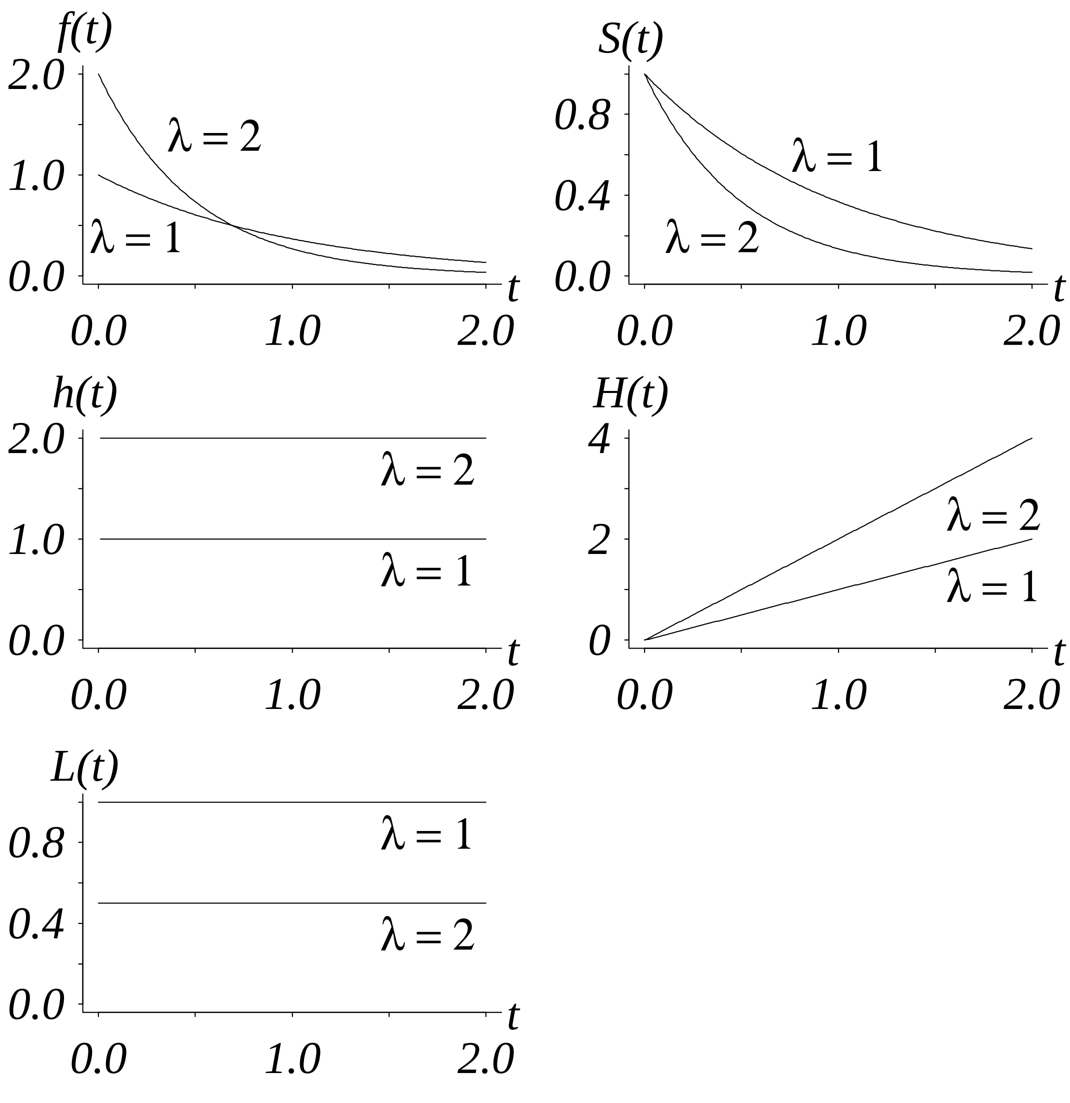

Figure 4.3 Lifetime distribution representations for the exponential distribution.

$$S(t) = e^{-\lambda t} \qquad f(t) = \lambda e^{-\lambda t} \qquad h(t) = \lambda$$

$$H(t) = \lambda t \qquad L(t) = \frac{1}{\lambda}$$

Property 4.1 (Memoryless property) If $T \sim \text{exponential}(\lambda)$ then

$$P[T \geq t] = P[T \geq t + s \,|\, T \geq s] \qquad t \geq 0;\ s \geq 0.$$

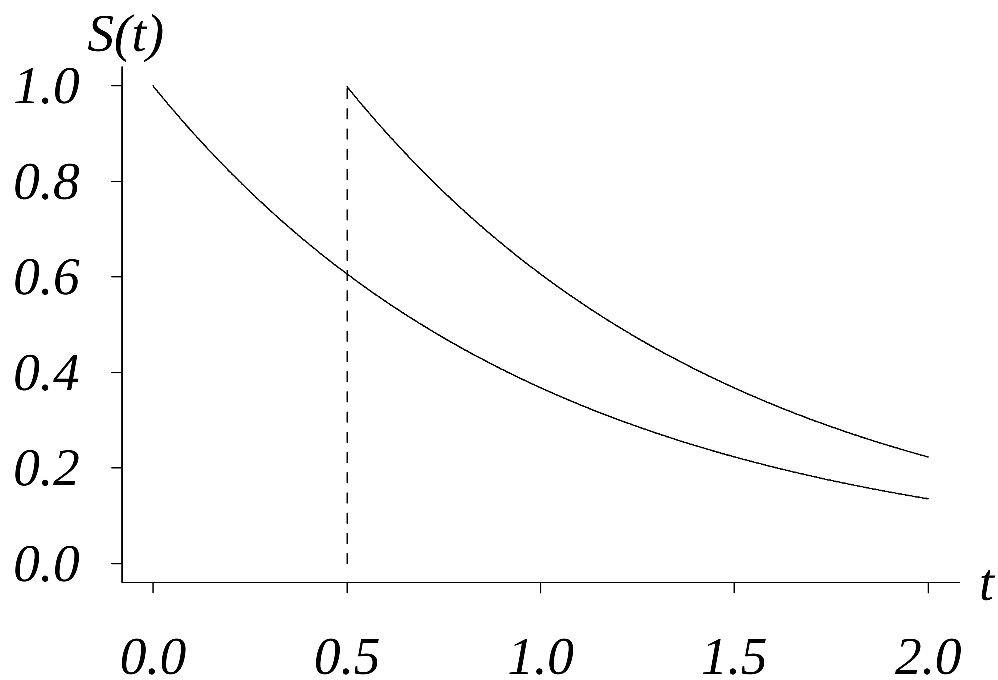

Figure 4.4 The memoryless property of the exponential distribution.

Property 4.2 The exponential distribution is the only continuous distribution with the memoryless property.

4.3 The Weibull distribution

Motivation

The exponential distribution's constant failure rate is often too restrictive or inappropriate.

$$S(t) = e^{-(\lambda t)^{\kappa}} \qquad f(t) = \kappa \lambda^{\kappa} t^{\kappa - 1} e^{-(\lambda t)^{\kappa}}$$

$$h(t) = \kappa \lambda^{\kappa} t^{\kappa - 1} \qquad H(t) = (\lambda t)^{\kappa}$$

for all $t \geq 0$.

Notes

- λ is a positive scale parameter
- κ is a positive shape parameter
- exponential distribution is a special case ($\kappa = 1$)
- hazard function increases from 0 when $\kappa > 1$ (IFR)
- hazard function decreases from ∞ to 0 when $\kappa < 1$ (DFR)
- $\kappa = 2$ known as the Rayleigh distribution
- when $3 < \kappa < 4$ the probability density function resembles that of a normal random variable
- the mode and median of the distribution are equal when $\kappa \approx 3.26$
- the *characteristic life* is a special fractile defined by $t_c = \frac{1}{\lambda}$; all Weibull survivor functions pass through the point $(\frac{1}{\lambda}, e^{-1})$

- since $H(t) = -\log S(t)$, all Weibull cumulative hazard functions pass through the point $(\frac{1}{\lambda}, 1)$
- if T has the Weibull distribution, then $Y = \log T$ has the *extreme value* distribution
- self-reproducing property: if $T_i \sim \text{Weibull}(\lambda_i, \kappa)$ for $i = 1, 2, \ldots, n$, then $\min\{T_1, T_2, \ldots, T_n\} \sim \text{Weibull}(\sum_{i=1}^{n} \lambda_i, \kappa)$
- moments

$$\mu = \frac{1}{\lambda}\Gamma\left(1 + \frac{1}{\kappa}\right) = \frac{1}{\lambda\kappa}\Gamma\left(\frac{1}{\kappa}\right)$$

$$\sigma^2 = \frac{1}{\lambda^2}\left\{\frac{2}{\kappa}\Gamma\left(\frac{2}{\kappa}\right) - \left[\frac{1}{\kappa}\Gamma\left(\frac{1}{\kappa}\right)\right]^2\right\}$$

$$\gamma = \frac{\left\{\frac{2}{\kappa}\Gamma\left(\frac{2}{\kappa}\right) - \left[\frac{1}{\kappa}\Gamma\left(\frac{1}{\kappa}\right)\right]^2\right\}^{1/2}}{\frac{1}{\kappa}\Gamma\left(\frac{1}{\kappa}\right)}$$

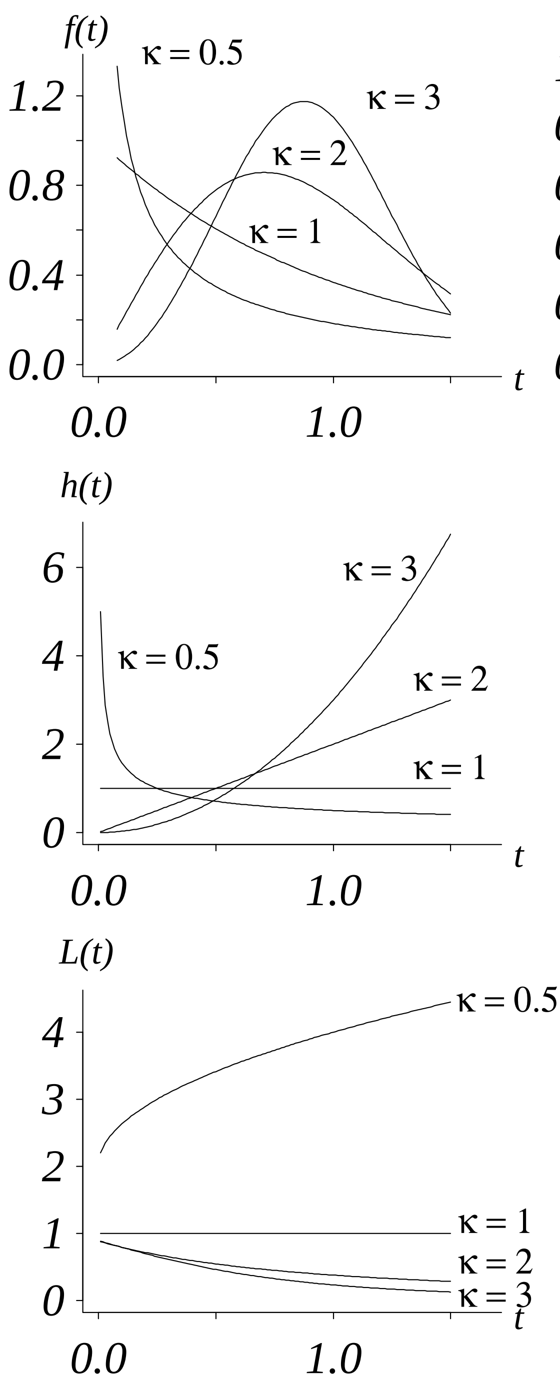

Figure 4.7 Lifetime distribution representations for the Weibull distribution.

4.5 Other lifetime distributions

Table 4.4 Distribution classes.

Distribution	IFR	DFR	BT	UBT
Exponential	**YES**	**YES**	**NO**	**NO**
Muth	**YES**	**NO**	**NO**	**NO**
Weibull	**YES**	**YES**	**NO**	**NO**
Gamma	**YES**	**YES**	**NO**	**NO**
Uniform	**YES**	**NO**	**NO**	**NO**
Log normal	**NO**	**NO**	**NO**	**YES**
Log logistic	**NO**	**YES**	**NO**	**YES**
Inv. Gaussian	**NO**	**NO**	**NO**	**YES**
Expon. power	**YES**	**NO**	**YES**	**NO**
Pareto	**NO**	**YES**	**NO**	**NO**
Gompertz	**YES**	**NO**	**NO**	**NO**
Makeham	**YES**	**NO**	**NO**	**NO**
IDB	**YES**	**YES**	**YES**	**NO**
Gen. Pareto	**YES**	**YES**	**NO**	**NO**

5. Specialized Models

Motivation

There are several ways to combine and extend the continuous lifetime models previously outlined.

Outline

- competing risks
- mixtures
- accelerated life
- proportional hazards

5.1 Competing risks

Notes

- causes of failure may be grouped into k classes
- an item is subject to k competing risks (or causes) $C_1, C_2, \ldots, C_k$
- can be thought of as a series system of components
- origins of competing risks theory traced to a study by Daniel Bernoulli in the 1700's concerning the impact of eliminating smallpox
- a second and equally appealing use of competing risks models is that they can be used to combine component distributions to

form more complicated models

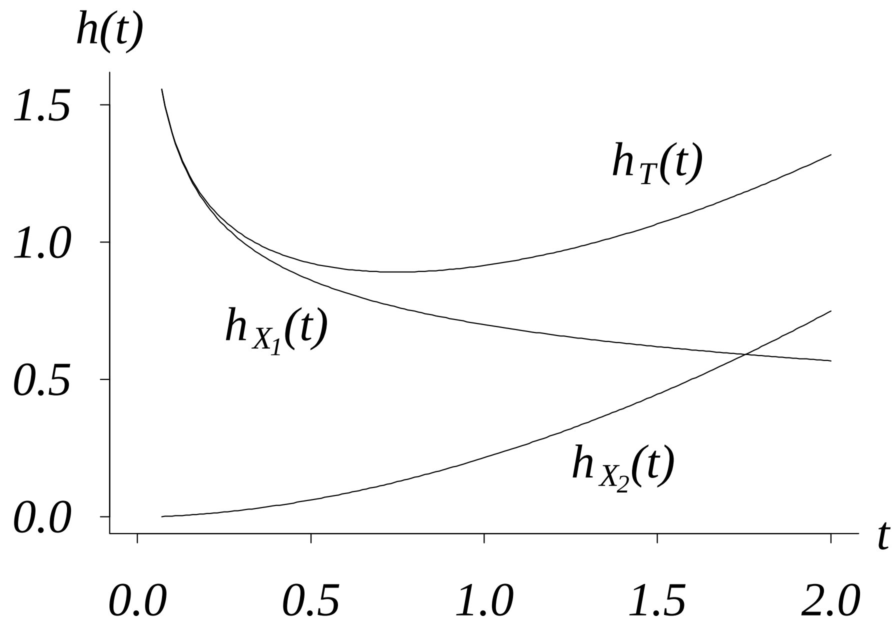

Figure 5.1 Hazard functions for a competing risks model.

- notation (for $j = 1, 2, \ldots, k$)

T	lifetime
k	number of risks
X_j	net life for risk j
Y_j	crude life for risk j
π_j	probability risk j causes failure

- *net lifetimes:* causes $C_1, \ldots, C_k$ are viewed individually
- *crude lifetimes:* lifetimes are considered in the presence of all other risks

5.2 Mixtures

Mixture models are appropriate when items ar drawn from one of several populations (finite mixtures) or can be differentiated by a continuous parameter.

Finite mixtures

$$f(t) = \sum_{l=1}^{m} p_l \, f_l(t \mid \theta_l)$$

where $\sum_{l=1}^{m} p_l = 1$, $p_l \ge 0$ for $l = 1, 2, \ldots, m$.

Continuous mixtures (stochastic parameters)

$$f(t) = \int_{all\theta} f(t|\theta) \, p(\theta) d\theta$$

where θ is called the *mix parameter* and $p(\theta)$ indicates the distribution of the mix parameter.

> **Example 5.4** If $m = 2$ facilities produce items with exponential(1) and exponential(2) lifetimes, respectively, and $1/3$ of the items come from facility 1 and $2/3$ come from facility 2, the probability density function of the time to failure of an item whose manufacturing site is unknown is
>
> $$f(t) = p_1 \, f_1(t \mid \lambda_1) + p_2 \, f_2(t \mid \lambda_2)$$

$$= \frac{1}{3} e^{-t} + \frac{4}{3} e^{-2t} \qquad t \geq 0$$

which is a finite mixture of the two populations. This model is a special case of the *hyperexponential* distribution.

Combining competing risks and finite mixtures

$$f(t) = \sum_{l=1}^{m} p_l \left[\sum_{j=1}^{k_l} h_{lj}(t)\, e^{-\int_0^t \sum_{j=1}^{k_l} h_{lj}(\tau) d\tau} \right]$$

where m is the number of populations, $\sum_{l=1}^{m} p_l = 1$, k_l is the number of risks acting within the l^{th} population, $h_{lj}(t)$ is the hazard function for the j^{th} risk within the l^{th} population.

Application: casualty insurance

- $m = 3$ populations of dwellings
  - single family dwellings
  - condominiums
  - apartments
- $k_1 = k_2 = k_3 = 5$ risks
  - fire
  - flood
  - tornado
  - earthquake

burglary

5.3 Accelerated life

The *accelerated life* and *proportional hazards* models are appropriate for including a vector of covariates in a lifetime model.

The $q \times 1$ vector $\mathbf{z} = (z_1, z_2, \ldots, z_q)'$ contains q covariates associated with a particular item.

Example Reliability
T: drill bit failure time
z_1: turning speed
z_2: feed rate
z_3: hardness of the material

Example Biostatistics
T: patient survival time
z_1: age
z_2: gender
z_3: cholesterol level

Example Recidivism
T: time to return to prison
z_1: age
z_2: time served
z_3: number of previous convictions

Notation

$\mathbf{z} = (z_1, z_2, \ldots, z_q)'$	covariates
$\beta = (\beta_1, \beta_2, \ldots, \beta_q)'$	regression coefficients
$\psi(\mathbf{z})$	link function
$S_0(t), f_0(t), h_0(t), H_0(t)$	baseline functions

How to link covariates to a lifetime distribution

- one lifetime model when $\mathbf{z} = \mathbf{0}$ (often called the *baseline* model)
- other models when $\mathbf{z} \neq \mathbf{0}$

The accelerated life model

$$S(t) = S_0(t\,\psi(\mathbf{z})) \qquad t \geq 0$$

Notes

- S_0 is a baseline survivor function
- $\psi(\mathbf{z})$ is a *link function* satisfying $\psi(\mathbf{0}) = 1$ and $\psi(\mathbf{z}) > 0$ for all $\mathbf{z}$
- a popular link function choice is the log-linear form $\psi(\mathbf{z}) = e^{\beta'\mathbf{z}}$
- the covariates accelerate the rate at which the item moves through time with respect to the baseline case when $\psi(\mathbf{z}) > 1$
- the covariates decelerate the rate at which the item moves through time with respect to the baseline case when $\psi(\mathbf{z}) < 1$
- application: situations when testing items at their operating environments is too time

consuming

5.4 Proportional hazards

Whereas accelerated life models modify the rate that the item moves through time based on the values of the covariates, proportional hazards models modify the hazard function by the factor $\psi(\mathbf{z})$.

The *proportional hazards* model can be defined by

$$h(t) = \psi(\mathbf{z})\, h_0(t).$$

Notes

- the covariates increase the risk when $\psi(\mathbf{z}) > 1$
- the covariates decrease the risk when $\psi(\mathbf{z}) < 1$
- the log-linear form $\psi(\mathbf{z}) = e^{\beta'\mathbf{z}}$ is still an appropriate choice for the link function

Table 5.1 Lifetime distribution representations for regression models.

	Accelerated Life	Proportional Hazards
$S(t)$	$S_0(t\psi(\mathbf{z}))$	$[S_0(t)]^{\psi(\mathbf{z})}$
$f(t)$	$\psi(\mathbf{z})\, f_0(t\psi(\mathbf{z}))$	$f_0(t)\psi(\mathbf{z})[S_0(t)]^{\psi(\mathbf{z})-1}$
$h(t)$	$\psi(\mathbf{z})h_0(t\psi(\mathbf{z}))$	$\psi(\mathbf{z})h_0(t)$
$H(t)$	$H_0(t\psi(\mathbf{z}))$	$\psi(\mathbf{z})\, H_0(t)$

Example 5.11 Consider the baseline hazard function (Cox and Oakes, 1984)

$$h_0(t) = \begin{cases} 1 & 0 \le t < 1 \\ t & t \ge 1 \end{cases}$$

Assumptions

- single binary covariate z
- when $z = 0$ (the control case), $\psi(z) = 1$
- when $z = 1$ (the treatment case), $\psi(z) = 2$

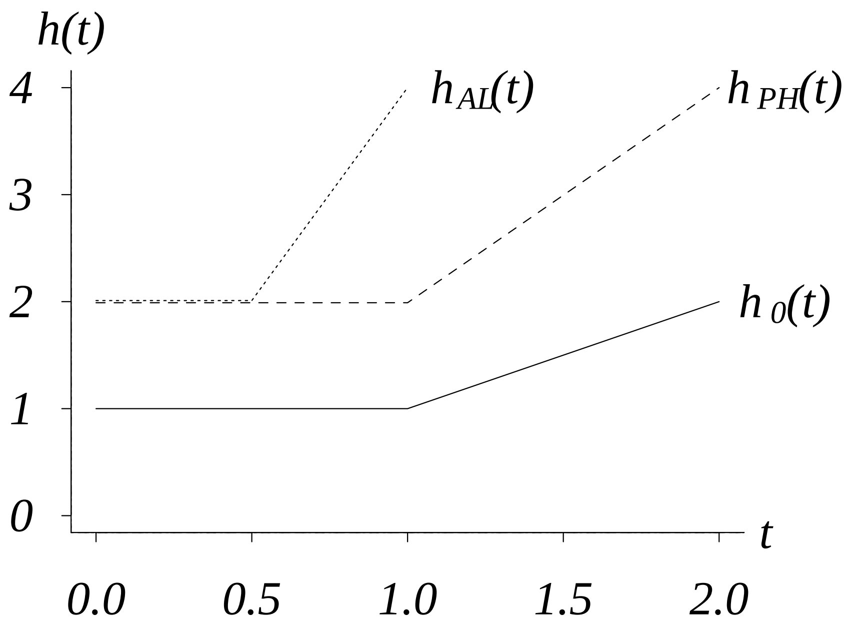

Figure 5.3 Hazard functions for a piecewise-continuous baseline hazard function.

$$h_{PH}(t) = \psi(z)\ h_0(t) \qquad\qquad t \geq 0$$

$$h_{AL}(t) = \psi(z) h_0(t\psi(z)) \qquad\qquad t \geq 0$$

6. Repairable Systems

Motivation

So far, only nonrepairable systems of components have been considered. Most systems are repairable.

Outline

- Introduction
- Point processes
- Availability

- Birth-death processes

6.1 Introduction

A *repairable item* may be returned to an operating condition after failure to perform a required function by any method other than replacement of the entire item.

Replacement models

- used when a nonrepairable item is replaced with another item upon failure
- "socket models"
- unlimited spares
- redundancy allocation problem (optimal number of spares)
- replacement policies

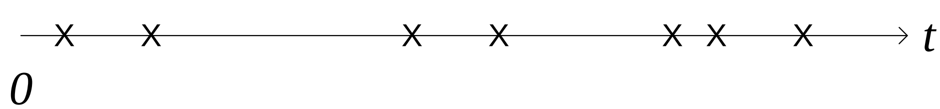

Figure 6.1 Failure replacement policy.

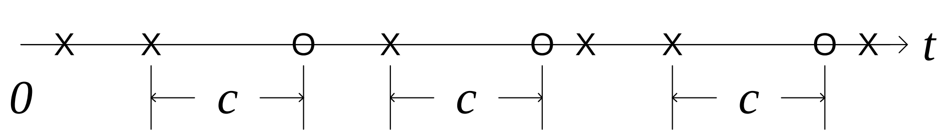

Figure 6.2 Age replacement policy.

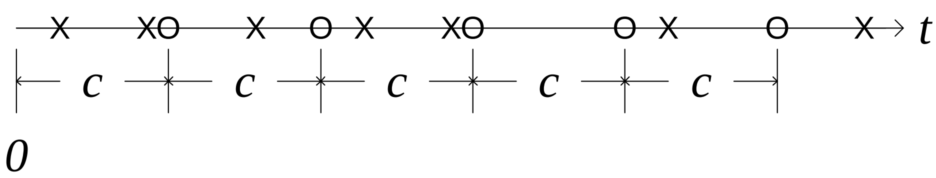

Figure 6.3 Block replacement policy.

- choice between these three replacement policies depends on the lifetime distribution, the cost of failure, administrative costs, etc.
- age and block replacement policies collapse to a failure replacement policy as $c \to \infty$
- expected number of items consumed (c fixed)

$$n_f(t) \le n_a(t) \le n_b(t) \qquad\qquad t > 0$$

6.2 Point processes

Hazard vs. intensity functions

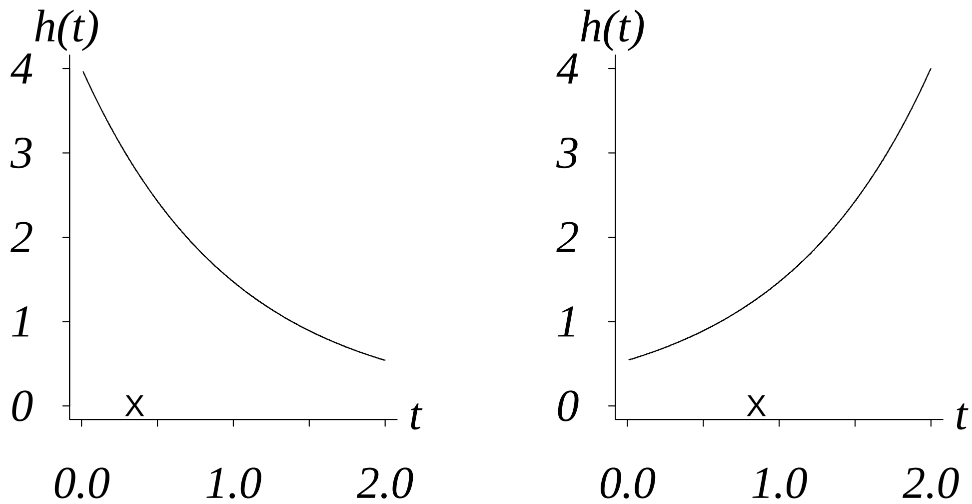

Figure 6.4 Hazard functions for an item with a DFR and IFR distribution.

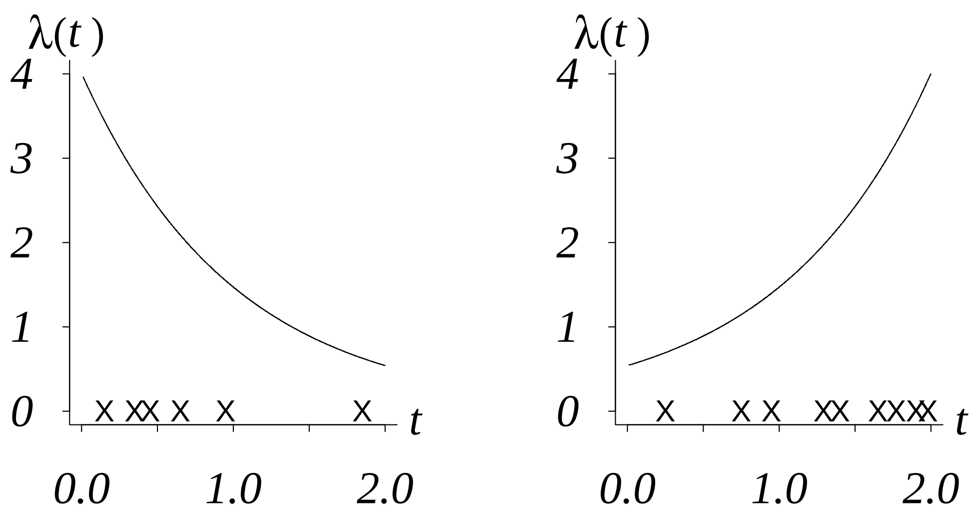

Figure 6.5 Intensity functions for an improving item and a deteriorating item.

Table 6.1 Terminology for nonrepairable and repairable items.

Status	Nonrepairable	Repairable
Gets better	burn-in $h'(t) \leq 0$	improving $\lambda'(t) \leq 0$
Gets worse	wear out $h'(t) \geq 0$	deteriorating $\lambda'(t) \geq 0$

Point process models

- Poisson processes
- renewal processes
- nonhomogeneous Poisson processes

Notation and assumptions

- failures occur at times $T_1, T_2, \ldots$
- the time to replace or repair an item is negligible
- the origin is defined to be $T_0 = 0$
- the times between the failures are $X_1, X_2, \ldots$
- $T_k = X_1 + X_2 + \ldots + X_k$, for $k = 1, 2, \ldots$
- the counting function $N(t)$ is the number of failures that occur in $(0, t]$

$$N(t) = \max\{\, k \mid T_k \leq t\}$$

  for $t > 0$
- $\{N(t), t > 0\}$ is often called a "counting process"

  * if $t_1 < t_2$ then $N(t_1) \leq N(t_2)$

* if $t_1 < t_2$ then $N(t_2) - N(t_1)$ is the number of failures in the interval $(t_1, t_2]$

- $\Lambda(t) = E[N(t)]$ is the expected number of failures that occur in the interval $(0, t]$
- $\lambda(t) = \Lambda'(t)$ is the rate of occurrence of failures

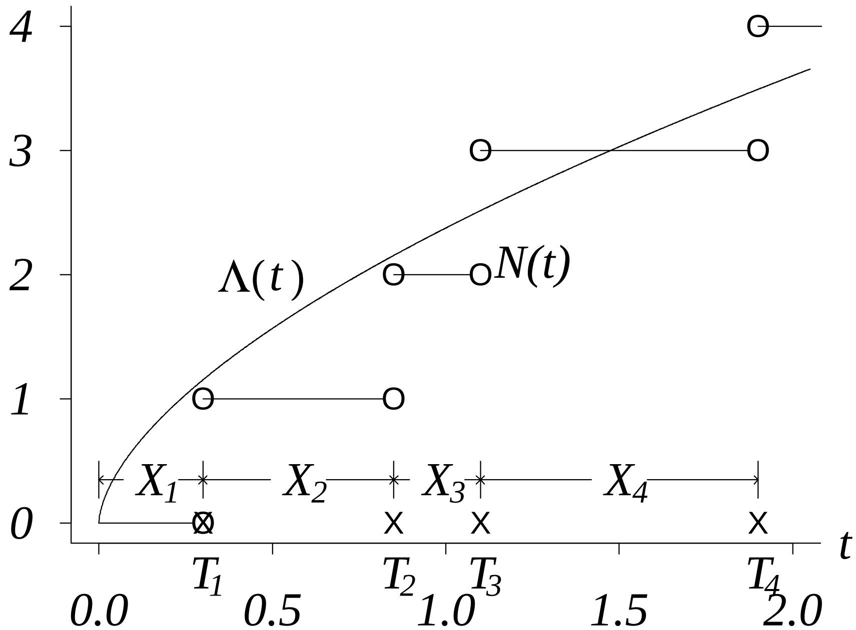

Figure 6.6 A point process realization.

Two important properties

- *independent increments:* the number of failures in mutually exclusive intervals are independent

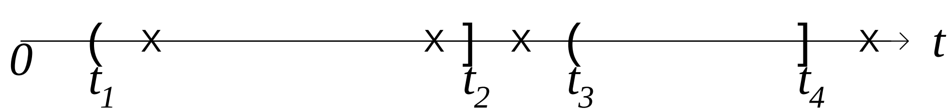

Figure 6.7 Independent increments.

- *stationarity:* the distribution of the number of failures in any time interval depends only on the length of the time interval

Homogeneous Poisson process (HPP)

Definition 6.1 A counting process is a Poisson process with parameter $\lambda > 0$ if

- $N(0) = 0$
- the process has independent increments
- the number of failures in any interval of length t has the Poisson distribution with parameter λt.

Implications

- the distribution of the number of events in $(t_1, t_2]$ has the Poisson distribution with parameter $\lambda(t_2 - t_1)$.
- $P[N(t_2) - N(t_1) = x] = \dfrac{[\lambda(t_2 - t_1)]^x e^{-\lambda(t_2 - t_1)}}{x!}$ for $x = 0, 1, 2, \ldots$
- $N(t)$ has the Poisson distribution with mean $\Lambda(t) = E[N(t)] = \lambda t$, where λ is often called the rate of occurrence of failures
- the intensity function is $\lambda(t) = \Lambda'(t) = \lambda$
- if $X_1, X_2, \ldots$ are independent and identically distributed exponential random variables,

then $N(t)$ corresponds to a Poisson process

- this model is sometimes called just a *Poisson process*

Nonhomogeneous Poisson process (NHPP)

Four reasons to consider an NHPP

- the HPP is a special case of an NHPP (stationarity assumption relaxed)
- the probabilistic model for an NHPP is mathematically tractable
- the statistical methods for an NHPP are also mathematically tractable
- the NHPP is capable of modeling improving and deteriorating systems

Intensity function: $\lambda(t)$

Cumulative intensity function: $\Lambda(t) = \int_0^t \lambda(\tau)d\tau$

Definition 6.4 A counting process is a nonhomogeneous Poisson process with intensity function $\lambda(t) \geq 0$ if

- $N(0) = 0$
- the process has independent increments
- the probability of exactly n events occurring in the interval $(a, b]$ is given by

$$P[N(b) - N(a) = n] = \frac{[\int_a^b \lambda(t)dt]^n \, e^{-\int_a^b \lambda(t)dt}}{n!}$$

for $n = 0, 1, \ldots.$

6.3 Availability

Notation

- X_i denotes the i^{th} time to failure, $i = 1, 2, \ldots$
- R_i denotes the i^{th} time to repair, $i = 1, 2, \ldots$

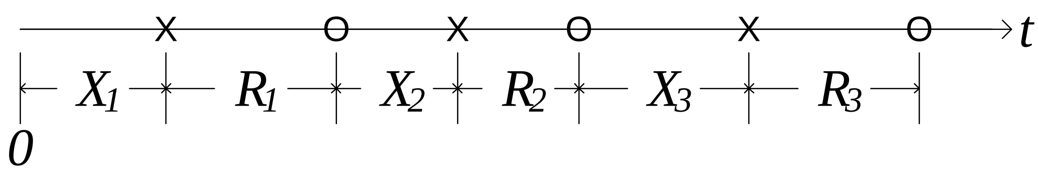

Figure 6.10 Failure and repair process realization.

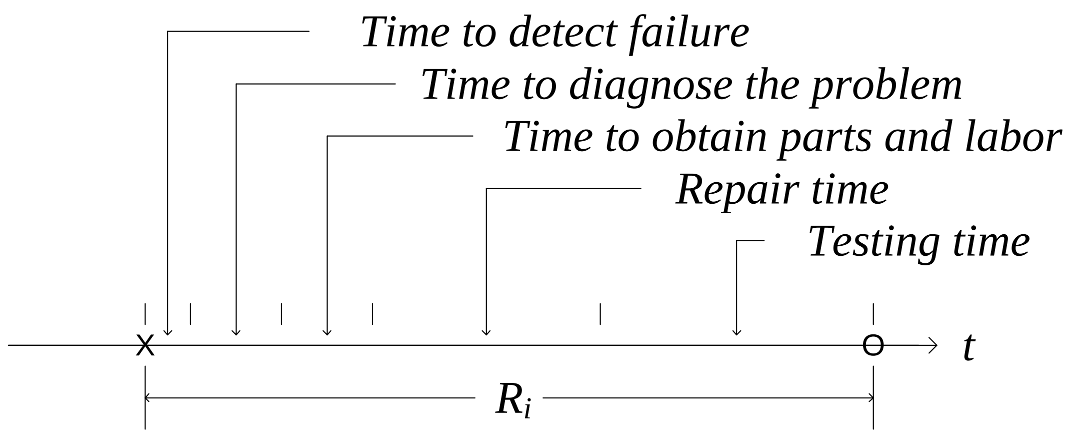

Figure 6.11 Partitioning the repair time.

7. Lifetime Data Analysis

Motivation

Parameters have been assumed to be known constants. The rest of the tutorial considers parameter estimation (e.g., component reliability, distribution parameter values).

Outline

- point estimation
- interval estimators
- likelihood function
- asymptotic properties of the likelihood function
- censoring

7.1 Point estimation

A *point estimator* is a statistic used to estimate a population parameter.

Definition 7.1 The point estimator $\hat{\theta}$ is an *unbiased estimator* of θ if and only if $E[\hat{\theta}] = \theta$.

Definition 7.2 Let $\hat{\theta}_1$ and $\hat{\theta}_2$ be two unbiased point estimators of the parameter θ. Then

$$\frac{V(\hat{\theta}_1)}{V(\hat{\theta}_2)}$$

is the *efficiency* of $\hat{\theta}_1$ relative to $\hat{\theta}_2$.

7.2 Interval estimation

Confidence intervals give bounds that contain a population parameter with a prescribed probability

$$L \leq \theta \leq U$$

Notes

- L and U are functions of
  the sample size n
  the lifetimes $t_1, \ldots, t_n$
  the nominal coverage of the interval $1 - \alpha$
- true value of the parameter θ is denoted by θ_0
- popular choices for α are 0.10 and 0.05
- confidence intervals: exact, approximate, asymptotically exact

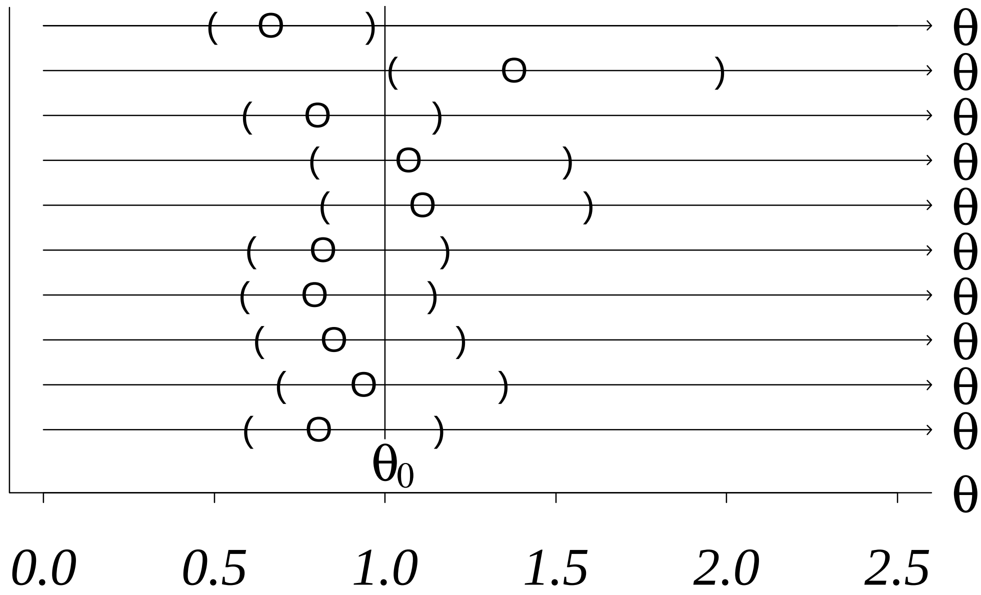

Figure 7.2 Ten 90% confidence intervals for θ ($n = 25$).

7.3 Likelihood theory

Notation

- $t_1, t_2, \ldots, t_n$ is a set of random lifetimes
- $\theta = (\theta_1, \theta_2, \cdots, \theta_p)'$ is a vector of unknown parameters
- $L(\mathbf{t}, \theta)$ is the likelihood function

$$L(\mathbf{t}, \theta) = \prod_{i=1}^{n} f(t_i, \theta)$$

- $\log L(\mathbf{t}, \theta)$ is the log likelihood function

$$\log L(\mathbf{t}, \theta) = \sum_{i=1}^{n} \log f(t_i, \theta)$$

- $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_p)'$ is the maximum likelihood estimator

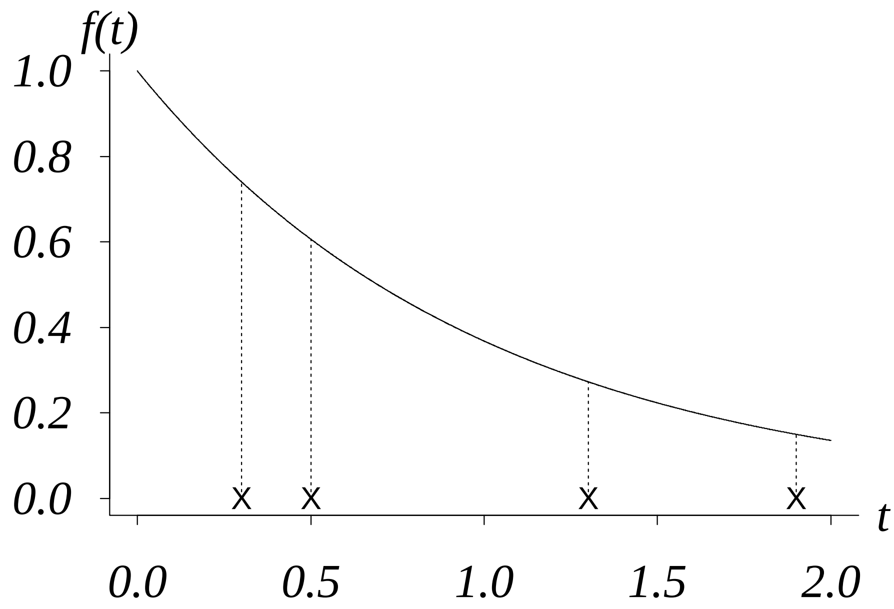

Figure 7.7 Maximum likelihood estimation.

- the i^{th} element of the score vector is

$$U_i(\theta) = \frac{\partial \log L(\mathbf{t}, \theta)}{\partial \theta_i} \qquad i = 1, 2, \ldots, p$$

• the score vector components have expectation

$$E[U_i(\theta)] = 0 \qquad i = 1, 2, \ldots, p$$

and variance-covariance matrix

$$I(\theta) = E[\mathbf{U}(\theta)\,\mathbf{U}'(\theta)]$$

• this variance-covariance matrix is called the *Fisher information matrix* with components

$$Cov(U_i(\theta), U_j(\theta)) = E\left[\frac{-\partial^2 \log L(\mathbf{t}, \theta)}{\partial \theta_i \, \partial \theta_j}\right]$$

for $i = 1, 2, \ldots, p$ and $j = 1, 2, \ldots, p$.

• the *observed information matrix* has components $O(\hat{\theta})$ is

$$\left[\frac{-\partial^2 \log L(\mathbf{t}, \theta)}{\partial \theta_i \, \partial \theta_j}\right]_{\theta = \hat{\theta}} \qquad \begin{matrix} i = 1, 2, \ldots, p \\ j = 1, 2, \ldots, p \end{matrix}$$

Example 7.7 Collect $t_1, t_2, \ldots, t_n$ from an exponential population with a single parameter θ

$$f(t;\theta) = \frac{1}{\theta}\, e^{-t/\theta} \qquad\qquad t > 0$$

The likelihood function is

$$\begin{aligned} L(\mathbf{t}, \theta) &= \prod_{i=1}^{n} f(t_i, \theta) \\ &= \prod_{i=1}^{n} \frac{1}{\theta}\, e^{-t_i/\theta} \\ &= \theta^{-n}\, e^{-\sum_{i=1}^{n} t_i/\theta} \end{aligned}$$

The log likelihood function is

$$\log L(\mathbf{t}, \theta) = -n \log \theta - \sum_{i=1}^{n} t_i/\theta$$

The score vector is

$$U(\theta) = \frac{\partial \log L(\mathbf{t}, \theta)}{\partial \theta} = -\frac{n}{\theta} + \frac{\sum_{i=1}^{n} t_i}{\theta^2}$$

The maximum likelihood estimator is

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} t_i$$

The derivative of the score vector is

$$\frac{\partial^2 \log L(\mathbf{t}, \theta)}{\partial \theta^2} = \frac{n}{\theta^2} - \frac{2 \sum_{i=1}^{n} t_i}{\theta^3}$$

The information matrix is

$$I(\theta) = E\left[\frac{-\partial^2 \log L(\mathbf{t}, \theta)}{\partial \theta^2}\right]$$

$$= E\left[-\frac{n}{\theta^2} + \frac{2 \sum_{i=1}^{n} t_i}{\theta^3} \right]$$

$$= -\frac{n}{\theta^2} + \frac{2}{\theta^3} E\left[\sum_{i=1}^{n} t_i\right]$$

$$= \frac{n}{\theta^2}$$

The observed information matrix is

$$O(\hat{\theta}) = \left[\frac{-\partial^2 \log L(\mathbf{t}, \theta)}{\partial \theta^2}\right]_{\theta = \hat{\theta}}$$

$$= \frac{n}{\hat{\theta}^2}$$

7.5 Censoring

A censored observation occurs when only a bound is known on the time of failure.

Notation

- n: number of items on test
- r: number of observed failures
- c: censoring time

A data set where all failure times are known is called a *complete data set.*

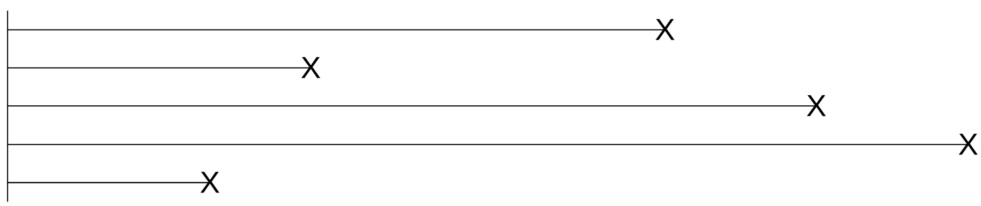

Figure 7.9 A complete data set with $n = 5$.

A data set containing one or more censored observations is called a *censored data set.* The most common type of censoring is *right censoring.*

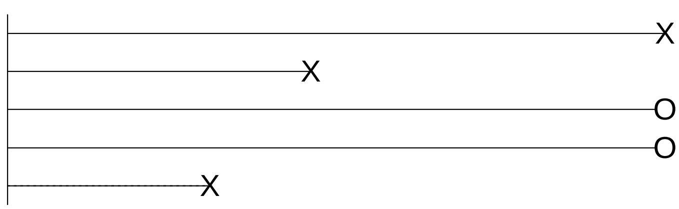

Figure 7.10 A single Type II right-censored data set with $n = 5$ and $r = 3$.

In single Type II censoring, the time to complete the test is random. The second special case is single Type I or *time* censoring.

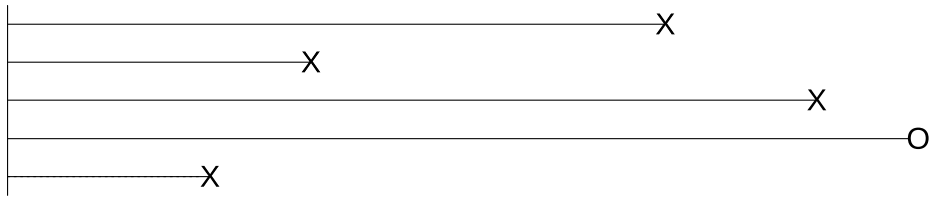

Figure 7.11 A single Type I right-censored data set with $n = 5$ and $r = 4$.

In single Type I censoring, the number of failures is random.

Random censoring occurs when individual items are withdrawn from the test at any time during the study. It is usually assumed that the i^{th} lifetime t_i and the i^{th} censoring time c_i are independent random variables.

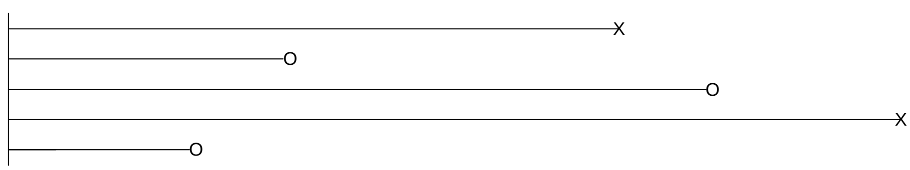

Figure 7.12 A randomly right-censored data set with $n = 5$ and $r = 2$.

8. Fitting Parametric Models to Data

Motivation

Find point and interval estimators for the exponential and Weibull distributions for sample data sets.

Outline

- sample data sets
- exponential distribution
- Weibull distribution

8.1 Sample data sets

Example 8.1 A complete data set of $n = 23$ ball bearing failure times to test the endurance of deep groove ball bearings (in 10^6 revolutions)

17.88	28.92	33.00	41.52	42.12	45.60
48.48	51.84	51.96	54.12	55.56	67.80
68.64	68.64	68.88	84.12	93.12	98.64
105.12	105.84	127.92	128.04	173.40	

Example 8.2 A Type II right-censored data set of $n = 15$ automotive a/c switches with $r = 5$. The observed failure times measured in number of cycles are

1410	1872	3138	4218	6971

Example 8.3 Determine the effect of 6-MP (6-mercaptopurine) on leukemia remission times. A sample of $n = 21$ patients were treated with 6-MP, and $r = 9$ remission times were observed. The survival times (in weeks) are

6 6 6 6* 7 9* 10 10* 11* 13 16 17*
19* 20* 22 23 25* 32* 32* 34* 35*

Control group: 21 other leukemia patients

1 1 2 2 3 4 4 5 5 8 8
8 8 11 11 12 12 15 17 22 23

Example 8.4 Forty motorettes were placed on test at 150°C, 170°C, 190°C and 220°C (ten motorettes at each temperature level and Type I censoring). The failure times (in hours) are

150°C: 8064* 8064* 8064* 8064*
8064* 8064* 8064* 8064* 8064* 8064*

170°C: 1764 2772 3444 3542 3780
4860 5196 5448* 5448* 5448*

190°C: 408 408 1344 1344 1440 1680*
1680* 1680* 1680* 1680*

220°C: 408 408 504 504 504 528*
528* 528* 528* 528*

Failure times are the midpoint of an

inspection period. Operating temperature: 130^oC.

8.2 The exponential distribution

Goal: find point and interval estimates for the $p = 1$ parameter λ.

The exponential distribution can be parameterized by either its failure rate λ or its mean $\mu = \theta = 1 / \lambda$.

$$S(t, \lambda) = e^{-\lambda t} \qquad f(t, \lambda) = \lambda e^{-\lambda t}$$

$$h(t, \lambda) = \lambda \qquad H(t, \lambda) = \lambda t$$

for all $t \geq 0$.

Complete data sets

A complete data set consists of failure times $t_1, t_2, \ldots, t_n$.

$$L(\lambda) = \prod_{i=1}^{n} f(t_i, \lambda)$$

The log likelihood function is

$$\log L(\lambda) = \sum_{i=1}^{n} [\log h(t_i, \lambda) - H(t_i, \lambda)]$$

or

$$\log L(\lambda) = \sum_{i=1}^{n} [\log \lambda - \lambda\, t_i] = n \log \lambda - \lambda \sum_{i=1}^{n} t_i$$

The single element score vector is

$$U(\lambda) = \frac{\partial \log L(\lambda)}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} t_i$$

The MLE is $\hat{\lambda} = \dfrac{n}{\sum_{i=1}^{n} t_i}$

Exact confidence intervals for λ

$$\frac{\hat{\lambda}\ \chi^2_{2n,\,1-\alpha/2}}{2n} < \lambda < \frac{\hat{\lambda}\chi^2_{2n,\,\alpha/2}}{2n}$$

Example 8.5 Consider the $n = 23$ ball bearing failure times. The maximum likelihood estimator is

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} t_i} = \frac{23}{1661.16} = 0.0138$$

failures per 10^6 revolutions. An exact 95% confidence interval is

$$0.00878 < \lambda < 0.0201$$

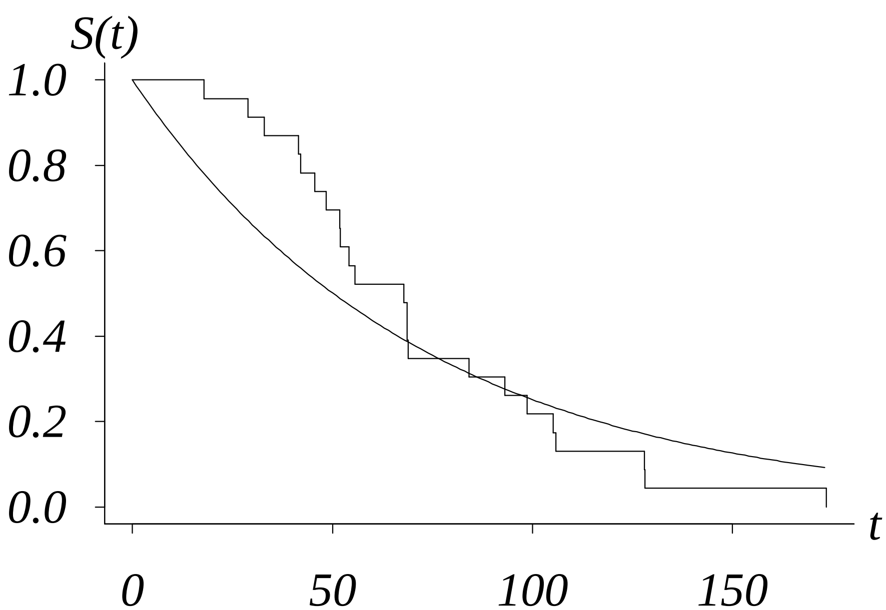

Figure 8.1 Empirical and exponential fitted survivor functions for the ball bearing

data set.

8.3 The Weibull distribution

Goal: find point and interval estimates for the $p = 2$ parameters λ and κ.

The hazard and cumulative hazard functions are

$$h(t, \lambda, \kappa) = \kappa\lambda(\lambda t)^{\kappa - 1} \quad \text{and} \quad H(t, \lambda, \kappa) = (\lambda t)^{\kappa}$$

for $t \geq 0$.

Notation

- n is the number of items on test
- r is the number of observed failures
- $t_1, t_2, \ldots, t_n$ are the failure times
- $c_1, c_2, \ldots, c_n$ are the censoring times
- $x_i = \min\{t_i, c_i\}$ for $i = 1, 2, \ldots, n$

Assuming random censoring

$$\log L(\lambda, \kappa) = \sum_{i \in U} \log h(x_i, \lambda, \kappa) - \sum_{i=1}^{n} H(x_i, \lambda, \kappa)$$

$$= r \log \kappa + \kappa r \log \lambda + (\kappa - 1) \sum_{i \in U} \log x_i - \lambda^{\kappa} \sum_{i=1}^{n} x_i^{\kappa}$$

The 2×1 score vector has elements

$$U_1(\lambda, \kappa) = \frac{\partial \log L(\lambda, \kappa)}{\partial \lambda} = \frac{\kappa r}{\lambda} - \kappa \lambda^{\kappa - 1} \sum_{i=1}^{n} x_i^{\kappa}$$

$$U_2(\lambda, \kappa) = \frac{\partial \log L(\lambda, \kappa)}{\partial \kappa}$$
$$= \frac{r}{\kappa} + r \log \lambda + \sum_{i \in U} \log x_i - \sum_{i=1}^{n} (\lambda x_i)^{\kappa} \log \lambda x_i$$

There is not a closed form solution for $\hat{\lambda}$ and $\hat{\kappa}$.

$$\frac{\kappa r}{\lambda} - \kappa \lambda^{\kappa - 1} \sum_{i=1}^{n} x_i^{\kappa} = 0$$

$$\frac{r}{\kappa} + r \log \lambda + \sum_{i \in U} \log x_i - \sum_{i=1}^{n} (\lambda x_i)^{\kappa} \log \lambda x_i = 0$$

The first equation can be solved for λ in terms of κ

$$\lambda = \left(\frac{r}{\sum_{i=1}^{n} x_i^{\kappa}} \right)^{1/\kappa}$$

The 2×2 information matrices are based on

$$\frac{-\partial^2 \log L(\lambda, \kappa)}{\partial \lambda^2} = \frac{\kappa r}{\lambda^2} + \kappa(\kappa - 1)\lambda^{\kappa - 2} \sum_{i=1}^{n} x_i^{\kappa}$$

$$\frac{-\partial^2 \log L(\lambda, \kappa)}{\partial \kappa^2} = \frac{r}{\kappa^2} + \sum_{i=1}^{n} (\lambda x_i)^{\kappa} (\log \lambda x_i)^2$$

$$\frac{-\partial^2 \log L(\lambda, \kappa)}{\partial \lambda \partial \kappa} =$$
$$-\frac{r}{\lambda} + \lambda^{\kappa - 1} \left[\kappa \sum_{i=1}^{n} x_i^{\kappa} \log x_i + (1 + \kappa \log \lambda) \sum_{i=1}^{n} x_i^{\kappa} \right]$$

Information matrices

• the expected values of these quantities are not tractable

• use $\hat{\lambda}$ and $\hat{\kappa}$ to obtain the observed information matrix

Example 8.15 Ball bearing data set. The fitted Weibull distribution: $\hat{\lambda} = 0.0122$ and $\hat{\kappa} = 2.10$.

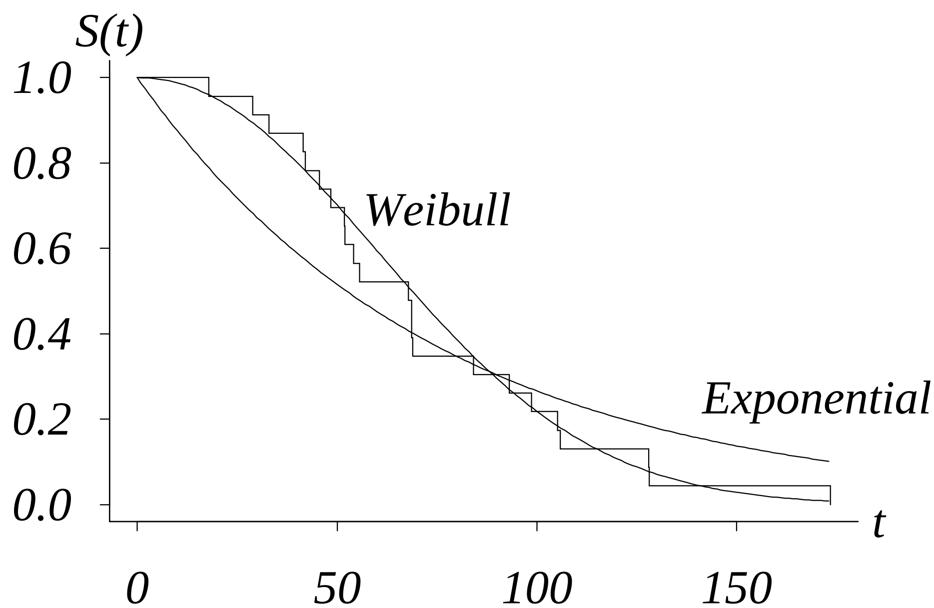

Figure 8.8 Exponential and Weibull fits to the ball bearing data.

The log likelihood function at the MLEs is

$$\log L(\hat{\lambda}, \hat{\kappa}) = -113.691$$

The observed information matrix is

$$O(\hat{\lambda}, \hat{\kappa}) = \begin{bmatrix} 681,000 & 875 \\ 875 & 10.4 \end{bmatrix}$$

Using the fact that the likelihood ratio statistic, $2[\log L(\hat{\lambda}, \hat{\kappa}) - \log L(\lambda, \kappa)]$, is asymptotically $\chi^2(2)$, a 95% confidence region is

$$2[-113.691 - \log L(\lambda, \kappa)] < 5.99$$

since $\chi^2_{2,\,0.05} = 5.99$.

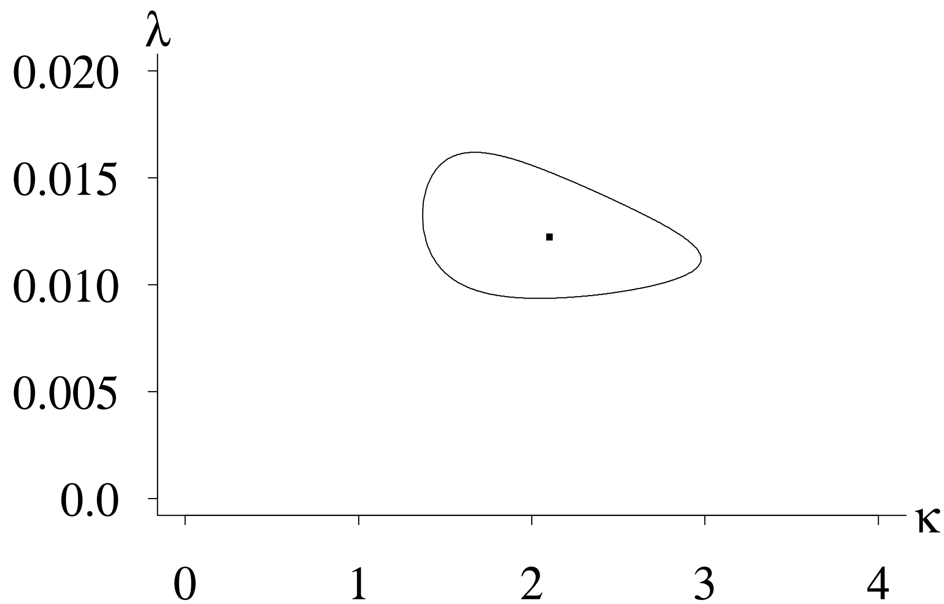

Figure 8.9 Confidence region for λ and κ ($\alpha = 0.05$).

Inverse of the observed information matrix:

$$O^{-1}(\hat{\lambda}, \hat{\kappa}) = \begin{bmatrix} 0.00000165 & -0.000139 \\ -0.000139 & 0.108 \end{bmatrix}$$

The standard errors of the parameter estimators are the square roots of the diagonal elements

$$\hat{\sigma}_{\hat{\lambda}} = 0.00128 \qquad \hat{\sigma}_{\hat{\kappa}} = 0.329$$

An asymptotic 95% confidence interval for κ is

$$2.10 - (1.96)(0.329) < \kappa < 2.10 + (1.96)(0.329)$$

or $1.46 < \kappa < 2.74$.

9. Parametric Estimation for Models with Covariates

Motivation

Estimate parameters for the accelerated life and proportional hazards models.

Outline

- model formulation
- accelerated life
- proportional hazards

9.1 Model formulation

Goal: estimate the vector of regression coefficients $\beta = (\beta_1, \beta_2, \ldots, \beta_q)'$

Applications

- determine which covariates significantly impact survival
- determine the impact of changing the values of covariates

The accelerated life model

$$S(t, \mathbf{z}) = S_0(t\,\psi(\mathbf{z}))$$

The proportional hazards model

$$h(t, \mathbf{z}) = \psi(\mathbf{z})\; h_0(t)$$

Notation (for $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, q$)

- $x_i = \min\{t_i, c_i\}$
- δ_i is a censoring indicator variable
- $\mathbf{z}_i = (z_{i1}, z_{i2}, \ldots, z_{iq})'$
- z_{ij} is the value of covariate j for item i
- extra parameters: $S(t, \mathbf{z}, \theta, \beta)$

Matrix formulation

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \quad \delta = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \cdot \\ \cdot \\ \cdot \\ \delta_n \end{bmatrix} \quad \text{and} \quad \mathbf{Z} = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1q} \\ z_{21} & z_{22} & \cdots & z_{2q} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ z_{n1} & z_{n2} & \cdots & z_{nq} \end{bmatrix}$$

$$L(\theta, \beta) = \prod_{i \in U} f(x_i, \mathbf{z}_i, \theta, \beta) \prod_{i \in C} S(x_i, \mathbf{z}_i, \theta, \beta)$$

The log likelihood function is

$$\log L(\theta, \beta) = \sum_{i \in U} \log f(x_i, \mathbf{z}_i, \theta, \beta) + \sum_{i \in C} \log S(x_i, \mathbf{z}_i, \theta, \beta)$$

or

$$\log L(\theta, \beta) = \sum_{i \in U} \log h(x_i, \mathbf{z}_i, \theta, \beta) - \sum_{i=1}^{n} H(x_i, \mathbf{z}_i, \theta, \beta)$$

Notes

- the maximum likelihood estimators for θ and β cannot be expressed in closed form
- the number of unique covariate vectors and n determine whether to use regression models

9.3 Proportional hazards

Example 9.2 A set of $n = 3$ light bulbs are placed on test. The first and second bulbs are 100-watt bulbs and the third bulb is a 60-watt bulb. A single ($q = 1$) covariate z_1 assumes the value 0 for a 60-watt bulb and 1 for a 100-watt bulb. Does the wattage have any influence on the survival distribution of the bulbs?

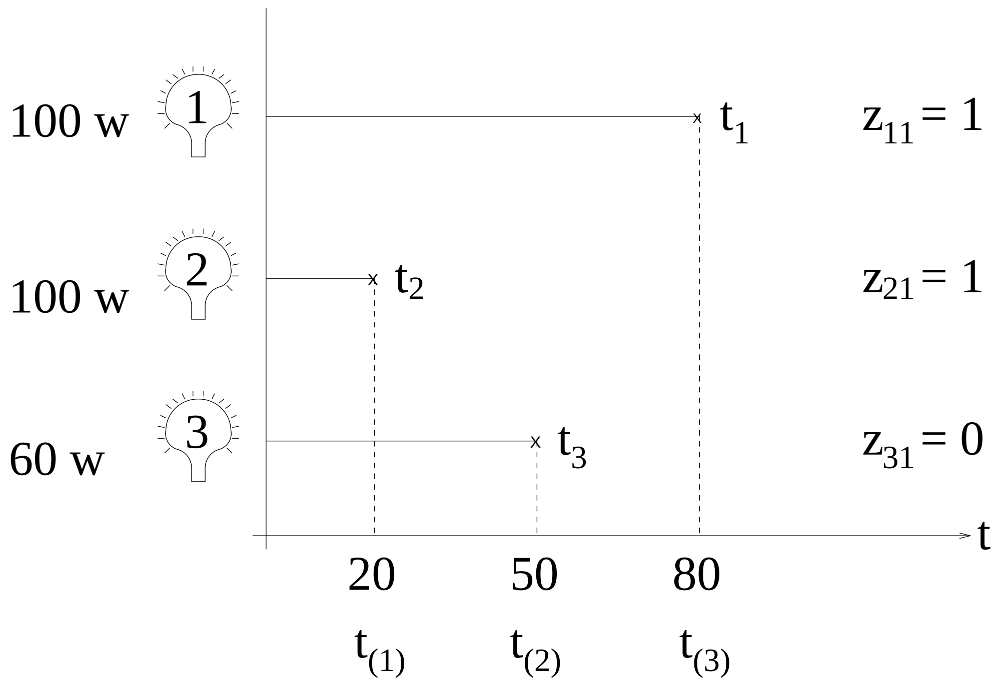

Figure 9.2 Proportional hazards parameter estimation notation.

$$\mathbf{x} = \begin{bmatrix} 80 \\ 20 \\ 50 \end{bmatrix} \qquad \delta = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{Z} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

Example 9.8 North Carolina collected recidivism data on $n = 1540$ prisoners in 1978 (Schmidt and Witte, 1988). T is the time of release until the time of return to prison. The purpose of the study is to assess the impact of the $q = 15$ covariates.

Table 9.2 North Carolina recidivism model.

z_i	Covar.	$\hat{\beta}$	$\sqrt{\hat{V}[\hat{\beta}]}$	$\frac{\hat{\beta}}{\sqrt{\hat{V}[\hat{\beta}]}}$	p-value
z_2	AGE	-3.34	0.52	-6.43	0.000
z_3	PRIORS	0.84	0.14	6.10	0.000
z_1	TSERV	1.17	0.20	5.96	0.000
z_6	WHITE	-0.44	0.09	-5.07	0.000
z_8	ALCHY	0.43	0.10	4.11	0.000
z_{13}	FELON	-0.58	0.16	-3.54	0.000
z_9	JUNKY	0.28	0.10	2.91	0.002
z_7	MALE	0.67	0.24	2.78	0.003
z_{15}	PROPTY	0.39	0.16	2.47	0.007
z_4	RULE	3.08	1.69	1.82	0.034
z_{10}	MARRY	-0.15	0.11	-1.42	0.077
z_5	SCHOOL	-0.25	0.19	-1.30	0.097
z_{12}	WORK	0.09	0.09	0.96	0.169
z_{14}	PERSON	0.07	0.24	0.30	0.381
z_{11}	SUPER	-0.01	0.10	-0.09	0.464

10. Nonparametric Methods

Motivation

Let the data "speak for itself", rather than approximating the lifetime distribution by one of the parametric models.

Outline

- nonparametric estimates of the survivor function
- life tables

10.1 Survivor function estimation

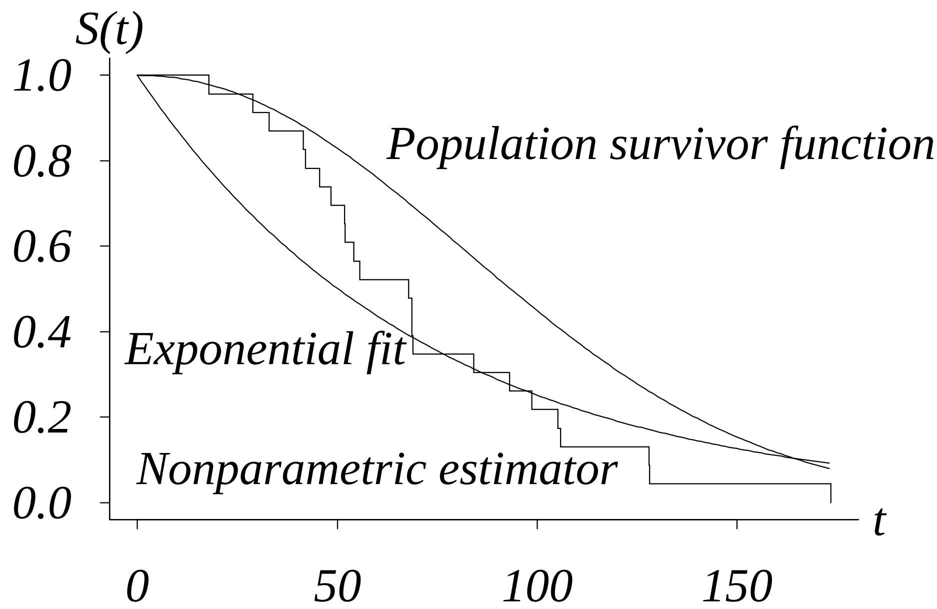

Figure 10.1 Parametric vs. nonparametric survivor function estimates.

CASE I: complete data set of n lifetimes

Notation

- $R(t)$, known as the risk set, contains the indices of all items at risk just prior to time t
- $n(t) = |R(t)|$ is the cardinality, or number of elements in $R(t)$

A nonparametric estimate for the survivor function is

$$\hat{S}(t) = \frac{n(t)}{n} \qquad\qquad t \ge 0$$

Notes

- often referred to as the *empirical* survivor function
- has a downward step of $\frac{1}{n}$ at each observed lifetime if there are no ties
- takes a downward step of $\frac{d}{n}$ if there are d tied observations at a particular time value

An asymptotically valid $100(1-\alpha)\%$ confidence interval for $S(t)$ is

$$\hat{S}(t) \pm z_{\alpha/2}\sqrt{\frac{\hat{S}(t)(1-\hat{S}(t))}{n}}$$

Example 10.1 For the ball bearing data set, find a nonparametric survivor function estimator and a 95% confidence interval for the probability that a ball bearing will last 50,000,000 cycles.

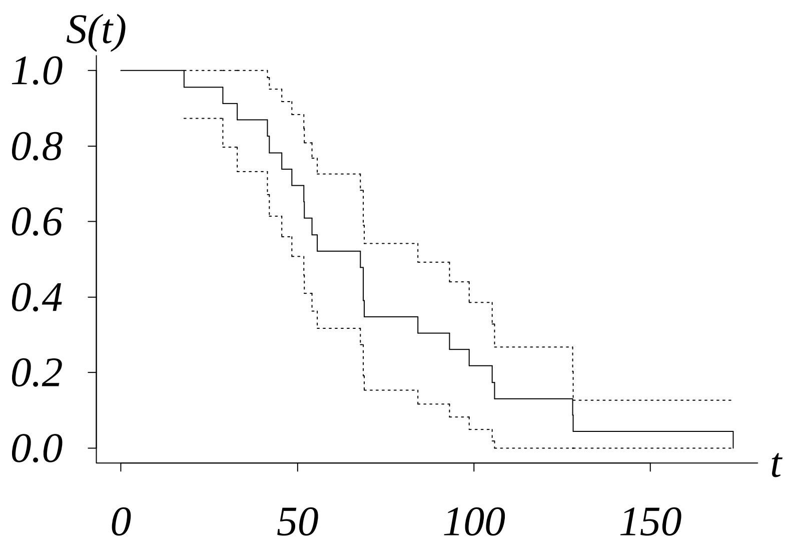

Figure 10.2 Ball bearing lifetime survivor function estimate.

Note

- the downward steps in $\hat{S}(t)$ have been connected by vertical lines
- some authors connect the survivor function estimates at the failure times with lines
- the survivor function takes a downward step of 1 / 23 at each

data value except 68.64, where it takes a downward step of 2 / 23

A point estimate for $S(50)$ is $\hat{S}(50) = \frac{16}{23} = 0.696$, and a 95% confidence interval for $S(50)$ is

$$\hat{S}(50) \pm 1.96 \sqrt{\frac{\hat{S}(50)(1 - \hat{S}(50))}{23}}$$

or

$$0.508 < S(50) < 0.884$$

Case II: Right censoring

Notation

- let $y_1 < y_2 < \ldots < y_k$ be the k distinct failure times
- let d_j denote the number of observed failures at y_j, $j = 1, 2, \ldots, k$
- let $n_j = n(y_j)$ denote the number of items on test just before time y_j, $j = 1, 2, \ldots, k$ and it is customary to include any values that are censored at y_j in this count
- $R(y_j)$ is the set of all indices of items that are at risk just before time y_j, $j = 1, 2, \ldots, k$

Kaplan-Meier (product-limit) estimator

$$\hat{S}(t) = \prod_{j \in R(t)'} [1 - \frac{d_j}{n_j}]$$

11. Model Adequacy

Motivation

Once a distribution has been fitted to a sample data set, the adequacy of the model should be accessed.

Outline

- all parameters known
- parameters estimated from data

11.1 All parameters known

Kolmogorov-Smirnov test

$$H_0: \ F(t) = F_0(t)$$

$$H_1: \ F(t) \neq F_0(t)$$

where $F(t)$ is the underlying population cumulative distribution function. For a complete data set, the test statistic is

$$D_n = \sup_t |\hat{F}(t) - F_0(t)|$$

Computational formulas

$$D_n^+ = \max_{i=1,2,\ldots,n} \left(\frac{i}{n} - F_0(t_{(i)}) \right)$$

$$D_n^- = \max_{i=1,\,2,\,\ldots,\,n}\left(F_0(t_{(i)}) - \frac{i-1}{n}\right)$$
$$D_n = \max\{D_n^+,\ D_n^-\}$$

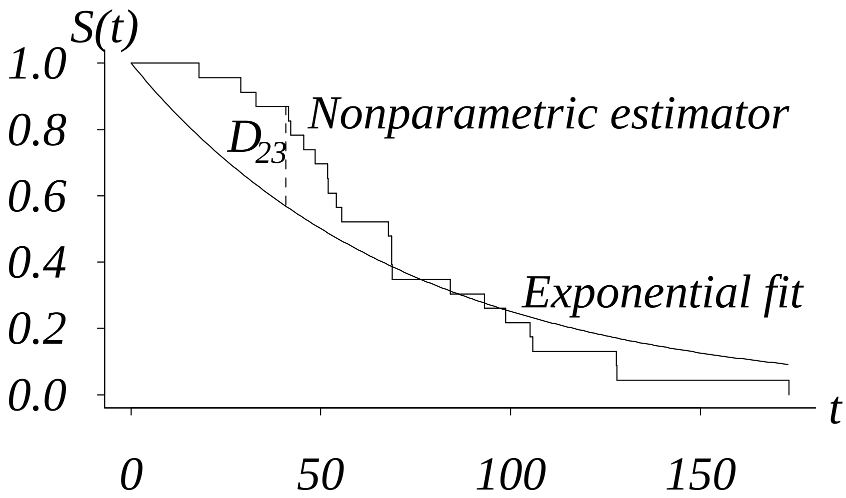

Figure 11.1 Empirical and fitted survivor functions.

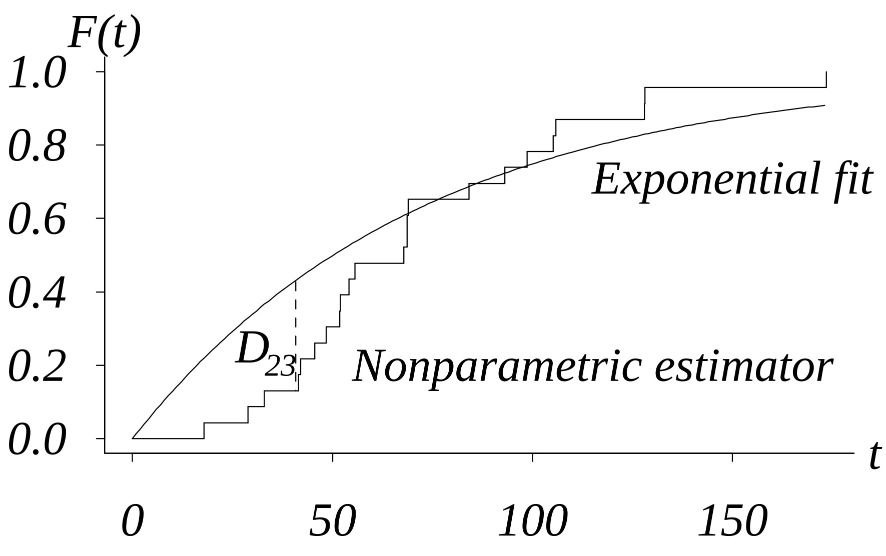

Figure 11.2 Empirical and fitted cdfs.

$D_{23} = 0.301$ occurs just to the left of

$t_{(4)} = 41.52.$

Table 11.1 Approximate K-S critical values (all parameters known).

	$1-\alpha$			
n	0.80	0.90	0.95	0.99
1	0.900	0.950	0.975	0.995
2	0.683	0.775	0.841	0.929
3	0.565	0.636	0.708	0.829
4	0.493	0.565	0.624	0.734
5	0.447	0.510	0.564	0.668
6	0.410	0.468	0.519	0.615
7	0.381	0.435	0.483	0.576
8	0.358	0.410	0.455	0.543
9	0.339	0.387	0.430	0.513
10	0.323	0.369	0.409	0.490
15	0.266	0.304	0.338	0.405
20	0.231	0.264	0.294	0.352
23	0.216	0.248	0.275	0.330
25	0.208	0.237	0.264	0.317
30	0.190	0.217	0.242	0.290
40	0.166	0.189	0.210	0.252
50	0.148	0.170	0.188	0.226

Table 11.1 gives estimates of the $1-\alpha$ fractiles of the distribution of D_n under H_0 determined by Monte Carlo simulation (500,000 replications).

Example 11.1 Use the K-S test to determine whether the ball bearing data set was drawn from a Weibull population with $\lambda = 0.01$ and $\kappa = 2$. Run the test at $\alpha = 0.10$.

The goodness-of-fit test is

$$H_0:\ F(t) = 1 - e^{-(0.01t)^2}$$

$$H_1:\ F(t) \neq 1 - e^{-(0.01t)^2}$$

The test statistic is $D_{23} = 0.274$. At $\alpha = 0.10$, the critical value is 0.248, so H_0 is rejected.

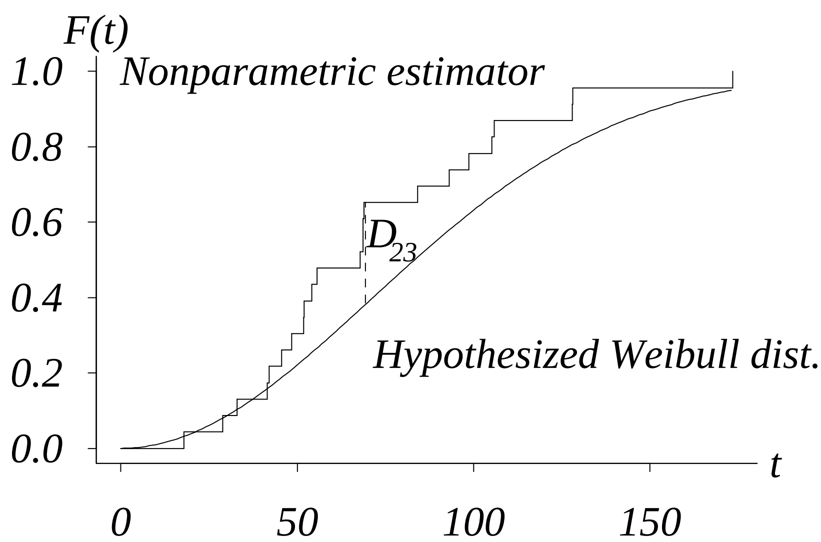

Figure 11.6 Empirical and fitted Weibull cumulative distribution functions for the ball bearings.