KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

Politechnika Wrocławska

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

SELECTED STOCHASTIC MODELS IN RELIABILITY

Alicja Jokiel-Rokita and Ryszard Magiera

„Zamawianie kształcenia na kierunkach technicznych, matematycznych, przyrodniczych - pilotaż

Projekt współfinansowany ze środków Unii Europejskiej
w ramach Europejskiego Funduszu Społecznego

KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

Politechnika Wrocławska

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

Recenzent:
Dr hab. Maciej Wilczyński

SELECTED STOCHASTIC MODELS IN RELIABILITY

ISBN 978-83-7493-679-8

Alicja Jokiel-Rokita and Ryszard Magiera

SELECTED STOCHASTIC MODELS IN RELIABILITY

Wrocław 2011

Alicja Jokiel-Rokita, Wroclaw University of Technology, Institute of Mathematics and Computer Science
e-mail: Alicja.Jokiel-Rokita@pwr.wroc.pl
Ryszard Magiera, Wroclaw University of Technology, Institute of Mathematics and Computer Science
e-mail: Ryszard.Magiera@pwr.wroc.pl

Preface

The aim of this handbook is to present most commonly used stochastic models for repairable systems and to consider some fundamental problems of estimating unknown parameters of these models. Special attention is paid to the trend-renewal process (TRP), which is recently widely discussed in the literature. The class of these processes covers non-homogeneous Poisson and renewal processes. As alternatives to the maximum likelihood (ML) method, three other method of estimation are proposed for the models considered. The alternative methods, called the least squares (LS) method, the constrained least squares (CLS) method and the method of moments (M method) can be useful when the ML method fails (as for some non-homogeneous Poisson software reliability models) or the renewal distribution of a TRP is unknown.

In Chapter 1 some basic notions from survival analysis are reminded and the classification of lifetime distributions in terms of ageing is presented (IFR, DFR, IFRA, DFRA, NBU, NWU and bathtub-shaped failure classes). Chapter 2 provides a review of parametric families of lifetime distributions.

The most commonly used stochastic models for repairable systems, such as the homogeneous Poisson, renewal, non-homogeneous Poisson (with its special cases, with power law intensity function and with bounded mean value function) and TRP's are described in Chapter 3. In the handbook a grater deal of attention is paid to the TRP's, whose basic properties are presented in Sections 3.2.6 and 3.2.7. Maximum likelihood estimation problem in the non-homogeneous Poisson process with power law intensity function is considered in Chapter 4.

Chapter 5 is devoted to estimation problems in TRP models. In Section 5.2 the form of likelihood function for a TRP process is presented. The likelihood function and the likelihood equations for estimating the parameters of the TRP with Weibull renewal distribution and power law trend function are given in Section 5.3. Mainly, in Chapter 5, methods of estimating unknown parameters of a trend function for TRP's are investigated in the case when the renewal distribution function of this process is unknown. If the renewal distribution is unknown, then the likelihood function of the TRP is unknown and consequently the ML method cannot be used. In such situation, in Section 5.4 three other methods of estimating the trend parameters are presented: the LS, CLS and M methods. The problem of estimating trend parameters of a TRP with unknown renewal distribution may be of interest in the situation when we observe several systems, of the same kind, working in different environments and we are interesting in examining and comparing their trend functions, whatsoever their renewal distribution is. The estimation problem of the trend parameters in some special case of the TRP is considered in Section 5.5. In Section 5.6 the estimators proposed are examined and compared with the ML estimators (obtained under the additional assumption that the renewal distribution has a known parametric form) through a computer simulation study. Some real data are examined in Section 5.6.3. Section 5.7 contains conclusions and some prospects. Chapter 5 contains the results of Jokiel-Rokita and Magiera (2011).

In Chapter 6 a subclass of non-homogeneous Poisson processes is considered. This subclass with bounded mean value function can be used to model software reliability. The ML estimators of intensity parameters for a software reliability model are given in Section 6.2. In certain cases the ML estimators do not exist. In such cases the alternative methods of estimating proposed in Section 5.4 can be used. In Sections 6.3 and 6.5 the LS and CLS methods are applied to the software reliability model considered and to its special case. In Section 6.6 some numerical results illustrating the accuracy of the proposed LS and CLS estimators with comparison to the ML estimators are presented for a special case of the Erlangian non-homogeneous Poisson process software reliability model. Chapter 6 contains the results of Jokiel-Rokita and Magiera (2010).

Wrocław, November 15, 2011 *Alicja Jokiel-Rokita* *Ryszard Magiera*

Contents

List of Figures

List of Tables

Chapter 1

Classes of Lifetime Distributions in Reliability Models

1.1 Functions characterizing lifetime distributions

1.1.1 Survival function

Let T be a non-negative valued random variable determining, for example, a working time of a device (or unit). Such random variable is called the **lifetime** or **survival time**. Let $F(t)$ denote the cumulative distribution function (cdf) of the random variable T.

Definition 1.1.1 The function

$$S(t) = P(T > t) = 1 - F(t)$$

is called a **reliability function** or **survival function**. $\square$

Reliability of a unit at time t is defined as the probability that the working time of this unit is greater than t.

1.1.2 Failure rate function

Assume that T is a continuous type random variable with density function $f(t)$.

Definition 1.1.2 The function defined by the formula

$$\rho(t) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)} = -\frac{S'(t)}{S(t)} \qquad (1.1)$$

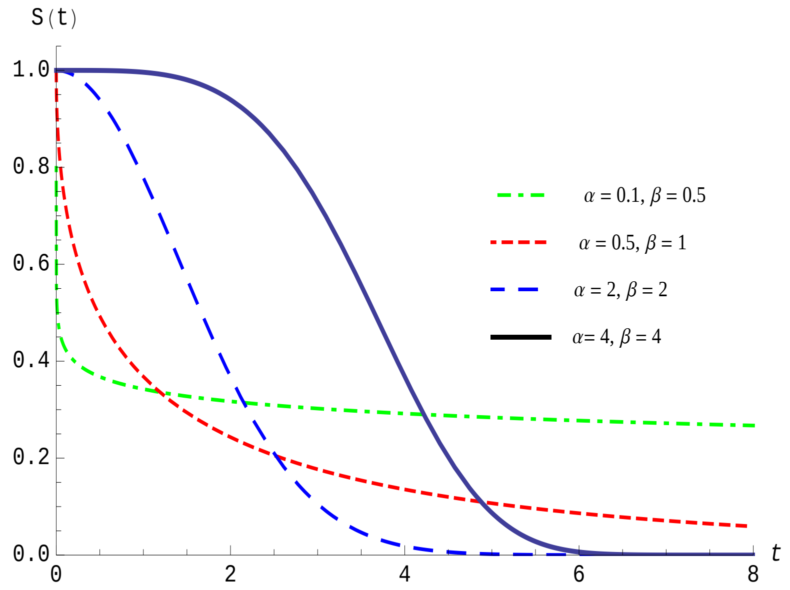

Figure 1.1: Survival functions of the Weibull $\mathcal{W}e(\alpha, \beta)$ distribution.

is called the **failure rate function**. □

The failure rate function is also called the **hazard function** or **hazard rate**.

The failure rate function at time t is the density of the probability distribution of the lifetime of a unit, given the lifetime is greater then t:

$$\begin{aligned}\rho(t) = \lim_{\Delta t \to 0} \frac{P(t < T \leq t + \Delta t)}{\Delta t P(T > t)} &= \lim_{\Delta t \to 0} \frac{f(t)\Delta t + o(\Delta t)}{\Delta t[1 - F(t)]} \\ &= \lim_{\Delta t \to 0} \left[\rho(t) + \frac{o(\Delta t)}{\Delta t}\right].\end{aligned}$$

The failure rate function has the following interpretation: if at time $t_0 = 0$ a unit is included in a system to work, and it then works without failure until time t, then the probability that it fails in the interval $(t, t + \Delta t]$ one can estimate by $\rho(t)\Delta t$.

Under the assumption that $S(t) > 0$, it follows from formula (1.1) that

$$S(t) = \exp\Big(-\int_0^t \rho(u)du\Big), \ t \geq 0.$$

Thus an integrable function $\rho(t)$ is the failure rate function if and only if

$$\rho(t) \geq 0, \quad t \geq 0, \quad \int_0^\infty \rho(t)dt = \infty. \tag{1.2}$$

The failure rate function for the ▷ *exponential distribution* (p. 13) $\mathcal{E}(\lambda)$ is the constant function $\rho(t) = 1/\lambda$. For the ▷ *Weibull distribution* (p. 17) $\mathcal{W}e(\alpha, \beta)$ the failure rate function is of the form $\rho(t) = \alpha\beta^{-\alpha}t^{\alpha-1}$ (see Table 1.1).

Distribution	$\rho(t)$
Exponential $\mathcal{E}(\lambda)$	$1/\lambda$
Weibull $\mathcal{W}e(\alpha, \beta)$	$\alpha\beta^{-\alpha}t^{\alpha-1}$
Gamma $\mathcal{G}(\alpha, \lambda)$	$\dfrac{t^{\alpha-1}\exp(-t/\lambda)}{\lambda^\alpha\Gamma(\alpha)[1-\gamma(t/\lambda, \alpha)]}$
Gompertz $\mathcal{G}om(\alpha, \beta)$	$\beta \mathrm{e}^{\alpha t}$

Table 1.1: Failure rate functions for some continuous distributions

The function $\gamma(x, \alpha)$ appearing in the formula for the failure rate function in the ▷ *gamma distribution* (p. 19) in Table 1.1 denotes the so called modified incomplete gamma function.

Definition 1.1.3 The function

$$H(t) = \int_0^t \rho(u)du = -\ln[S(t)]$$

is called **cumulative hazard function**. □

1.1.3 Mean residual life function

Suppose that a unit has worked without failure up to time t, and denote by T_t its residual lifetime up to next failure. Let S_t be the survival function corresponding to T_t. We have

$$S_t(u) = P(T_t > u) = P(T > t + u | T > t) = \frac{S(t+u)}{S(t)}.$$

Definition 1.1.4 **The mean residual life function** is defined by

$$L_F(t) = E(T_t) = \int_0^\infty S_t(u)du = \int_0^\infty \frac{S(t+u)}{S(t)}du, \quad t \geq 0.$$

□

We then have the following relations:

$$\begin{aligned} L_F(0) &= E(T), \\ \rho(t) &= 1 + \frac{L'_F(t)}{L(t)}. \end{aligned}$$

1.2 Classification of lifetime distributions in terms of ageing

In this section we present a classification of lifetime distributions which is based on monotonicity property of the failure rate function $\rho(\cdot)$. Such classification is useful to survival and reliability analysis.

1.2.1 Increasing and decreasing failure rate classes

Definition 1.2.5 A lifetime X is said to have **increasing failure rate (IFR)**, or the distribution of X is said to belong to the class IFR, if the failure rate function $\rho(x)$ is increasing for $x \geq 0$. □

Example 1.2.1 Consider the **Gompertz-Makeham distribution**. This distribution is defined by the density function

$$f(x) = (\alpha \mathrm{e}^{\beta x} + \lambda) \exp\left[-\lambda x - \frac{\alpha}{\beta}(\mathrm{e}^{\beta x} - 1)\right], \quad x \in \mathbb{R}^+,$$

$\alpha > 0, \beta > 0, \lambda > 0$. The corresponding cdf is given by

$$F(x) = 1 - \exp\left[-\lambda x - \frac{\alpha}{\beta}(\mathrm{e}^{\beta x} - 1)\right],$$

and the corresponding failure rate function is

$$\rho(x) = \alpha \mathrm{e}^{\beta x} + \lambda.$$

The Gompertz-Makeham law states that the death rate is the sum of an age-independent component (the Makeham term) and an age-dependent component (the Gompertz function) which increases exponentially with age. In a protected environment where external causes of death are rare (laboratory conditions, low mortality countries, etc.), the age-independent mortality component is often negligible. In this case the formula simplifies to a Gompertz law of mortality. The

Gompertz law is the same as a **Fisher-Tippett distribution** for the negative of age, restricted to negative values for the random variable (positive values for age). □

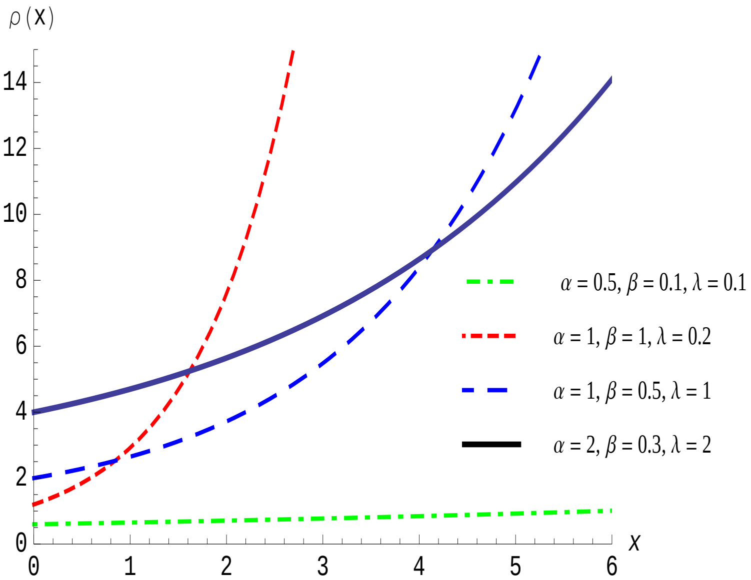

Figure 1.2: Hazard rate functions of the Gompertz-Makeham distribution.

Analogously, a lifetime X is said to have **decreasing failure rate (DFR)**, or the distribution of X is said to belong to the DFR class, if $\rho(x)$ is non-increasing for $x \geq 0$.

Examples of IFR distributions are the Weibull distribution $\mathcal{W}e(\alpha, \beta)$ for $\alpha > 1$ and the gamma distribution $\mathcal{G}(\alpha, \lambda)$ for $\alpha > 1$. For $\alpha < 1$, these distributions belong to the DFR class.

The exponential distribution $\mathcal{E}(\lambda)$ has the constant failure rate function $\rho(x) = 1/\lambda$ and it is the IFR as well as the DFR distribution. If a lifetime has a constant failure rate function we say that we observe **no ageing**.

If a lifetime of units has an IFR distribution, then we say that the units are ageing and we deal with an **ageing process** or with **positive ageing**.

If a lifetime of units has an DFR distribution, then we say that the units are running in and we deal with an **improvement process** or with **negative ageing**.

Definition 1.2.6 A lifetime X is said to have a distribution with **increasing failure rate on the average (IFRA)**, or the distribution of X is said to belong to the class IFRA, if the function $H(x)/x$ is non-decreasing for $x > 0$, where $H(x)$ denotes the cumulative hazard function. □

The class **DFRA** is defined in an analogous way.

Theorem 1.2.1 If F is the cdf of a continuous type random variable X and $F(0-) = 0$, then X has the IFRA distribution if and only if

$$S(bx) \geq [S(x)]^b, \tag{1.3}$$

for $x > 0,\ 0 < b < 1$. □

For the DFRA distribution, the reverse inequality sign in (1.3) holds.

1.2.2 New better than used and new worse than used classes

A wider class than the distribution class IFR (IFRA) or DFR (DFRA) is the class *new better than used.*

Definition 1.2.7 A non-negative valued random variable X is said to have **new better than used (NBU) distribution** if

$$P(X \geq x + y | X \geq x) \leq P(X \geq y), \quad x, y \geq 0.$$

□

This means that if X is the lifetime of a unit, then it has NBU distribution if for all $x, y \geq 0$ the probability that the unit which worked without failure up to time x, it will work further without failure in the time interval of length at least y, is less or equal to the probability that a new unit will work without failure in the time interval of length at least y.

If F is a cdf of a continuous type random variable X with $F(0-) = 0$, then X has the NBU distribution if

$$S(x + y) \leq S(x)S(y) \quad \text{for all} \quad x,\ y \geq 0.$$

An analogous to the NBU class is the class of **new worse than used (NWU)** distributions.

The following implications holds between the failure rate classes:

$$\begin{array}{ccccc} \textbf{IFR} & \Longrightarrow & \textbf{IFRA} & \Longrightarrow & \textbf{NBU}, \\ \textbf{DFR} & \Longrightarrow & \textbf{DFRA} & \Longrightarrow & \textbf{NWU}. \end{array}$$

1.2.3 Bathtub-shaped failure rate functions

The bathtub-shaped failure rate functions are widely used in reliability systems. A formal definition of a bathtub-shaped failure rate is:

Definition 1.2.8 A failure rate function $\rho(x)$ is said to have a **bathtub shape** if there exist x_1 and x_2 such that $0 \le x_1 \le x_2 \le \infty$ and

$$\rho(x) \text{ is } \begin{cases} \text{strictly decreases if} & 0 \le x \le x_1, \\ \text{almost constant if} & x_1 \le x \le x_2, \\ \text{strictly increases if} & x \ge x_2. \end{cases} \tag{1.4}$$

□

The points x_1 and x_2 are called the first and second **change points**, respectively. The time interval $[0, x_1]$ is called the **infant mortality period**; the interval $[x_1, x_2]$, where $\rho(x)$ is (almost) flat and attains its minimum value, is called the **normal operating life** or the **useful life**; and the interval $[x_2, \infty]$ is called the **wear-out period**. In other words, the bathtub curve describes a particular form of the hazard function which comprises three parts: the first part is a decreasing failure rate, known as **early failures**; the second part is an almost constant failure rate, known as **normal failures**; the third part is an increasing failure rate, known as **wear-out failures**. The name is derived from the cross-sectional shape of a bathtub.

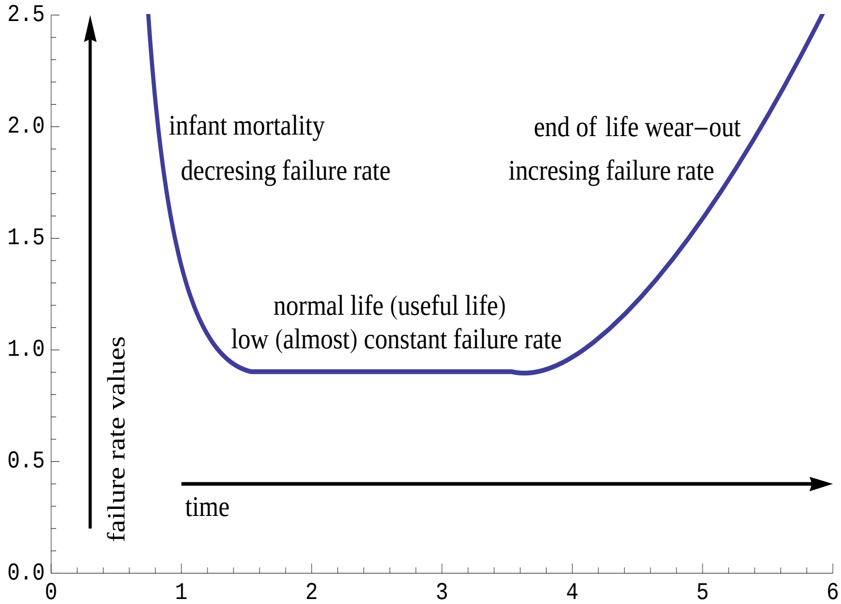

Figure 1.3: Bathtub-shaped failure rate function.

The bathtub curve is generated by mapping the rate of early "infant mortality" failures when first introduced, the rate of random failures with constant failure rate during its "useful life", and finally the rate of "wear out" failures as the product exceeds its design lifetime.

In less technical terms, in the early life of a product adhering to the bathtub curve, the failure rate is high but rapidly decreasing as defective products are identified and discarded, and early sources of potential failure such as handling and installation error are surmounted. In the mid-life of a product – generally, once it reaches consumers – the failure rate is low and constant. In the late life of the product, the failure rate increases, as age and wear take their toll on the product. Many consumer products strongly reflect the bathtub curve, such as computer processors.

While the bathtub curve is useful, not every product or system follows a bathtub curve hazard function, for example if units are retired or have decreased use during or before the onset of the wear-out period, they will show fewer failures per unit calendar time (not per unit use time) than the bathtub curve.

Let us consider some examples of bathtub-type hazard rate functions.

Example 1.2.2 Let a lifetime X have the cdf defined by

$$F(x) = 1 - \mathrm{e}^{\lambda\left(1-\mathrm{e}^{x^\beta}\right)} \quad (x > 0),$$

where $\lambda > 0$ and $\beta > 0$ are the parameters. This type of distribution was considered by Chen (2000) and we call it the **Chen distribution**. The corresponding failure rate function of this distribution is

$$\rho(x) = \lambda\beta x^{\beta-1}\mathrm{e}^{x^\beta} \quad (x > 0).$$

Since

$$\rho'(x) = \lambda\beta x^{\beta-2}\mathrm{e}^{x^\beta}\left[(\beta-1)+\beta x^\beta\right] \quad (x > 0),$$

$\rho(x)$ may have a bathtub shape when $\beta < 1$. The distribution has increasing failure rate function when $\beta \geq 1$. Figures 1.4 and 1.5 show the failure rate functions for various values of λ and β. $\square$

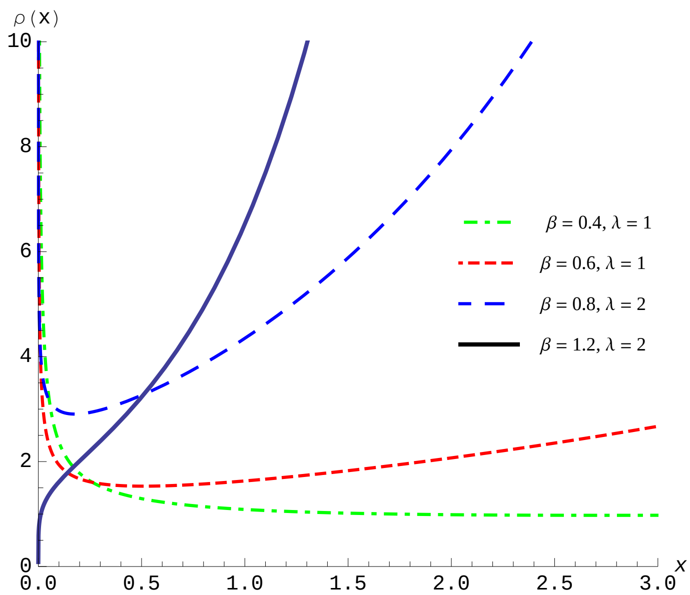

Figure 1.4: Hazard rate functions of the Chen distribution for some $\lambda \geq 1$.

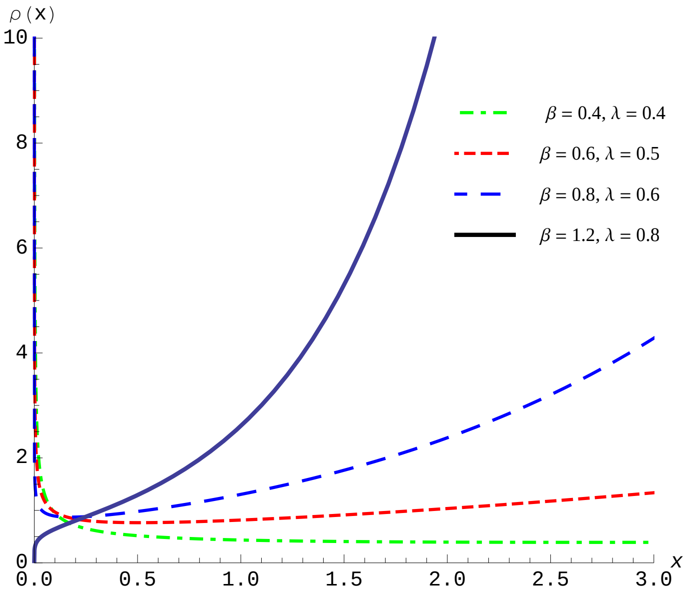

Figure 1.5: Hazard rate functions of the Chen distribution for some $\lambda < 1$.

Example 1.2.3 Lai et al. (2003) introduced a **modified Weibull distribution**, defined by the following form of cdf:

$$F(x) = 1 - \exp\left(-\alpha x^{\beta} \mathrm{e}^{\lambda x}\right),$$

where $\alpha > 0$, $\beta \geq 0$, $\lambda > 0$ are the parameters. The corresponding hazard rate function is given by the formula

$$\rho(t) = \alpha(\beta + \lambda x) x^{\beta - 1} \mathrm{e}^{\lambda x},$$

which has a turning point (i.e., minimum) at $x^* = (\sqrt{\beta} - \beta)/\lambda$ for $\beta < 1$. $\square$

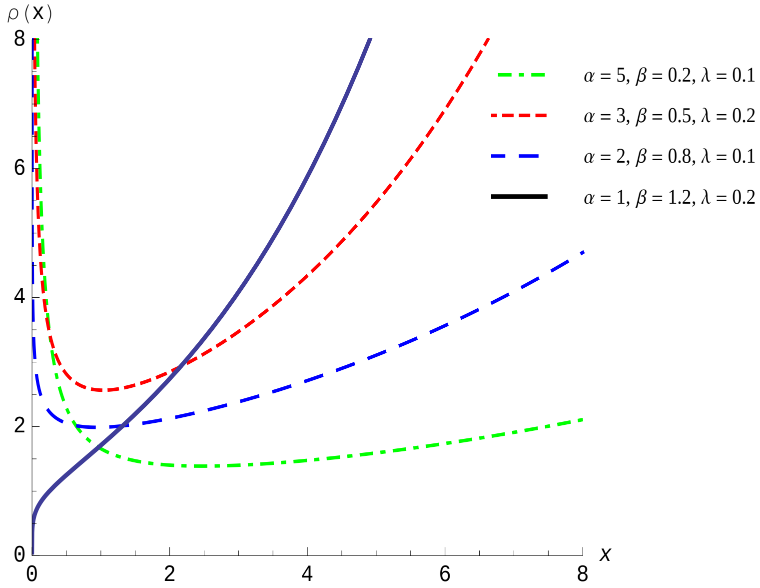

Figure 1.6: Hazard rate functions of a modified Weibull distribution.

Chapter 2

Parametric Families of Lifetime Distributions

2.1 Gamma family

2.1.1 Exponential distribution

A random variable X is said to have *exponential distribution* with parameter λ ($\lambda > 0$), which will be denoted by $\mathcal{E}(\lambda)$, if its density is defined by

$$f(x) = f(x;\lambda) = \frac{1}{\lambda}\exp\left(-\frac{x}{\lambda}\right)\mathbf{1}_{(0,\infty)}(x).$$

The cumulant distribution function is of the form

$$F(x) = \left[1 - \exp\left(-\frac{x}{\lambda}\right)\right]\mathbf{1}_{(0,\infty)}(x).$$

The characteristic function is

$$\phi(t) = \frac{1}{1 - \iota\lambda t}.$$

The k-th moment of the random variable X is defined by $E(X^k) = \lambda^k k!$. In particular,

$$E(X) = \lambda, \quad \operatorname{Var}(X) = \lambda^2.$$

The mode and median are zero and $\lambda \ln 2$, respectively, the skewness $\gamma_1 = 2$ and the excess $\gamma_2 = 6$.

The exponential distribution $\mathcal{E}(\lambda)$ is a special case of the gamma distribution, the Weibull distribution and the $\triangleright$ *negative exponential distribution* (p. 14). Namely, it is the $\mathcal{G}(1,\lambda)$, $\mathcal{W}e(1,\lambda)$ and $\mathcal{NE}(0,\lambda)$ distribution.

If $X_1, \ldots, X_n$ are independent and identically exponential $\mathcal{E}(\lambda)$ distributed, then $\sum_{i=1}^{n} X_i$ has the ▷ *gamma distribution* $\mathcal{G}(n, \lambda)$ which is equivalent to the ▷ *Erlang distribution* $\mathcal{E}r(n, \lambda)$).

In the class of continuous distributions, the exponential distribution is the only one having the following property

$$P(X > s + t | X > s) = P(X > t), \quad s > 0, t > 0,$$

which is called the **lack of memory property**. Interpreting the random variable X as a lifetime of a unit, the lack of memory property asserts that, independently of the current working time, the future working time is independent of the past and has the same distribution as the common distribution of the working time of the unit. This property plays an important role in theory of stochastic processes, especially in renewal theory and in theory of mass service and reliability.

2.1.2 Negative exponential distribution

A random variable X is said to have the *negative exponential distribution* with location parameter μ and scale parameter σ ($\mu \in \mathbb{R}, \sigma > 0$), which we denote by $\mathcal{NE}(\mu, \sigma)$, if its density is of the form

$$f(x) = f(x; \mu, \sigma) = \frac{1}{\sigma} \exp\left(-\frac{x - \mu}{\sigma}\right) \mathbf{1}_{(\mu, \infty)}(x). \tag{2.1}$$

The cdf is defined by

$$F(x) = \left[1 - \exp\left(-\frac{x - \mu}{\sigma}\right)\right] \mathbf{1}_{(\mu, \infty)}(x).$$

The negative exponential distribution is also called the **two-parameter exponential distribution** or **location and scale exponential distribution**.

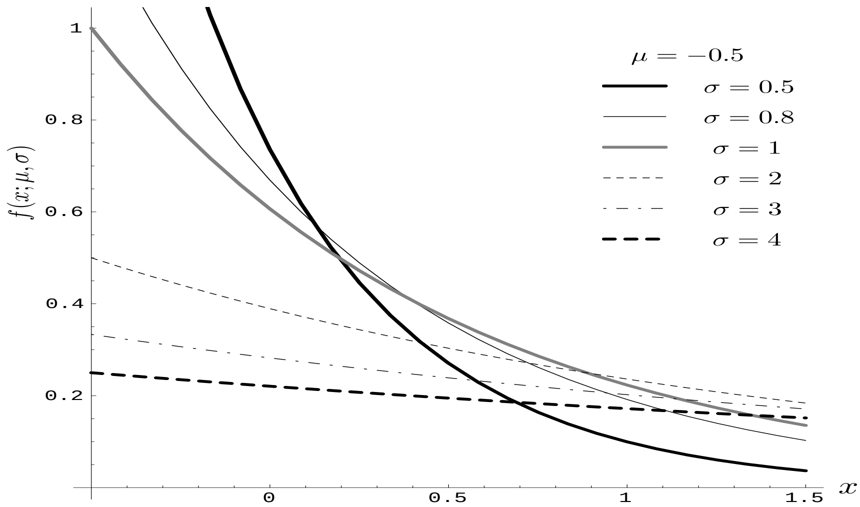

Figure 2.1: Density functions of the negative exponential distribution $\mathcal{NE}(\mu,\sigma)$

The characteristic function is of the form

$$\phi(t) = \frac{\mathrm{e}^{\iota t\mu}}{1-\iota t\sigma}.$$

The expected value and the variance of the negative exponential distribution $\mathcal{NE}(\mu,\sigma)$ are equal to

$$E(X) = \mu + \sigma, \quad \mathrm{Var}(X) = \sigma^2,$$

respectively.

There are many applications of the statistical model defined by (2.1) in reliability and lifetime examinations. The parameter μ is often interpreted as a threshold value of a minimal lifetime and it is then natural to assume that $\mu > 0$. If $\mu \geq 0$, then the negative exponential distribution is called the **left truncated exponential distribution**. In the case $\mu = 0$ the $\mathcal{NE}(\mu,\sigma)$ distribution becomes the exponential distribution $\mathcal{E}(\sigma)$.

2.1.3 Erlang distribution

A random variable X is said to have the *Erlang distribution* with parameters k and μ ($k \in \mathbb{N}$, $\mu > 0$), which will be denoted by $\mathcal{E}r(k,\mu)$, if its density has the following form

$$f(x) = \frac{1}{\mu^k(k-1)!}x^{k-1}\exp\left(-\frac{x}{\mu}\right)\mathbf{1}_{(0,\infty)}(x).$$

The expected value and the variance are equal to

$$E(X) = k\mu, \ \operatorname{Var}(X) = k\mu^2,$$

respectively.

The Erlang distribution is a special case of the $\triangleright$ *gamma distribution* (p. 19), namely it is equivalent to the gamma distribution $\mathcal{G}(k, \mu)$ for $k \in \mathbb{N}$. When $k = 1$, the distribution $\mathcal{E}r(k, \mu)$ becomes the exponential distribution $\mathcal{E}(\mu)$.

The distribution of the total lifetime until k events of the Poisson process with intensity $\lambda = 1/\mu$ have been observed has the Erlang distribution $\mathcal{E}r(k, \mu)$. The Erlang distribution is used in the theory of mass service and in examination of service times and incoming streams.

2.1.4 Rayleigh distribution

A random variable X is said to have *Rayleigh distribution* with parameter σ ($\sigma > 0$), which will be denoted by $\mathcal{R}a(\sigma)$, if its density is defined by

$$f(x) = f(x; \sigma) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right) \mathbf{1}_{(0,\infty)}(x).$$

The cdf has the form

$$F(x) = \left\{1 - \exp\left(-\frac{x^2}{2\sigma^2}\right)\right\} \mathbf{1}_{(0,\infty)}(x).$$

The parameter σ is the scale parameter.

The moments of order k are

$$E(X^k) = 2^{k/2}\sigma^k \Gamma\left(\frac{k}{2} + 1\right)$$

(Γ denotes the gamma function). In particular,

$$E(X) = \sqrt{\frac{\pi}{2}}\sigma \ (\approx 1.25331\sigma), \quad \operatorname{Var}(X) = \left(2 - \frac{\pi}{2}\right)\sigma^2 \ (\approx 0.429204\sigma^2).$$

The mode and the median are equal to σ and $\sqrt{2\ln 2}\sigma$ ($\approx$ 1.17741σ), respectively.

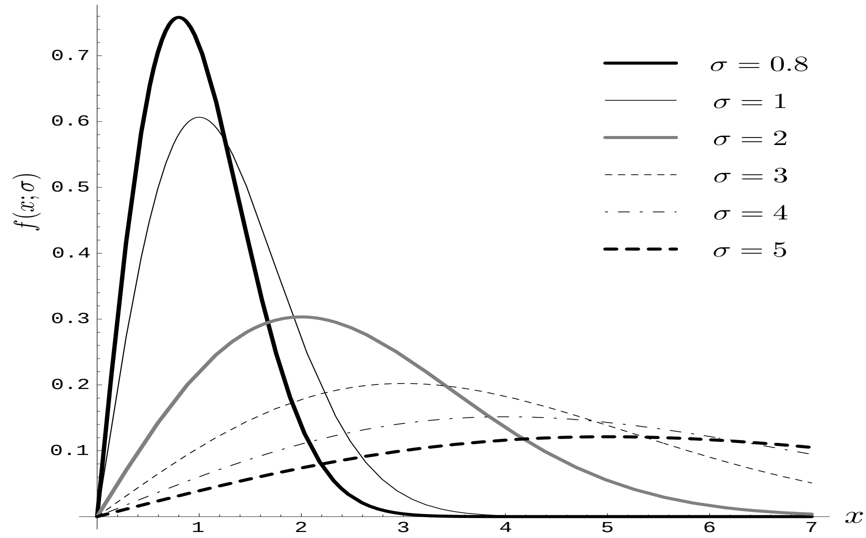

Figure 2.2: Density functions of the Rayleigh $\mathcal{R}a(\sigma)$ distribution

The Rayleigh distribution is a special case of the Weibulla distribution, namely it is equivalent to the $\mathcal{W}e(2, \sqrt{2}\sigma)$ distribution. It is used first of all in reliability theory.

2.1.5 Weibull distribution

A random variable X is said to have *Weibull distribution* with parameters α and β ($\alpha > 0, \beta > 0$), which will be denoted by $\mathcal{W}e(\alpha, \beta)$, if its density is of the form

$$\begin{aligned} f(x) = f(x; \alpha, \beta) &= \alpha\beta^{-\alpha}x^{\alpha-1} \exp\left[-\left(\frac{x}{\beta}\right)^{\alpha}\right] \mathbf{1}_{(0,\infty)}(x) \\ &= \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} \exp\left[-\left(\frac{x}{\beta}\right)^{\alpha}\right] \mathbf{1}_{(0,\infty)}(x). \end{aligned}$$

The parameter α is the shape parameter, and the parameter β is the scale parameter. The cdf of the Weibull distribution $\mathcal{W}e(\alpha, \beta)$ is

$$F(x) = \left\{1 - \exp\left[-\left(\frac{x}{\beta}\right)^{\alpha}\right]\right\} \mathbf{1}_{(0,\infty)}(x).$$

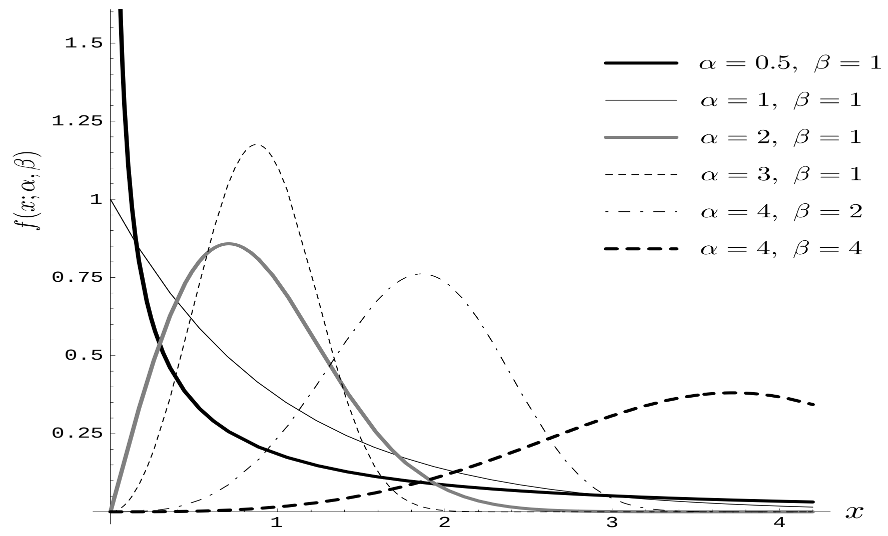

Figure 2.3: Density functions of the Weibull distribution $\mathcal{W}e(\alpha, \beta)$

The moments are defined by

$$E(X^k) = \beta^k \Gamma\left(1 + \frac{k}{\alpha}\right).$$

In particular,

$$E(X) = \beta\Gamma\left(1 + \frac{1}{\alpha}\right), \quad \mathrm{Var}(X) = \beta^2\left[\Gamma\left(1 + \frac{2}{\alpha}\right) - \Gamma^2\left(1 + \frac{1}{\alpha}\right)\right]$$

(Γ denotes the gamma function). The mode and the median are equal to 0 and $\beta(\ln 2)^{1/\alpha}$, respectively.

The moment generating function of the random variable $Y = \ln X$, when X has the Weibull distribution $\mathcal{W}e(\alpha, 1)$, is defined by

$$E\left(\mathrm{e}^{t \ln X}\right) = E(X^t) = \Gamma\left(1 + \frac{t}{\alpha}\right).$$

Special cases of the Weibull distribution are:

$\mathcal{W}e(1, \lambda)$ – $\triangleright$ *exponential distribution* $\mathcal{E}(\lambda)$;
$\mathcal{W}e(2, \sqrt{2}\sigma)$ – $\triangleright$ *Rayleigh distribution* $\mathcal{R}a(\sigma)$.

If X has the $\mathcal{W}e(\alpha, \beta)$ distribution, then the random variable $Y = -\ln X$ has a ▷ *double exponential distribution* (p. 21) $\mathcal{DE}(-\ln\beta, 1/\alpha)$ (i.e. the ▷ *extreme value distribution of type I* (p. 23)), and the random variable $-X$ has the ▷ *extreme value distribution of type III* (p. 24) with parameter α.

The Weibull distribution one obtains as a limit distribution of the minimum of independent and identically distributed random variables. It finds important applications in reliability theory by description of failure-free working times of devices.

2.1.6 Three-parameter Weibull distribution

A more general form of the Weibull distribution is the *three-parameter Weibull distribution* with additional location parameter x_0 ($x_0 \in \mathbb{R}$). We deal with such distribution if the random variable $X - x_0$ has the usual two-parameter Weibull distribution.

2.1.7 Gamma distribution

A random variable X is said to have *gamma distribution* with parameters α and λ ($\alpha > 0, \lambda > 0$), which will be denoted by $\mathcal{G}(\alpha, \lambda)$, if its density is of the form

$$f(x) = f(x; \alpha, \lambda) = \frac{1}{\lambda^\alpha \Gamma(\alpha)} x^{\alpha-1} \exp\left(-\frac{x}{\lambda}\right) \mathbf{1}_{(0,\infty)}(x),$$

where $\Gamma(\alpha)$ is the gamma function. The parameter α is the shape parameter; λ is the scale parameter.

The characteristic function is

$$\phi(t) = \frac{1}{(1 - \iota\lambda t)^\alpha}.$$

The k-th moment of the random variable X is defined by

$$E(X^k) = \frac{\lambda^k \Gamma(k + \alpha)}{\Gamma(\alpha)} = \alpha(\alpha+1)\cdots(\alpha+k-1)\lambda^k.$$

In particular,

$$E(X) = \alpha\lambda,\ E(X^2) = \alpha(\alpha+1)\lambda^2,\ \mathrm{Var}(X) = \alpha\lambda^2.$$

Special cases of the gamma distributions are:

$\mathcal{G}(1, \lambda)$ – ▷ *exponential distribution* with the parameter $\lambda, \mathcal{E}(\lambda)$;

$\mathcal{G}(k, \mu)$ – ▷ *Erlang distribution* with the parameters $(k, \mu), k \in \mathbb{N}, \mu > 0$;

The distribution $\mathcal{G}\left(\frac{n}{2}, 2\right)$ is called the **chi-square distribution** (χ^2) with n degrees of freedom.

2.1.8 Three-parameter gamma distribution

A random variable X is said to have the *three-parameter gamma distribution* with parameters α, λ and γ ($\alpha > 0, \lambda > 0, \gamma > 0$), if its density is of the form

$$f(x) = f(x; \alpha, \lambda, \gamma) = \frac{\gamma}{\lambda^{\alpha/\gamma}\Gamma(\alpha/\gamma)} x^{\alpha-1} \exp\left(-\frac{x^\gamma}{\lambda}\right) \mathbf{1}_{(0,\infty)}(x).$$

The parameter λ is the scale parameter, the parameters α and γ are the shape parameters. This distribution we denote by $\mathcal{GG}am(\alpha, \lambda, \gamma)$. The three-parameter gamma distribution is also called the **generalized gamma distribution**.

Special cases of the generalized gamma distribution are

$\mathcal{GG}am(\alpha, \lambda, 1)$ – gamma distribution $\mathcal{G}(\alpha, \lambda)$;
$\mathcal{GG}am(n/2, 2, 1)$ – chi-square distribution with n degrees of freedom $\chi^2(n)$;
$\mathcal{GG}am(\alpha, \beta^\alpha, \alpha)$ – Weibull distribution $\mathcal{W}e(\alpha, \beta)$;
$\mathcal{GG}am(2, 2\sigma^2, 2)$ – Rayleigh distribution $\mathcal{R}a(\sigma)$.

2.1.9 Four-parameter generalized gamma distribution

A random variable X is said to have the *four-parameter generalized gamma distribution* with parameters α, λ, μ and γ ($\alpha > 0, \lambda > 0, \mu \in \mathbb{R}, \gamma > 0$), if its density has the following form:

$$\begin{aligned} f(x) &= f(x; \alpha, \lambda, \gamma, \mu) \\ &= \frac{\gamma}{\lambda^{\alpha/\gamma}\Gamma(\alpha/\gamma)} (x-\mu)^{\alpha-1} \exp\left[-\frac{(x-\mu)^\gamma}{\lambda}\right] \mathbf{1}_{(\mu,\infty)}(x). \end{aligned}$$

We denote this distribution by $\mathcal{GG}am_1(\alpha, \lambda, \mu, \gamma)$. When $\mu = 0$ one obtains the density of the generalized gamma distribution, i.e. $\mathcal{GG}am_1(\alpha, \lambda, 0, \gamma) = \mathcal{GG}am(\alpha, \lambda, \gamma)$. A random variable X has the four-parameter generalized gamma distribution $\mathcal{GG}am_1(\alpha, \lambda, \mu, \gamma)$ if and only if the random variable $X - \mu$ has the generalized (three-parameter) gamma distribution $\mathcal{GG}am(\alpha, \lambda, \gamma)$.

2.2 Inverse gamma distribution

A random variable X is said to have *inverse gamma distribution* with parameters α and λ ($\alpha > 0, \lambda > 0$), which will be denoted by $\mathcal{IG}am(\alpha, \lambda)$, if its density has the following form

$$f(x) = f(x;\alpha,\lambda) = \frac{1}{\lambda^\alpha \Gamma(\alpha)} x^{-\alpha-1} \exp\left(-\frac{1}{\lambda x}\right) \mathbf{1}_{(0,\infty)}(x).$$

The expected value and the variance are given by

$$E(X) = \frac{1}{(\alpha-1)\lambda}, \text{ jeżeli } \alpha > 1; \quad \mathrm{Var}(X) = \frac{1}{(\alpha-1)^2(\alpha-2)\lambda^2}, \text{ jeżeli } \alpha > 2,$$

respectively. The random variable $1/X$ has the ▷ *gamma distribution* (p. 19) $\mathcal{G}(\alpha,\lambda)$.

2.3 Gompertz distribution

A random variable X follows the *Gompertz distribution* with parameters α and β ($\alpha > 0, \beta > 0$), denoted by $\mathcal{G}om(\alpha,\beta)$, if its density function is defined by

$$f(x) = \beta e^{\alpha x} \exp\left[\frac{\beta}{\alpha}(1 - e^{\alpha x})\right] \mathbf{1}_{[0,\infty)}(x).$$

This distribution was introduced by Gompertz in 1825 to describe mortality curves.

2.4 Gumbel distributions

2.4.1 Double exponential distribution

A random variable X is said to have *double exponential distribution* with location parameter μ and scale parameter σ ($\mu \in \mathbb{R}, \sigma > 0$), which will be denoted by $\mathcal{DE}(\mu,\sigma)$, if its density function has the following form:

$$\begin{aligned} f(x) = f(x;\mu,\sigma) &= \frac{1}{\sigma} \exp\left[-\exp\left(-\frac{x-\mu}{\sigma}\right) - \frac{x-\mu}{\sigma}\right] \\ &= \frac{1}{\sigma} \exp\left(-\frac{x-\mu}{\sigma}\right) \exp\left[-\exp\left(-\frac{x-\mu}{\sigma}\right)\right], \quad x \in \mathbb{R}. \end{aligned}$$

The cdf is defined by

$$F(x) = \exp\left[-\exp\left(-\frac{x-\mu}{\sigma}\right)\right].$$

The characteristic function has the following form:

$$\phi(t) = \mathrm{e}^{\iota\mu t}\Gamma(1 - \iota\sigma t)$$

(Γ denotes the gamma function).

The expected value and the variance of the $\mathcal{DE}(\mu, \sigma)$ distribution are equal to

$$E(X) = \mu + \sigma C, \quad \mathrm{Var}(X) = \frac{\pi^2\sigma^2}{6},$$

respectively, where $C = -\Gamma'(1) = 0.57721566490\ldots$ denotes the Euler constant. The skewness of this distribution is $\gamma_1 = 12\sqrt{6}\zeta(3)/\pi^3 \approx 1.1395$, and the excess $\gamma_2 = 2.4$. The mode and the median are equal to μ and $\mu - \sigma \ln\ln 2$, respectively.

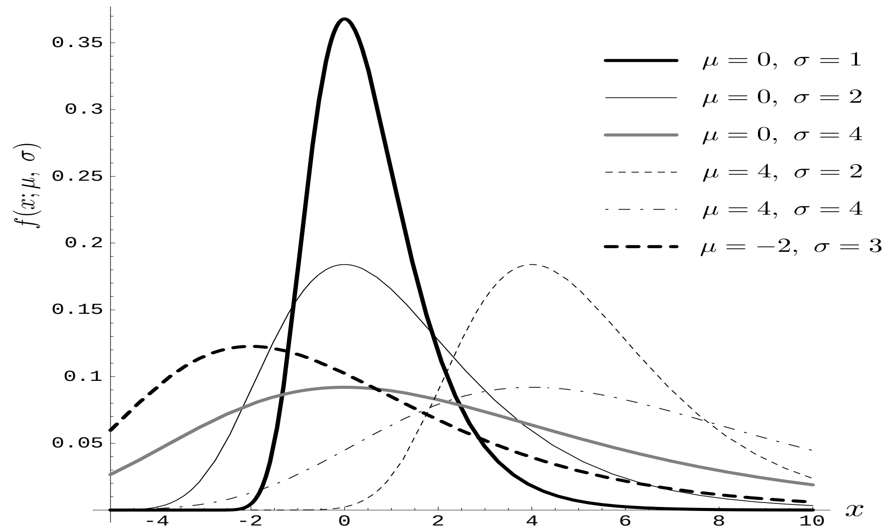

Figure 2.4: Density functions of the double exponential distribution $\mathcal{DE}(\mu, \sigma)$

If X follows the $\mathcal{DE}(\mu, \sigma)$ distribution, then the random variable $Y = \exp(-X)$ has the Weibull distribution $\mathcal{W}e(1/\sigma, \exp(-\mu))$, which becomes the exponential distribution $\mathcal{E}(\exp(-\mu))$ for $\sigma = 1$. The double exponential distribution is equivalent to the $\triangleright$ *extreme value distribution of type I* (p. 23). For this reason the double exponential distribution is often called the **extreme value distribution** or the **Gumbel distribution** $\mathcal{G}um(\mu, \sigma)$.

The double exponential distribution is used in reliability theory and among others in evaluating maximal precipitations, water levels and earthquakes.

2.4.2 Generalized Gumbel distribution

The *generalized Gumbel distribution* is usually meant as the $\triangleright$ *generalized extreme value distribution* (p. 24) $\mathcal{GEV}(\mu, \sigma, \gamma)$. The name of generalized Gumbel distribution is often referred to the following slight modification of the generalized extreme value distribution. A random variable X is said to have the **generalized Gumbel distribution** with parameters μ, σ and ν ($\mu \in \mathbb{R}$, $\sigma > 0$, $\nu \in \mathbb{R}\backslash\{0\}$), if its density function has the following form:

$$f(x) = f(x; \mu, \sigma, \nu)$$
$$= \begin{cases} \dfrac{1}{\sigma}\Big(1 + \nu\dfrac{x-\mu}{\sigma}\Big)^{(1/\nu)-1} \exp\left[-\Big(1 + \nu\dfrac{x-\mu}{\sigma}\Big)^{1/\nu}\right] \mathbf{1}_{[\mu-\sigma/\nu,\infty)}(x), & \nu > 0, \\ \dfrac{1}{\sigma}\Big(1 + \nu\dfrac{x-\mu}{\sigma}\Big)^{(1/\nu)-1} \exp\left[-\Big(1 + \nu\dfrac{x-\mu}{\sigma}\Big)^{1/\nu}\right] \mathbf{1}_{(-\infty,\mu-\sigma/\nu]}(x), & \nu < 0. \end{cases}$$

The distribution defined above we denote by $\mathcal{GG}um(\mu, \sigma, \nu)$. In the limit case, when $\nu \to 0$, one obtains the Gumbel distribution $\mathcal{G}um(\mu, \sigma)$ ($\triangleright$ *extreme value distribution* (p. 23)).

A random variable X has the $\mathcal{GG}um(\mu, \sigma, \nu)$ distribution when

$$\left(1 + \nu\frac{X-\mu}{\sigma}\right)^{1/\nu} \sim \mathcal{E}(1).$$

2.4.3 Extreme value distribution

Let X_1, X_2, ... be a sequence of independent and identically distributed random variables with a cdf $F(x)$. Denote $W_n = \max\{X_1, \ldots, X_n\}$ and let $(a_n), (b_n), b_n > 0, n = 1, 2, \ldots,$ be sequences of real numbers. If the sequence of random variables $Y_n = (W_n - a_n)/b_n$ converges in law to a random variable X with non-degenerated distribution defined by a cdf G, then G can take only one of the following form

(1) *extreme value distribution of type I*:

$$G_1(x) = \exp\left[-\exp\left(-\frac{x-\mu}{\sigma}\right)\right], \; x \in \mathbb{R};$$

(2) *extreme value distribution of type II*:

$$G_2(x) = \exp\left[-\left(\frac{x-\mu}{\sigma}\right)^{-\alpha}\right] \mathbf{1}_{[\mu,\infty)}(x);$$

(3) *extreme value distribution of type III*:

$$G_3(x) = \begin{cases} \exp\left[-\left(-\dfrac{x-\mu}{\sigma}\right)^{\alpha}\right] & \text{for } x \le \mu, \\ 1 & \text{for } x > \mu, \end{cases}$$

where $\mu \in \mathbb{R}$, $\sigma > 0$ and $\alpha > 0$ are parameters. The corresponding distributions of the random variable $-X$ are also called the extreme value distributions.

The existence of the sequences $(a_n), (b_n), b_n > 0, n = 1, 2, \ldots$, such that the cdf's of the random variables Y_n converge to a cdf of non-degenerated distribution, as well as the type of the limit distribution determined by G depends on the cdf F.

The extreme value distribution of type I is called the **Gumbel distribution**. We denote this distribution by $\mathcal{G}um(\mu, \sigma)$. The distributions of type II and III can be come down to the $\mathcal{G}um(\mu, \sigma)$ distribution by using the simple transformations $Z = \log(X - \mu)$ and $Z = -\log(X - \mu)$. Because of this, all the types of distributions above are often called the Gumbel distributions. The extreme value distribution of type II is also the **Fréchet distribution**.

The extreme value distribution of type I corresponds to the ▷ *double exponential distribution* (p. 21). If X has the distribution of type I, then $Z = \exp[-(X-\mu)/\sigma]$ follows the exponential distribution $\mathcal{E}(1)$. If Y has the extreme value distribution of type III, then the cdf of $-Y$ is defined by $x \mapsto 1 - G_3(-x)$, which leads to the ▷ *Weibull distribution* (p. 17).

The extreme value distributions and their generalizations (▷ *generalized extreme value distribution* (p. 24)) have applications among others in natural phenomena analysis (such as rapid heavy rains, floods, hurricanes, earthquakes, air pollution, corrosion), in reliability analysis and in insurance analysis.

2.4.4 Generalized extreme value distribution

A random variable X is said to have *generalized extreme value distribution* with parameters μ, σ and γ ($\mu, \gamma \in \mathbb{R}$, $\sigma > 0$), if its cdf is defined by

$$\begin{aligned} F(x) &= F(x; \mu, \sigma, \gamma) \\ &= \begin{cases} \exp\left[-\left(1 - \gamma\dfrac{x-\mu}{\sigma}\right)^{1/\gamma}\right] \mathbf{1}_{(-\infty, \mu+\sigma/\gamma]}(x), & \text{if } \gamma > 0, \\ \exp\left[-\left(1 - \gamma\dfrac{x-\mu}{\sigma}\right)^{1/\gamma}\right] \mathbf{1}_{[\mu+\sigma/\gamma, \infty)}(x), & \text{if } \gamma < 0, \\ \exp\left[-\exp\left(-\dfrac{x-\mu}{\sigma}\right)\right], \ x \in \mathbb{R}, & \text{if } \gamma = 0. \end{cases} \end{aligned}$$

This distribution we denote by $\mathcal{GEV}(\mu,\sigma,\gamma)$. The case $\gamma<0$ covers the extreme value distribution of type II; the case $\gamma>0$ corresponds to the extreme value distribution of type III; the case $\gamma=0$ leads to the extreme value distribution of type I (the Gumbel distribution).

The density function corresponding to the $\mathcal{GEV}(\mu,\sigma,\gamma)$ distribution has the following form:

$$
\begin{aligned}
&f(x) = f(x;\mu,\sigma,\gamma)\\
&= \begin{cases}
\dfrac{1}{\sigma}\Big(1-\gamma\dfrac{x-\mu}{\sigma}\Big)^{(1/\gamma)-1}\exp\left[-\Big(1-\gamma\dfrac{x-\mu}{\sigma}\Big)^{1/\gamma}\right]\mathbf{1}_{(-\infty,\mu+\sigma/\gamma]}(x), & \gamma>0,\\
\dfrac{1}{\sigma}\Big(1-\gamma\dfrac{x-\mu}{\sigma}\Big)^{(1/\gamma)-1}\exp\left[-\Big(1-\gamma\dfrac{x-\mu}{\sigma}\Big)^{1/\gamma}\right]\mathbf{1}_{[\mu+\sigma/\gamma,\infty)}(x), & \gamma<0,\\
\dfrac{1}{\sigma}\exp\Big(-\dfrac{x-\mu}{\sigma}\Big)\exp\Big[-\exp\Big(-\dfrac{x-\mu}{\sigma}\Big)\Big],\quad x\in\mathbb{R}, & \gamma=0.
\end{cases}
\end{aligned}
$$

2.5 Lognormal distributions

A random variable X is said to have the *lognormal distribution* with parameters μ and σ^2 $(\mu\in\mathbb{R},\sigma>0)$, which we denote by $\mathcal{LN}(\mu,\sigma^2)$, if its density function is of the form

$$f(x)=f(x;\mu,\sigma^2)=\frac{1}{\sqrt{2\pi}\sigma x}\exp\left[-\frac{(\ln x-\mu)^2}{2\sigma^2}\right]\mathbf{1}_{(0,\infty)}(x).$$

The moments of order k $(k=1,2,\ldots)$ are defined by

$$E(X^k)=\exp\left(\frac{1}{2}k^2\sigma^2+k\mu\right).$$

In particular,

$$E(X)=\exp\left(\frac{1}{2}\sigma^2+\mu\right),\quad \operatorname{Var}(X)=\exp(\sigma^2+2\mu)\left(\exp\sigma^2-1\right).$$

The mode is equal to $\exp(\mu-\sigma^2)$, and the median is $\exp(\mu)$.

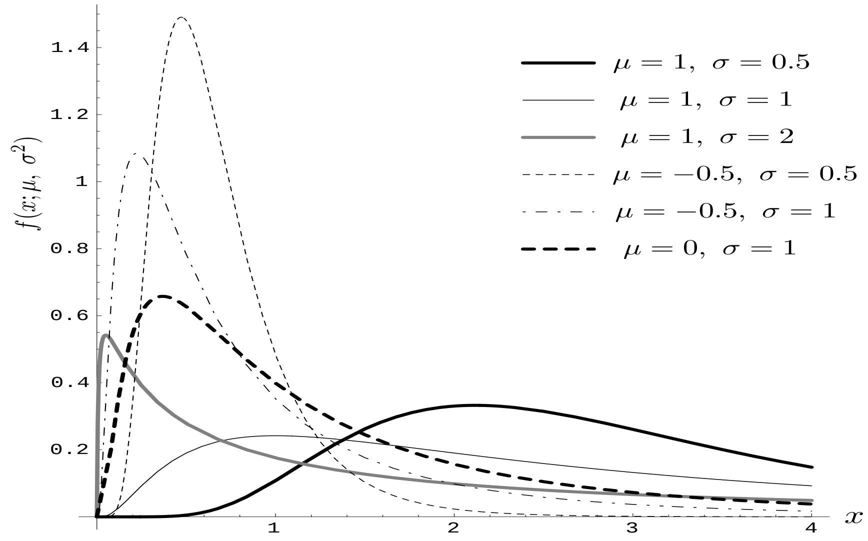

Figure 2.5: Density functions of the lognormal distribution $\mathcal{LN}(\mu, \sigma^2)$

A random variable X has the $\mathcal{LN}(\mu, \sigma^2)$ distribution, if $\ln X$ follows the normal distribution $\mathcal{N}(\mu, \sigma^2)$. If Y has the normal distribution $\mathcal{N}(0, 1)$, then $X = \exp(\sigma Y + \mu)$ is $\mathcal{LN}(\mu, \sigma^2)$ distributed.

If $X_1, \ldots, X_k$ are independent random variables with X_i $(i = 1, \ldots, k)$ having the $\mathcal{LN}(\mu_i, \sigma_i^2)$ distribution, then $\prod_{i=1}^k a_i X_i$, where $a_i, i = 1, \ldots, k$, are any positive constants, has the lognormal distribution $\mathcal{LN}(\sum_{i=1}^k (\mu_i + \ln a_i), \sum_{i=1}^k \sigma_i^2)$.

The lognormal distribution is used in describing the life data resulting from a single semiconductor failure mechanism or a closely related group of failure mechanisms. This is a suitable model for patients of tuberculosis or other diseases where the potential for death increases early in the disease and then decreases when the effect of the treatment is evident. This distribution is widely applied in statistical examinations in physics, geology, economics and biology.

2.6 Pareto family

2.6.1 Pareto distribution

A random variable X is said to have the *Pareto distribution* with parameters x_0 and α $(x_0 > 0, \alpha > 0)$, which we denote by $\mathcal{P}a(x_0, \alpha)$, if its density function has the following form:

$$f(x) = \frac{\alpha}{x_0}\left(\frac{x_0}{x}\right)^{\alpha+1}\mathbf{1}_{(x_0,\infty)}(x).$$

The cdf is defined by

$$F(x) = \left[1 - \left(\frac{x_0}{x}\right)^{\alpha}\right]\mathbf{1}_{(x_0,\infty)}(x), \quad x_0 > 0, \alpha > 0.$$

The moment generating function of the random variable $Y = \ln X$, when X has the $\mathcal{P}a(x_0,\alpha)$ distribution, is expressed by the formula

$$\psi_Y(t) = E\left(\mathrm{e}^{t\ln X}\right) = E(X^t) = \frac{\alpha x_0^t}{\alpha - t}.$$

The expectation and the variance of the random variable X are given by

$$E(X) = \frac{\alpha x_0}{\alpha - 1} \quad \text{for} \quad \alpha > 1, \quad \mathrm{Var}(X) = \frac{\alpha x_0^2}{(\alpha-1)^2(\alpha-2)} \quad \text{for } \alpha > 2,$$

respectively. The median of the random variable X equals $2^{1/\alpha}x_0$.

If X has the $\mathcal{P}a(x_0,\alpha)$ distribution, then the random variable $\ln(X/x_0)$ follows the exponential distribution $\mathcal{E}(1/\alpha)$.

The Pareto distributed random variable can take its values only above a positive level x_0. minimalnym.

2.6.2 Generalized Pareto distribution

The *generalized Pareto distribution* is defined by the following form of the cdf

$$F(x) = F(x;\beta,\gamma) = \begin{cases} \left[1 - \left(1 - \frac{\gamma x}{\beta}\right)^{1/\gamma}\right]\mathbf{1}_{(0,\beta/\gamma)}(x), & \text{gdy } \gamma > 0, \\ \left[1 - \left(1 - \frac{\gamma x}{\beta}\right)^{1/\gamma}\right]\mathbf{1}_{(0,\infty)}(x), & \text{gdy } \gamma < 0, \\ \left(1 - \mathrm{e}^{-x/\beta}\right)\mathbf{1}_{(0,\infty)}(x), & \text{gdy } \gamma = 0, \end{cases}$$

where $\beta > 0$. This distribution we denote by $\mathcal{GP}a(\beta,\gamma)$. The density function of the $\mathcal{GP}a(\beta,\gamma)$ distribution has the form

$$f(x) = f(x;\beta,\gamma) = \begin{cases} \frac{1}{\beta}\left(1 - \frac{\gamma x}{\beta}\right)^{(1/\gamma)-1}\mathbf{1}_{(0,\beta/\gamma)}(x), & \text{gdy } \gamma > 0, \\ \frac{1}{\beta}\left(1 - \frac{\gamma x}{\beta}\right)^{(1/\gamma)-1}\mathbf{1}_{(0,\infty)}(x), & \text{gdy } \gamma < 0, \\ \frac{1}{\beta}\mathrm{e}^{-x/\beta}\mathbf{1}_{(0,\infty)}(x), & \text{gdy } \gamma = 0. \end{cases}$$

Let us note that for $\gamma = 0$ this is the exponential distribution with mean β, and for $\gamma = 1$ this is the uniform distribution on the interval $(0, \beta)$, i.e. $\mathcal{GP}a(\beta, 0) = \mathcal{E}(\beta)$ and $\mathcal{GP}a(\beta, 1) = \mathcal{U}(0, \beta)$.

If $X \sim \mathcal{GP}a(\beta, \gamma)$, then the random variable $X - u$, under the condition $X > u$, has the generated Pareto distribution $\mathcal{GP}a(\beta - \gamma u, \gamma)$, given $\beta - \gamma u > 0$. This property is called **threshold-stability** (compare with the lack of memory property of the exponential distribution).

The ▷ *failure rate function* (p. 2) of the $\mathcal{GP}a(\beta, \gamma)$ distribution has the form $\rho(x) = (\beta - \gamma x)^{-1}$. Thus, this is the IFR distribution for $\gamma > 0$, and the DFR distribution for $\gamma < 0$. For $\gamma = 0$ this is the IFR as well as the DFR distribution.

If $1 + \gamma k > 0$, then there exists the moment of order k of the random variable X distributed according to $\mathcal{GP}a(\beta, \gamma)$, and it is defined by the formula

$$E\left[\left(1 - \frac{\gamma X}{\beta}\right)^k\right] = \frac{1}{1 + \gamma k}.$$

In particular,

$$E(X) = \frac{\beta}{1 + \gamma}, \quad \mathrm{Var}(X) = \frac{\beta^2}{(1 + \beta)^2 (1 + 2\beta)}.$$

2.7 Burr distributions

From among the distributions belonging to the family of *Burr distributions* we present the following ones:

(1) the distribution $\mathcal{B}u^{(1)}(\lambda, \mu)$ with shape parameters λ and μ ($\lambda > 0, \mu > 0$), whose density function is defined by

$$f(x) = \lambda \mu x^{\lambda - 1} (1 + x^\lambda)^{-(\mu + 1)} \mathbf{1}_{(0,\infty)}(x);$$

(2) the distribution $\mathcal{B}u^{(2)}(\lambda, \mu)$ with parameters λ and μ ($\lambda > 0, \mu > 0$), whose density function is defined by

$$f(x) = \lambda \mu x^{\lambda - 1} (\mu + x^\lambda)^{-2} \mathbf{1}_{(0,\infty)}(x).$$

The cdf of the $\mathcal{B}u^{(2)}(\lambda, \mu)$ distribution has the form

$$F(x) = \frac{x^\lambda}{\mu + x^\lambda} \mathbf{1}_{(0,\infty)}(x).$$

2.8 Time-transformed exponential family

Consider the exponential family of distributions with the following density

$$f(x;\omega) = s'(x)\omega \exp[-\omega s(x)]\mathbf{1}_{(\mu,\infty)}(x),$$

where $s(x)$ is strictly increasing and differentiable with $s(\mu) = 0$. The cumulative distribution function is of the form

$$F(x;\omega) = \{1 - \exp[-\omega s(x)]\}\mathbf{1}_{(\mu,\infty)}(x).$$

The parameter ω is unknown and μ is known. This exponential family covers many distributions serving as lifetime distributions in reliability models. Some of important members of this family are contained in Table 2.1.

Distribution and its density	$s(x)$	ω
Exponential $\mathcal{E}(\eta)$, $\eta > 0$ $f(x;\eta) = \eta\exp(-\eta x)\mathbf{1}_{(0,\infty)}(x)$	x	η
Rayleigh $\mathcal{R}a(\eta)$ $f(x;\eta) = \frac{x}{\eta^2}\exp(-\frac{x^2}{2\eta^2})\mathbf{1}_{(0,\infty)}(x)$	$x^2/2$	η^{-2}
Pareto $\mathcal{P}a(x_0,\eta)$, $x_0 > 0$ – known $f(x;\eta) = \frac{\eta}{x_0}(\frac{x_0}{x})^{\eta+1}\mathbf{1}_{(x_0,\infty)}(x)$	$\ln(x/x_0)$	η
Weibull $\mathcal{W}e(\alpha,\eta)$, $\alpha > 0$ – known $f(x;\eta) = \alpha\eta^\alpha x^{\alpha-1}\exp[-(x\eta)^\alpha]\mathbf{1}_{(0,\infty)}(x)$	x^α	η^α
Location-and-scale parameter exponential $\mathcal{NE}(\mu,\eta)$, $\mu \in \mathbb{R}$ – known $f(x;\eta) = \eta\exp[-\eta(x-\mu)]\mathbf{1}_{(\mu,\infty)}(x)$	$x - \mu$	η

Table 2.1: Members of the time-transformed exponential family

Chapter 3

Most Commonly used Models for Repairable Systems

3.1 Perfect repair and minimal repair models

Let $N(t)$ denote the number of failures (events) in the time interval $(0, t]$ and let T_i be the time of the ith failure. The times T_i are called **failure times** or **event times**. Define $T_0 = 0$ and denote $X_i = T_i - T_{i-1}$, $i = 1, 2, \ldots,$ – the time between failure number $i-1$ and failure number i. The times X_i are called **working times** or **waiting times** and also **inter-arrival times**.

The observed sequence $\{T_i, i = 1, 2, \ldots\}$ of occurrence times $T_1, T_2, \ldots$ (failure times) forms a point process, and $\{N(t), t \geq 0\}$ is the corresponding counting process.

Following Ascher and Feingold (1984), a **repairable system** is defined to be a system which, after failing to perform one or more of its functions satisfactorily, can be restored to fully satisfactory performance by any method, other than replacement of the entire system.

A restoration wherein a failed system (device) is returned to operable condition is called 'repair'. It could involve replacing failed components by working ones, restoring broken connections, mending it or any part of it by machining, cleaning, lubricating etc.

In the context of failure-repair models it is assumed here that all repair times are equal to 0. In practice this corresponds to the situation, when repair actions are conducted immediately or the repair times can be neglected with comparison to the working times X_i.

A repair under which a failed system is replaced with a new identical one is called a **perfect repair**.

By a **minimal repair** we mean a repair of limited effort wherein the device is returned to the operable state it was in just before failure.

Formal definitions of minimal repair are given below. For a comparison of definitions of this notion, see, for example, Arjas (2002).

A definition of minimal repair used by Sheu (1990) and then by others, e.g., Bagai and Jain (1994), Bae and Lee (2001), uses the survival function as below.

Definition 3.1.1 If F is a cdf of lifetime distribution of a device, the cdf of lifetime X following a **perfect repair** is always F; but the cdf of lifetime distribution following a **minimal repair** performed at age s is given by

$$P(X > t | X > s) = S(t|s) = \frac{S(s+t)}{S(s)}.$$

□

The definition of minimal repair written in a more formal way by Nakagawa and Kowada (1983) as follows: suppose the system begins to operate at time 0 and that the time for repair is negligible. Let $T_0, T_1, T_2, \ldots$ ($T_0 = 0$) denote the system of failure times and let $X_i = T_i - T_{i-1}$, $i = 1, 2, \ldots$, denote, as above, the times between failures (the working times).

Definition 3.1.2 Let $F(t) = P(X_1 \le t)$, $t \ge 0$. The system undergoes **minimal repairs** at failures if and only if

$$P\Big(X_n < u \Big| \sum_{i=1}^{n-1} X_i = t\Big) = \frac{F(t+u) - F(t)}{F(t)}, \quad n = 2, 3, \ldots$$

for $u > 0$, $t \ge 0$, such that $F(t) < 1$. □

Following Andersen et al. (1993, Section II.4) the **conditional intensity** $\lambda(t|\mathcal{F}_{t-})$ of a point process $\{N(t), t \ge 0\}$ is defined by

$$\begin{aligned} \gamma(t) &:= \lambda(t|\mathcal{F}_{t-}) \\ &= \lim_{\Delta t \downarrow 0} \frac{P(\text{failure in a point process in } [t, t+\Delta t)|\mathcal{F}_{t-})}{\Delta t}, \quad t > 0, \quad (3.1) \end{aligned}$$

where $\mathcal{F}_{t-} = \sigma\{N(u), u < t\}$. Thus, $\gamma(t)\Delta t$ is approximately the probability of failure in the time interval $[t, t+\Delta t)$, conditional on the experienced failure history before time t.

For a ▷ *renewal process* (p. 35) with inter-arrival distribution F it is well known that $\gamma(t) = \rho(t - T_{N(t-)})$, where ρ is the hazard rate corresponding to the distribution F, and $t - T_{N(t-)}$ is the time since the last failure strictly before time t. Thus at each failure, the conditional intensity is returned to what it was at time 0, which means that the system is as good as new immediately after each failure. This is a **perfect repair** model.

On the other hand, a $\triangleright$ *non-homogeneous Poisson process* (p. 36) with intensity function $\lambda(t)$ has the conditional intensity $\gamma(t) = \lambda(t)$. This means that the conditional intensity after a failure is exactly as if no failure had ever occurred. Thus at failures, the system is only restored to a condition where it is exactly as good (or bad) as it was immediately before the failure. this is a **minimal repair** model.

3.2 Models for repairable systems

In this section we recall the definitions of the most common point processes which are used to model of repairable systems. The two point processes, non-homogeneous Poisson process $\{N(t), t \geq 0\}$ and renewal process $\{T_i, i = 1, 2, \ldots\}$, are widely investigated in the literature on reliability to model minimal repairs and perfect repairs (renewals). In Section 3.2.6, a wider class of processes, the class of so called trend-renewal processes, is described, which covers non-homogeneous Poisson and renewal processes.

3.2.1 Homogeneous Poisson process

Definition 3.2.3 The process $N(t), t \geq 0$ is said to be a (homogeneous) **Poisson process** with rate (or intensity) $\lambda > 0$ if

(i) $N(0) = 0$;

(ii) the number of events $N(s_2) - N(s_1)$ and $N(t_2) - N(t_1)$ in disjoint time intervals $(s_1, s_2]$ and $(t_1, t_2]$ are independent (independent increments);

(iii) the distribution of the number of events in a certain interval depends only on the length of the interval and not on its position (stationary increments);

(iv) $\displaystyle\lim_{\Delta t \to 0} \frac{P(N(\Delta t) = 1)}{\Delta t} = \lambda$;

(v) $\displaystyle\lim_{\Delta t \to 0} \frac{P(N(\Delta t) \geq 2)}{\Delta t} = 0.$

□

The process defined above will be denoted by HPP(λ). Note that if the random variables $X_1, X_2, \ldots$ (waiting times) are independent and exponentially distributed $\mathcal{E}(1/\lambda)$, the counting process $\{N(t), t \geq 0\}$ is the HPP(λ). The corresponding sequence $\{T_i, i = 1, 2, \ldots\}$ is also called the HPP(λ).

Let us note that in the HPP(λ) the number of events in an interval of length t is a random variable having the Poisson distribution $\mathcal{P}(\lambda t)$, and that the time of the n-th event is a random variable $\mathcal{G}(n, 1/\lambda)$ distributed.

Major weakness of a HPP are the constant rate assumption and the fact that the distribution of the number of events in an interval depends only on the length of the interval and not on its position.

Poisson process simulation

To generate n first successive event times of the HPP(λ) one can use the fact that the inter-arrival times $X_i, i = 1, 2, \ldots,$ are independent random variables having the exponential distribution $\mathcal{E}(1/\lambda)$. One of the simulation methods of the HPP(λ) is then generating the inter-arrival times $X_i, i = 1, 2, \ldots,$ according to the inversion method formula $X_i = -\frac{1}{\lambda} \ln U_i$, where U_i are random numbers from the uniform distribution $\mathcal{U}(0, 1)$. Then

$$T_j = \sum_{i=1}^{j} X_i, j = 1, \ldots, n,$$

are the successive n event times of the HPP(λ).

Let us note that to generate the event times on a given interval time $(0, t)$, the number of failures in this time interval is determined by the formula

$$\inf \left\{ n : \sum_{i=1}^{n} X_i > t \right\} - 1.$$

Thus in generating the HPP(λ) in the interval $(0, t)$ one can use the following algorithm:

1: $s = 0, n = 0$;

2: generate the random number $U \sim \mathcal{U}(0, 1)$;

3: set $X = -\frac{1}{\lambda} \ln U$;

4: $s = s + X$; if $s > t$, then stop the procedure;

5: $n = n + 1$, $T_n = s$;

6: go to step 2.

The last value n in the algorithm above represents the number of failures in the process in the interval $(0, t)$, and $T_1, \ldots, T_n$ are the successive n event times of the HPP(λ) in this time interval.

Generating the HPP(λ) until a given number of failures appears is conducted according to the following main formula:

$$T_i = \left(T_{i-1} + \frac{1}{\lambda} \ln \frac{1}{1 - U_i} \right), \quad i = 1, 2, \ldots, \tag{3.2}$$

$T_0 = 0$, where U_i are random numbers from uniform distribution $\mathcal{U}(0, 1)$. The generating formula is equivalent to

$$T_i = \left(T_{i-1} - \frac{1}{\alpha} \ln U_i \right), \quad i = 1, 2, \ldots,$$

but for numerical computation reasons the first formula is more useful.

3.2.2 Renewal process

Definition 3.2.4 The process $\{N(t), t \geq 0\}$ is a **renewal process** if the random variables $X_1, X_2, \ldots$ are independent and identically distributed with cumulative distribution function (cdf) F with $F(0) = 0$. □

The renewal process will be denoted by RP(F). The sequence $\{T_i, i = 1, 2, \ldots\}$ is called the RP(F) too. If F is the cdf of the exponential distribution $\mathcal{E}(\lambda)$, then RP(F) is HPP(λ).

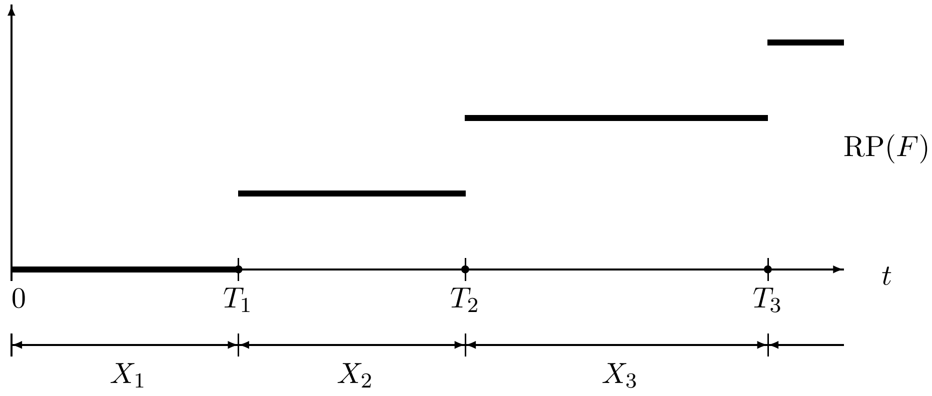

Figure 3.1: The idea of the RP model

3.2.3 Non-homogeneous Poisson process

Definition 3.2.5 A process $\{N(t), t \geq 0\}$ is a **non-homogeneous Poisson process** with intensity function $\lambda(t), t \geq 0$, if

(i) $N(t) = 0$, i.e. there are no events at time 0;

(ii) the numbers of events $N(s_2) - N(s_1)$ and $N(t_2) - N(t_1)$ in disjoint intervals $(s_1, s_2]$ and $(t_1, t_2]$ are independent random variables (independent increments);

(iii) there exists a function $\lambda(t)$ such that

$$\lim_{\Delta t \to 0} \frac{P(N(t + \Delta t) - N(t) = 1)}{\Delta t} = \lambda(t);$$

(iv) for each $t > 0$, $\lim_{\Delta t \to 0} \dfrac{P(N(t + \Delta t) - N(t) \geq 2)}{\Delta t} = 0.$

□

The process defined above will be denoted by NHPP($\lambda(\cdot)$).

The important consequence of conditions (i)–(iv) is that the number of failures in the interval $(s, t]$, $t > s$, has the Poisson distribution with the parameter

$\int_s^t \lambda(u)du$, i.e.,

$$P(N(t) - N(s) = k) = \frac{\left(\int_s^t \lambda(u)du\right)^k}{k!} \exp\left(-\int_s^t \lambda(u)du\right)$$

The functions $\lambda(t)$ and $\Lambda(t) = \int_0^t \lambda(u)du$ are called the **intensity function** and the **cumulative intensity function** of the process, respectively.

Of course, if $\lambda(t) \equiv \lambda$, then the NHPP($\lambda(\cdot)$) is the HPP($\lambda$).

Note that the definition of a NHPP relaxes the stationarity assumption of a HPP.

Fact 3.2.1 The process $\{N(t), t \geq 0\}$ is an NHPP($\lambda(\cdot)$) if the time-transformed process $\Lambda(T_1), \Lambda(T_2), \ldots$ forms an HPP(1), i.e., if $N(t) = \widetilde{N}(\Lambda(t))$, where $\{\widetilde{N}(t), t \geq 0\}$ is an HPP(1). □

Theorem 3.2.1 Suppose that events are occurring according to an HPP(λ). Suppose that, independently of anything that occurred before, an event that happens at time t is counted with probability $p(t)$. Then the process $\{N(t), t \geq 0\}$ of counted events constitutes an NHPP($\lambda p(t)$). □

The mean value function of an NHPP($\lambda(\cdot)$) is $E(N(t)) = \Lambda(t)$. This function gives the mean (expected) number of failures up to time t.

Let F_i denote the distribution function of the occurrence times $T_i, i = 1, 2, \ldots,$ and f_i be the density function. We have

$$\begin{aligned}
F_i(t) &= P\left(T_i \leq t\right) = P\left(N(t) \geq i\right) = 1 - \sum_{k=1}^{i-1} P\left(N(t) = k\right) \\
&= 1 - \sum_{k=0}^{i-1} \frac{\left(\int_0^t \lambda(y)\,dy\right)^k}{k!} e^{-\left(\int_0^t \lambda(y)\,dy\right)} = 1 - \sum_{k=0}^{i-1} \frac{[\Lambda(t)]^k}{k!} e^{-\Lambda(t)}
\end{aligned}$$

and

$$\begin{aligned}
f_i(t) &= \frac{dF_i(t)}{dt} = -\sum_{k=0}^{i-1} \left(\frac{k[\Lambda(t)]^{k-1}\Lambda'(t)}{k!} e^{-\Lambda(t)} + \frac{[\Lambda(t)]^k}{k!} e^{-\Lambda(t)}(-\Lambda'(t))\right) \\
&= \Lambda'(t) e^{-\Lambda(t)} \left(\frac{[\Lambda(t)]^{i-1}}{(i-1)!}\right) = \lambda(t) e^{-\Lambda(t)} \left(\frac{[\Lambda(t)]^{i-1}}{(i-1)!}\right).
\end{aligned}$$

If F_{Λ_i} denotes the distribution function of $\Lambda\left(T_i\right)$, then

$$F_{\Lambda_i}(t) = P\left(\Lambda\left(T_i\right) \leq t\right) = P\left(T_i \leq \Lambda^{-1}(t)\right) = 1 - \sum_{k=0}^{i-1} \frac{t^k}{k!} e^{-t}.$$

On the other hand,

$$F_{\Lambda_i}(t) = P\left(\widetilde{N}(t) \geq i\right) = \sum_{k=i}^{\infty} \frac{t^k}{k!} e^{-t} = \frac{1}{(i-1)!} \int_0^t u^{i-1} e^{-u} du.$$

It then follows that $\Lambda(T_i) = \sum_{k=1}^{i} W_k$, $i = 1, 2, \ldots$, has the gamma distribution $\mathcal{G}(i, 1)$; $W_k = \Lambda(T_k) - \Lambda(T_{k-1})$, $k = 1, 2, \ldots$, are exponentially distributed $\mathcal{E}(1)$ and that the transformed counting process $\{N(\Lambda(t)), t \geq 0\}$ is the Poisson process with intensity 1, i.e., it is the HPP(1) as a special renewal process RP($1-\exp(-t)$).

Special classes of NHPP's which play important role in modeling reliability systems will be presented in the next two subsections.

Simulation of a non-homogeneous Poisson process

A. Random sampling or thinning approach. The procedura follows the following steps: 1) choose a λ such that $\lambda(t) \leq T$ for all $t \leq T$; 2) generate events according to a HPP(λ); 3) accept the event, say at time t, with probability $\frac{\lambda(t)}{\lambda}$, independently of what has happened before. The process of the counted events then forms the NHPP($\lambda(t)$), $\lambda(t), t \leq T$.

The algorithm is:

1: $t = 0, n = 0$;

2: generate a random number $U \sim \mathcal{U}(0, 1)$;

3: set $X = -\frac{1}{\lambda} \ln U$;

4: $t = t + X$; If $t > T$ then stop;

5: generate a random number $U \sim \mathcal{U}(0, 1)$;

6: if $U \leq \lambda(t)/\lambda$, set $n = n + 1$, $T_n = t$;

7: go to step 2.

In the output, n is the number of events in the time interval $(0, T)$, and $T_1, \ldots, T_n$ are the successive n event times of the NHPP($\lambda(t)$) in this time interval.

B. The method of direct generation of successive event times T_i, $i = 1, 2, \ldots$, relies on the Fact 3.2.1, i.e., on the formula $\Lambda(T_i) - \Lambda(T_{i-1}) \sim \mathcal{E}(1) \sim -\ln U_i$. This corresponds to the formula: $T_i = \Lambda^{-1}\Big(\Lambda(T_i) - \ln U_i\Big)$.

3.2.4 Power law process (PLP)

Definition 3.2.6 The NHPP($\lambda(\cdot)$) for which $\lambda(t) = \lambda(t;\alpha,\beta) = \alpha\beta t^{\beta-1}, \alpha > 0, \beta > 0$, is called the **Power Law Process**. □

This process will be denoted by PLP(α,β). This NHPP is also known as **Weibull NHPP**. The cumulative intensity function for this process is $\Lambda(t;\alpha,\beta) = \alpha t^{\beta}$.

Figures 3.2, 3.3 and 3.4 show the plots of the intensity functions $\lambda(t;\alpha,\beta)$, the cumulative intensity functions $\Lambda(t;\alpha,\beta)$ and the plots of trajectories of the PLP(α,β) for three chosen pairs of α and β.

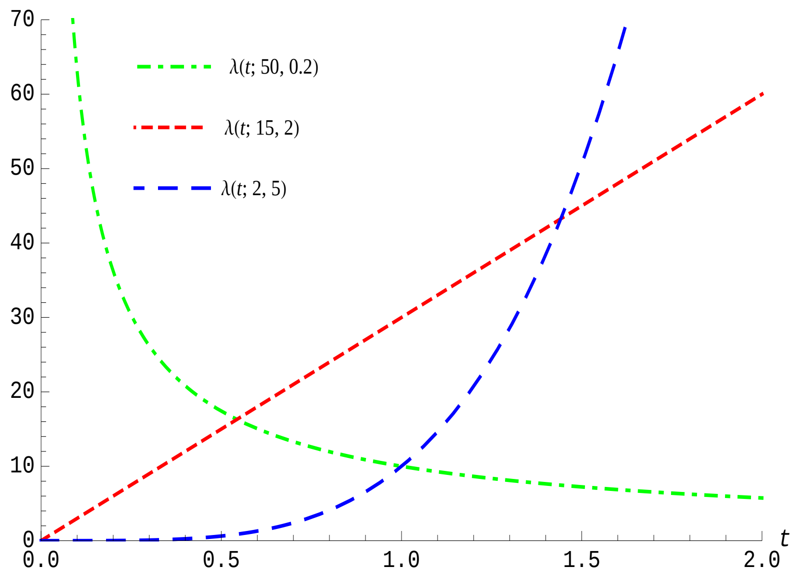

Figure 3.2: Intensity functions $\lambda(t;\alpha,\beta)$ of the PLP(α,β) for three chosen pairs of α and β.

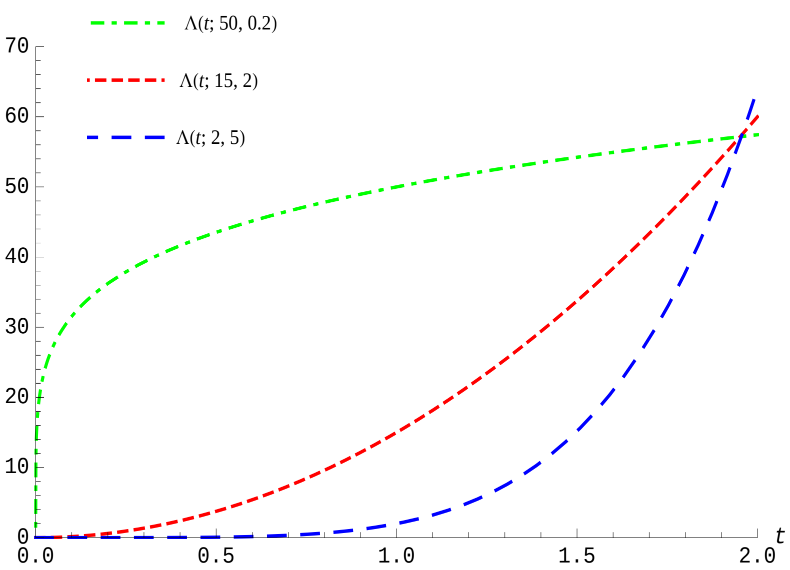

Figure 3.3: Cumulative intensity functions $\Lambda(t;\alpha,\beta)$ of the PLP(α,β) for three chosen pairs of α and β.

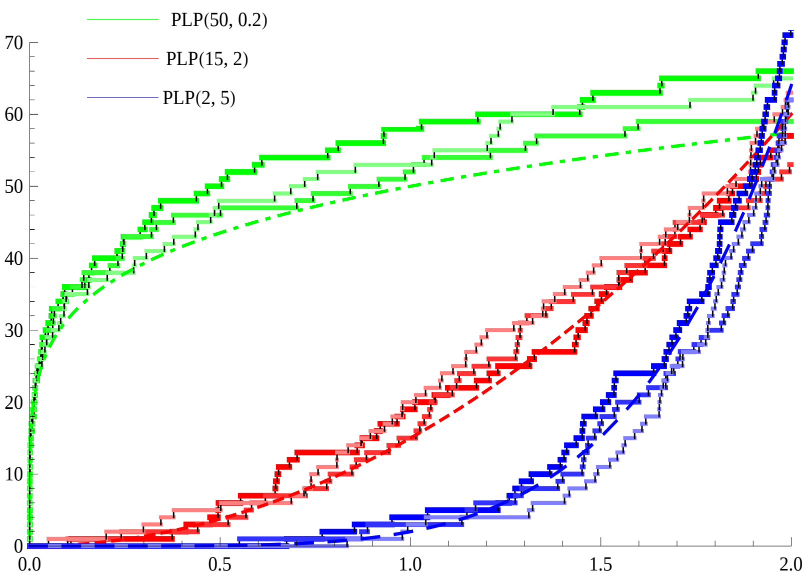

Figure 3.4: Sample paths of the PLP(α,β) for three chosen pairs of α and β.

Figure 3.4 shows three sample paths of the PLP(α, β) for each of the three chosen combinations $(50, 0.2), (15, 2)$ and $(2, 5)$ of the pair (α, β). The experimental average numbers of jumps evaluated from 100 generated repetitions of the PLP for the three chosen pairs of α and β in the simulation are equal to 48, 61, 52, respectively (compare with the values of $\widehat{n}$ in Table 4.1 for $T = 2$).

Simulation of a power law process

Generating the PLP(α, β) process until a given number of jumps (failures) is based on the following formula:

$$T_i = \left(T_{i-1}^{\beta} + \frac{1}{\alpha} \ln \frac{1}{1 - U_i} \right)^{1/\beta}, \quad i = 1, 2, \ldots, \tag{3.3}$$

$T_0 = 0$, where U_i are random numbers from uniform distribution $\mathcal{U}(0, 1)$. The generating formula is equivalent to

$$T_i = \left(T_{i-1}^{\beta} - \frac{1}{\alpha} \ln U_i \right)^{1/\beta}, \quad i = 1, 2, \ldots,$$

but for numerical computation reasons the previous formula is more useful.

3.2.5 Non-homogeneous Poisson processes with bounded mean value function

Let $\{N(t), t \geq 0\}$ be a NHPP with intensity function $\lambda(t)$ and the mean value function $\Lambda(t) = \int_0^t \lambda(u) du$ (cumulative intensity) having the following parametric form

$$\Lambda(t; \alpha, \beta) = \alpha F(t/\beta), \tag{3.4}$$

where $\alpha, \beta > 0$ are unknown parameters, and $F(\cdot)$ is a known continuous, differentiable distribution function. The corresponding intensity function is given by

$$\lambda(t; \alpha, \beta) = \frac{\alpha}{\beta} f(t/\beta), \tag{3.5}$$

where $f(\cdot)$ is the differential coefficient of $F(\cdot)$. It is obvious that the model defined by (3.4) has bounded mean value function. This is the reason for which this model is used as a software reliability model, because a software system contains only a finite number of faults. The parameter α is called the expected number of faults to be eventually detected (see Yamada and Osaki (1985)) and β is a scale parameter.

3.2.6 Trend-renewal process (TRP)

The TRP was introduced and investigated first by Lindqvist (1993) and by Lindqvist et al. (1994) (see also Lindqvist and Doksum (2003)).

Let $\lambda(t)$ be a nonnegative function defined for $t \geq 0$, and let $\Lambda(t) = \int_0^t \lambda(u)du$.

Definition 3.2.7 The process $\{N(t), t \geq 0\}$ is called a **trend-renewal process** TRP$(F, \lambda(\cdot))$ if the time-transformed process $\Lambda(T_1), \Lambda(T_2), \ldots$ is an RP(F), i.e. if the random variables $\Lambda(T_i) - \Lambda(T_{i-1}), i = 1, 2, \ldots$ are i.i.d. with cdf F. □

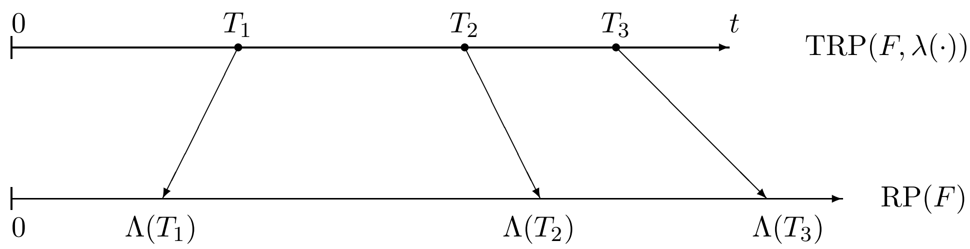

Figure 3.5: The idea of the TRP model.

The cdf F is meant as the **renewal distribution function**, and $\lambda(t)$ is called the **trend function**.

Let us note that for $F(t) = 1 - \exp(-t)$ the TRP$(1 - \exp(-t), \lambda(\cdot))$ becomes the NHPP$(\lambda(\cdot))$. Let us also remark that in particular, the TRP$(F, 1)$ is the RP(F).

The class of TRP's is defined by properties of the sequence $\{T_i, i = 1, 2, \ldots\}$ and includes the NHPP's and RP's. Equivalently, the corresponding counting process $\{N(t), t \geq 0\}$ can be considered, where $N(t) = \widetilde{N}(\Lambda(t))$ and $\{\widetilde{N}(t), t \geq 0\}$ represents an RP.

Note that the representation TRP$(F, \lambda(\cdot))$ is not unique. For uniqueness we assume that the expected value of the renewal distribution defined by F equals 1.

The form of conditional intensity function of the TRP$(F, \lambda(\cdot))$ one obtains from formula (3.1). Namely, we have

$$\begin{aligned}\gamma(t) &= \lim_{\Delta t\to 0}\frac{P(\text{failure in TRP in } (t, t+\Delta t)|\mathcal{F}_{t-})}{\Delta t}\\ &= \lim_{\Delta t\to 0}\frac{P(\text{failure in RP}(F) \text{ in } (\Lambda(t), \Lambda(t+\Delta t))|\mathcal{F}_{t-})}{\Delta t}\\ &= \lim_{\Delta t\to 0}\frac{P(\text{failure in RP}(F) \text{ in } (\Lambda(t), \Lambda(t+\Delta t))|\mathcal{F}_{t-})}{\Delta\Lambda(t)}\frac{\Delta\Lambda(t)}{\Delta t}.\end{aligned}$$

The conditional intensity function of the RP(F) is given by $\gamma(t) = \rho(t - T_{N(t-)})$, where $\rho(t)$ is the ▷ *hazard rate* (p. 1) corresponding to F. It then follows that for the TRP($F, \lambda(\cdot)$) the conditional intensity function is

$$\begin{aligned}\gamma(t) &= \rho(\Lambda(t) - \Lambda(T_{N(t-)}))\lim_{\Delta t\to 0}\frac{\Lambda(t+\Delta t)-\Lambda(t)}{\Delta t}\\ &= \rho(\Lambda(t) - \Lambda(T_{N(t-)}))\lambda(t). \qquad (3.6)\end{aligned}$$

3.2.7 Weibull-power law TRP

Let us consider the TRP($F, \lambda(\cdot)$) with

$$\lambda(t;\alpha,\beta) = \alpha\beta t^{\beta-1}, \quad \alpha>0,\ \beta>0, \quad \Lambda(t;\alpha,\beta) = \alpha t^{\beta} \qquad (3.7)$$

and

$$F(x) = F(x;\gamma) = 1 - \exp\big[-\big(\Gamma(1+1/\gamma)x\big)^{\gamma}\big] \quad (\gamma > 0), \qquad (3.8)$$

studied by Lindqvist et al. (2003). The renewal distribution function F corresponds to the Weibull distribution $\mathcal{W}e(\gamma, 1/\Gamma(1+1/\gamma))$ with the parametrization resulting in the expectation 1.

Definition 3.2.8 The TRP($F, \lambda(\cdot)$) with $\lambda(\cdot)$ and F defined by 3.7 and 3.8 is called the **Weibull-Power Law TRP**. □

The Weibull-Power Law TRP will be denoted shortly by **WPLP**(α, β, γ).

The hazard function corresponding to F is

$$\rho(x) = \rho(x;\gamma) = (\Gamma(1+1/\gamma))^{\gamma}\,\gamma x^{\gamma-1}. \qquad (3.9)$$

In the case $\gamma = 1$ the renewal distribution function F corresponds to the exponential distribution $\mathcal{E}(1)$ and the WPLP($\alpha, \beta, 1$) becomes the NHPP($\lambda(t)$) with $\lambda(t) = \alpha\beta t^{\beta-1}$, i.e., it becomes the PLP(α, β). Note that in this case $\varphi = \alpha$.

If $\gamma = 1$ and $\beta = 1$, then the WPLP($\alpha, 1, 1$) is the TRP($1 - \exp(-t), \alpha$), i.e., it is the HPP(α).

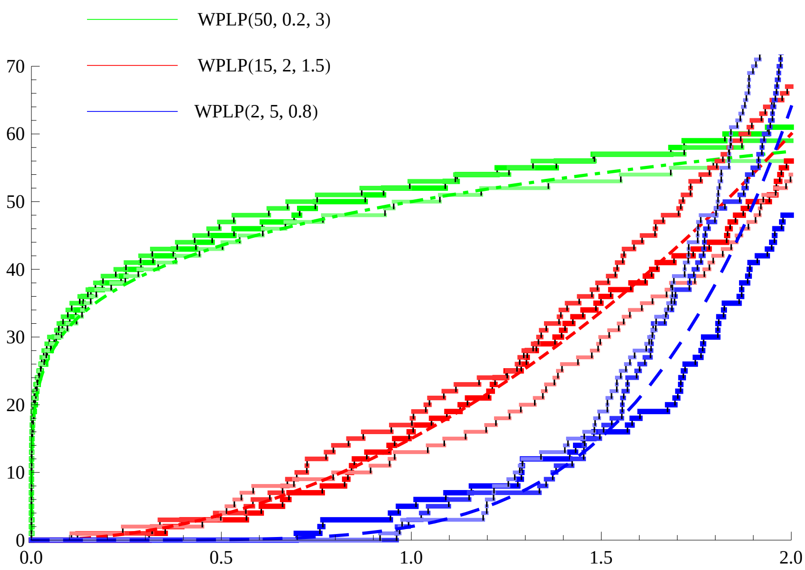

Figure 3.6: Sample paths of the WPLP(α, β, γ) for three chosen triplets of α, β and γ.

Figure 3.6 shows three sample paths of the WPLP(α, β, γ) for each of the three chosen combinations $(50, 0.2, 3), (15, 2, 1.5)$ and $(2, 5, 0.8)$ of the triplet (α, β, γ). The experimental average numbers of jumps evaluated from 100 generated repetitions of the WPLP for the three chosen triplets of α, β and γ in the simulation are equal to 51, 50, 53, respectively (compare with the values of $\widehat{n}$ in Table 4.1 for $T = 2$).

Denote by $s = s(\gamma)$ the variance of the renewal distribution defined by F. For the WPLP(α, β, γ) we then have

$$s == s(\gamma) = \frac{\Gamma(1 + 2/\gamma)}{\Gamma^2(1 + 1/\gamma)} - 1. \tag{3.10}$$

In particular, $s = 1$ for the PLP(α, β).

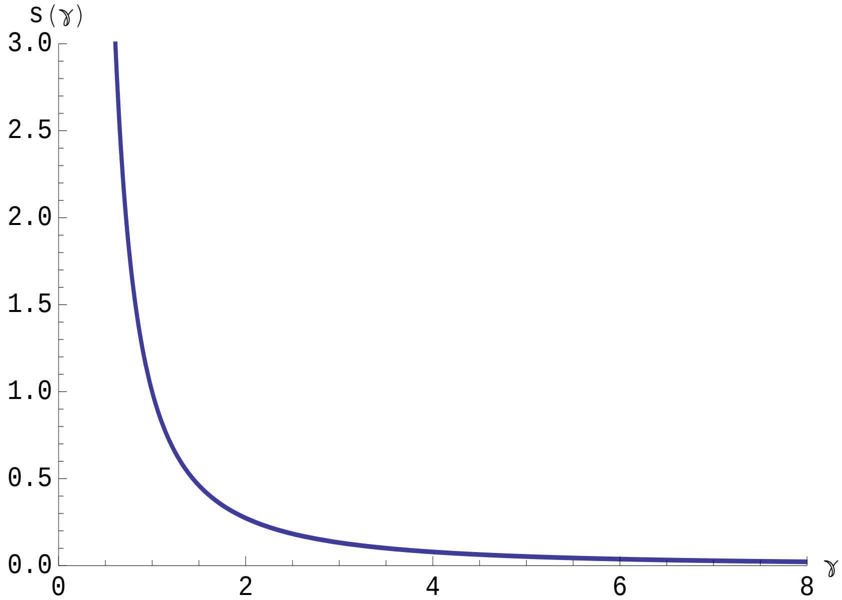

Figure 3.7: Variance function of the renewal distribution in the WPLP.

Figures 3.8 – 3.10 illustrate how variability of a WPLP is reflected by the variance of the renewal distribution. The variability of the process becomes evidently greater as the variance s of the renewal distribution defined by F decreases. The variance function $s(\gamma)$ is strictly decreasing for $\gamma > 0$ and increases rapidly for $\gamma < 1$ (see formula (3.10) and Figure 3.7).

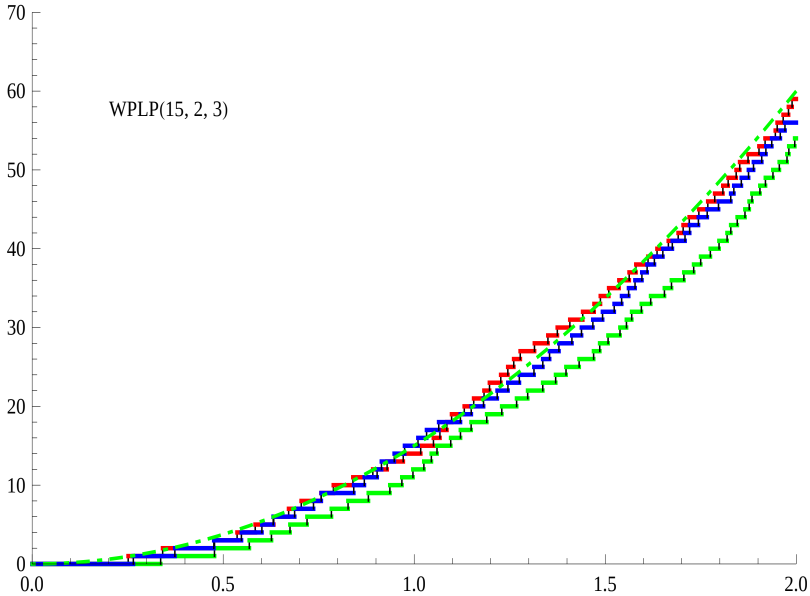

Figure 3.8: Sample paths of the WPLP$(15, 2, 3)$.

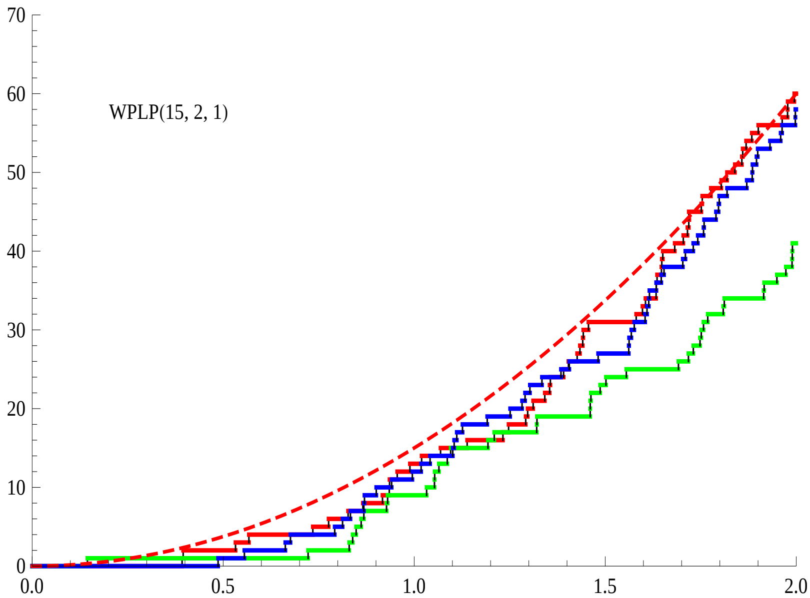

Figure 3.9: Sample paths of the WPLP$(15, 2, 1)$.

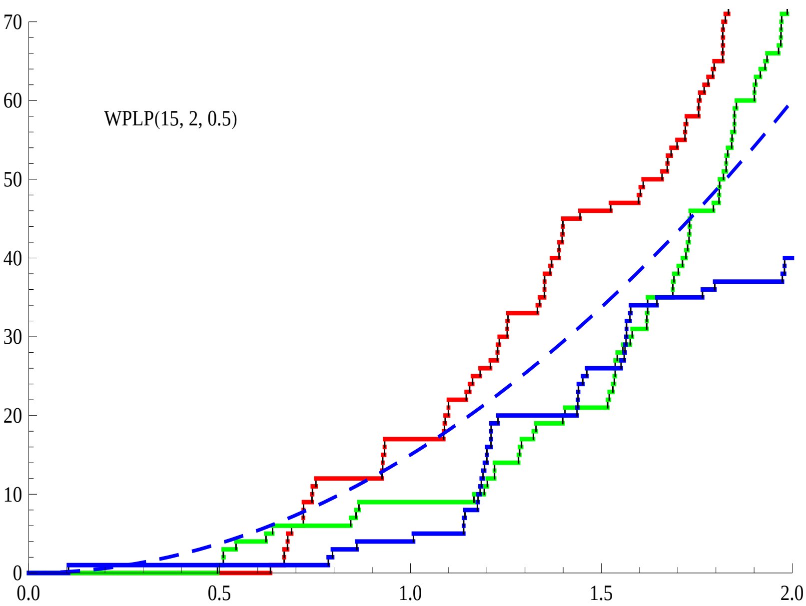

Figure 3.10: Sample paths of the WPLP$(15, 2, 0.5)$.

It follows from formula (3.9) that the hazard function $\rho(x;\gamma)$ corresponding to F of the WPLP is decreasing in x for $\gamma < 1$ (belonging then to DFR class) and increasing in x for $\gamma > 1$ (belonging then to IFR class). For $\gamma = 2$ it is the linear function $\rho(x;2) = 1.5708x$, and for $\gamma = 1$ it is a constant function which equals 1 and corresponds to the exponential distribution $\mathcal{E}(1)$.

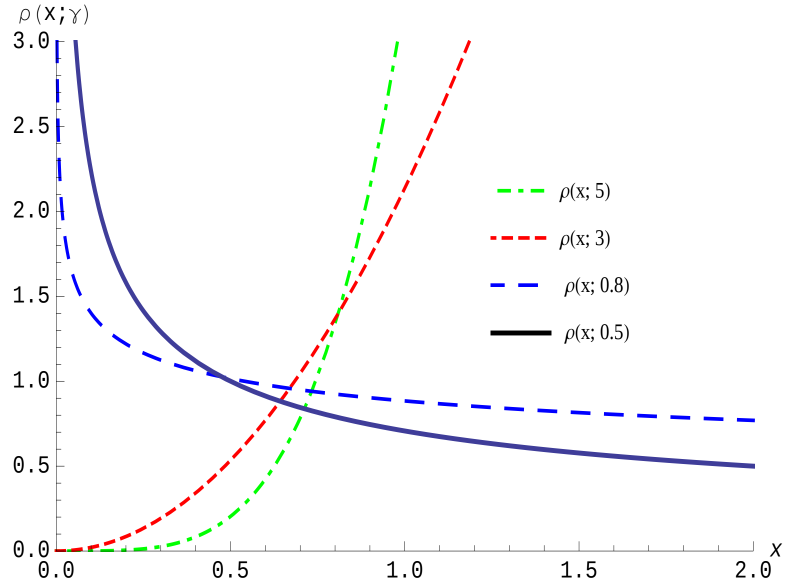

Figure 3.11: Hazard functions of the renewal distribution in the WPLP.

Simulation of a Weibull-power law TRP

The successive jump times $T_1, T_2, \ldots$ of the WPLP(α, β, γ) are generated according to the following formula:

$$T_i = \left[T_{i-1}^{\beta} + \frac{1}{\alpha\Gamma(1+1/\gamma)} \left(\ln \frac{1}{1-U_i} \right)^{1/\gamma} \right]^{1/\beta}, \quad i = 1, 2, \ldots, \tag{3.11}$$

$T_0 = 0$, where U_i are random numbers from uniform distribution $\mathcal{U}(0,1)$. The generating formula is equivalent to

$$T_i = \left[T_{i-1}^{\beta} - \frac{1}{\alpha\Gamma(1+1/\gamma)} (\ln U_i)^{1/\gamma} \right]^{1/\beta}, \quad i = 1, 2, \ldots,$$

but for numerical computation reasons formula (3.11) is more useful.

Chapter 4

Parameter Estimation in Non-homogeneous Poisson Process with Power Law Intensity Function

4.1 Likelihood function for stochastic processes

Let $\{Y(s), s \geq 0\}$ be a stochastic process observed in the time interval $[0, t]$. Suppose that for each t the distribution P^t of the process $\{Y(s), s \geq 0\}$ observed up to time t is an element of a family of distributions $\{P^t_\vartheta, \vartheta \in \Theta\}$. Assume that for some fixed $\vartheta_0 \in \Theta$, all the measures $P^t_\vartheta, \vartheta \in \Theta$, are absolutely continuous with respect to the measure $P^t_{\vartheta_0}$. The corresponding density (the Radon-Nikodym derivative)

$$L(t; \vartheta) = L(t, y(\cdot); \vartheta, \vartheta_0) = \frac{dP^t_\vartheta}{dP^t_{\vartheta_0}}(y(\cdot)) \tag{4.1}$$

regarded as a function of ϑ given the data (realization) $y(\cdot)$ is called the **likelihood function** of the process $\{Y(s), s \geq 0\}$ observed in the time interval $[0, t]$.

In the discrete time process $\{Y(s), s = 0, 1, 2, \ldots\}$ the likelihood function always exists if the distributions P^t_ϑ, $P^t_{\vartheta_0}$ have positive densities, and it coincides with the ratio of these two densities.

The likelihood function can be often calculated using the limit relation

$$L(t; \vartheta) = \lim_{n \to \infty} \frac{p_\vartheta\left(y(t_1), \ldots, y(t_n)\right)}{p_{\vartheta_0}\left(y(t_1), \ldots, y(t_n)\right)},$$

where p_ϑ denotes the density of the vector $(y(t_1), \ldots, y(t_n))$, while $\{t_1, \ldots, t_n\}$ is a dense set in $[0, t]$.

Note that by the fundamental identity of sequential analysis for stochastic processes, which is a consequence of the optional stopping theorem, the form of the likelihood function of (4.1) remains the same, where the end time t is replaced with a random finite stopping time τ with respect to the family $\mathcal{F}_t = \sigma\{Y(s), s \leq t\}, t \geq 0,$.

If $\{Y_0, Y_1, Y_2, \ldots\}$ forms a homogeneous Markov chain, then under quite general assumptions the likelihood function is

$$\frac{dP^t_\vartheta}{dP^t_{\vartheta_0}}(y(\cdot)) = p_0(Y_0; \vartheta) p(Y_1|Y_0; \vartheta) \cdots p(Y_t|Y_{t-1}; \vartheta),$$

where p_0 and p are the initial and transition densities of the distribution.

Let $\{Y(s), s \geq 0\}$ be a homogeneous Markov process with a discrete state space and differentiable transition probabilities $p_{i,j}(s)$. The transition probability matrix is determined by the matrix $\mathbf{Q} = \left(q_{i,j}\right)$, where $q_{i,j} = p'_{i,j}(0)$. Denote $q_i = -q_{i,i}$ and let $y_0 = i_0$ be independent of $\mathbf{Q}$ at the initial time. By choosing any matrix $\mathbf{Q}_0 = \left(q^0_{i,j}\right)$, one finds

$$\begin{aligned}\frac{dP^t_{\mathbf{Q}}}{dP^t_{\mathbf{Q}_0}}(y(\cdot)) = \exp\left[\left(q^0_{y(t_{N(t)})} - q_{y(t_{N(t)})}\right)t\right] \sum_{k=0}^{N(t)-1} \frac{q_{y(t_k), y(t_{k+1})}}{q^0_{y(t_k), y(t_{k+1})}} \\ \cdot \exp\left[x_k\left(q_{y(t_{N(t)})} - q_{y(t_k)} - q^0_{y(t_k)} + q^0_{y(t_{N(t)})}\right)\right],\end{aligned}$$

where $N(t)$ is the number of jumps of the process $\{Y(s), s \geq 0\}$ in the interval $[0, t]$; t_k is the k-th jump time; $x_k = t_{k+1} - t_k$ is the inter-arrival time. The ML estimators of the parameters $q_{i,j}$ are $\widehat{q_{i,j}} = \dfrac{K_{i,j}(t)}{S_i}$, where $K_{i,j}(t)$ is the number of transitions from state i to state j in the time interval $[0, t]$, and $S_i(t)$ is the time spent in state i before t.

Let $\{Y(s), s \geq 0\}$ be the linear birth-and-death process with birth rate $\lambda > 0$ and death rate $\mu > 0$. The counting process $N(t)$ in this model is the number of births and deaths in the time interval $[0, t]$. For the process Y we have $q_{i,i+1} = i\lambda$, $q_{i,i-1} = i\mu$, $q_{i,i} = 1 - i(\lambda + \mu)$ and $q_{i,j} = 0$ if $|i - j| > 1$. Assume that $Y(0) = 1$.

The likelihood function is

$$\begin{aligned} L((\lambda,\mu);t) &= \frac{dP^t_{(\lambda,\mu)}}{dP^t_{(\lambda_0,\mu_0)}}(y(\cdot)) \\ &= \left(\frac{\lambda}{\lambda_0}\right)^{B(t)}\left(\frac{\mu}{\mu_0}\right)^{D(t)} \exp\left[-(\lambda+\mu-\lambda_0-\mu_0)S(t)\right] \\ &= \exp\left[\vartheta_1 B(t)+\vartheta_2 D(t)-\left(\mathrm{e}^{\vartheta_1}+\mathrm{e}^{\vartheta_2}\right)S(t)\right] =: L(\vartheta;t), \end{aligned}$$

where $(\vartheta = (\vartheta_1,\vartheta_2) = (\log\lambda, \log\mu)$, $B(t)$ is the number of births (jumps of measure +1) in $[0,t]$, $D(t)$ is the number of deaths (jumps of measure -1) in $[0,t]$, and

$$S(t) = \int_0^t y(s)ds$$

is the total time lived in the population before time t. The ML estimators of λ and μ are

$$\widehat{\lambda}(t) = \frac{B(t)}{S(t)}, \quad \widehat{\mu}(t) = \frac{D(t)}{S(t)}$$

If $\{N(s), s \geq 0\}$ is the HPP(λ), then the likelihood function is

$$L(\lambda;t) = \lambda^{N(t)}\exp(-\lambda t) = \exp\left[\vartheta N(t) - \mathrm{e}^{\vartheta}t\right] =: L(\vartheta;t),$$

where $\vartheta = \log\lambda$. The ML estimator of λ is $\widehat{\lambda} = \dfrac{N(t)}{t}$.

4.2 Maximum likelihood estimators in a PLP

Let $\{N(t), t \geq 0\}$ be a NHPP with a mean value function $\Lambda(t;\vartheta)$ and intensity function $\lambda(t;\vartheta)$, where ϑ is an unknown vector parameter. For this process observed on the time interval $[0,\tau]$, the likelihood function based on the observed arrival times $t_1, t_2, \ldots, t_{N(\tau)}$ and $N(\tau)$ is of the form

$$L(\tau;\vartheta) = L(t_1, t_2, \ldots, t_{N(\tau)}, N(\tau);\vartheta) = \left[\prod_{i=1}^{N(\tau)} \lambda(t_i;\vartheta)\right] \exp\left[-\Lambda(\tau;\vartheta)\right], \quad (4.2)$$

dropping the factor which does not depend on an unknown parameter. The form of the likelihood function follows from the formula given by Andersen et al.

(1993) for the likelihood function of a point process (see also Thompson (1988)). That likelihood function of a point process is used in the next section. Formula (4.2) one obtains from formula (5.1) by assuming that the conditional intensity $\gamma(t) = \lambda(t)$.

In particular, for the PLP(α, β) the **likelihood function** defined by (4.2) takes the form

$$L(\tau; \alpha, \beta) = (\alpha\beta)^{N(\tau)} \prod_{i=1}^{N(\tau)} t_i^{\beta-1} \exp\left(-\alpha\tau^{\beta}\right)$$

For the PLP(α, β) the ML estimators of α and β can be explicitly determined (see e.g. Rigdon and Basu (2000), pp. 136–137).

The **log-likelihood function** for the PLP(α, β) is

$$\begin{aligned} \ell(\tau; \vartheta) &:= \log L(\tau; \vartheta) = \log L(\tau; \alpha, \beta) \\ &= N(\tau)(\ln\alpha + \ln\beta) + (\beta - 1)\sum_{i=1}^{N(\tau)} \ln t_i - \alpha\tau^{\beta}. \end{aligned}$$

Solving the **likelihood equations**

$$\left\{ \begin{aligned} \frac{\partial\ell(\tau; \alpha, \beta)}{\partial\alpha} &= 0, \\ \frac{\partial\ell(\tau; \alpha, \beta)}{\partial\beta} &= 0. \end{aligned} \right. \tag{4.3}$$

we obtain the form of the ML estimators given in the following fact:

Fact 4.2.1 For the PLP(α, β) the **ML estimators** $\widehat{\alpha}_{ML}$ and $\widehat{\beta}_{ML}$ of α and β, based on the observation up to any stopping time τ, are of the form

$$\widehat{\alpha}_{ML} = \frac{N(\tau)}{\tau^{\widehat{\beta}_{ML}}} \tag{4.4}$$

and

$$\widehat{\beta}_{ML} = N(\tau)\left(\ln\frac{\tau^{N(\tau)}}{\prod_{i=1}^{N(\tau)} T_i}\right)^{-1} = N(\tau)\left(\sum_{i=1}^{N(\tau)} \ln\frac{\tau}{T_i}\right)^{-1}. \tag{4.5}$$

□

In the case when the observation is finished at the n-th failure time point, i.e., $\tau = T_{N(\tau)}$ and $N(\tau) = n$, in other words if the observation time is the random stopping time $\tau = \inf\{t \geq 0; N(t) = n\}$, then we say that the stopping time τ

determines the so called **inverse estimation plan** . This way of terminating the observation of the process, as soon as a predetermined number of failures has been observed, is called **failure truncation**.

In the case when the observation is finished at a predetermined time, say T, i.e., $\tau = const = T$, then we say that the stopping time τ determines the so called **simple estimation plan** . This way of terminating the observation of the process is called **time truncation**.

Fact 4.2.2 In the inverse estimation plan for the PLP(α, β) the **ML estimators** of α and β are defined by

$$\widehat{\alpha}^I_{ML} = \frac{n}{T_n^{\widehat{\beta}^I_{ML}}}, \quad \text{where} \quad \widehat{\beta}^I_{ML} = n\Big(\ln \frac{T_n^n}{\prod_{i=1}^n T_i}\Big)^{-1} = n\Big(\sum_{i=1}^{n-1} \ln \frac{T_n}{T_i}\Big)^{-1}. \tag{4.6}$$

In the simple estimation plan for the PLP(α, β) the **ML estimators** of α and β are defined by

$$\widehat{\alpha}^S_{ML} = \frac{N(T)}{T^{\widehat{\beta}^S_{ML}}}, \tag{4.7}$$

where

$$\widehat{\beta}^S_{ML} = N(T)\Big(\ln \frac{T^{N(T)}}{\prod_{i=1}^{N(T)} T_i}\Big)^{-1} = N(T)\Big(\sum_{i=1}^{N(T)} \ln \frac{T}{T_i}\Big)^{-1}. \tag{4.8}$$

□

For the HPP(α) the **ML estimator** of α in the inverse estimation plan is

$$\widehat{\alpha}^I_{ML} = \frac{n}{T_n}, \tag{4.9}$$

and in the simple estimation plan is

$$\widehat{\alpha}^S_{ML} = \frac{N(T)}{T}. \tag{4.10}$$

Let us note that in the simple estimation plan for a PLP the ML estimators depend on $N(T)$, the number of failures up to time T, and on the failure times $T_i, i = 1, \ldots, N(T)$; $0 < T_1 < T_2 < \cdots < T_{N(T)} \leq T$. Remark that the number $N(T)$ as well as the failure times T_i are random, whereas in the inverse estimation plan the ML estimators are determined only by $T_1, \ldots, T_n$, where n is the number of failures given at advance.

4.3 Expected number of failures

The expected number of failures in the PLP(α,β) up to time T can be estimated by the formula $\widehat{E(N(T))} = \widehat{\alpha}T^{\widehat{\beta}}$, where $\widehat{\alpha}$ and $\widehat{\beta}$ are estimates of the parameters α and β.

Table 4.1 contains estimates of α for given values of β and the expected number of failures $\widehat{n}$ at the termination time T for the PLP(α,β). The estimates $\widehat{\alpha}$ of the parameter α are evaluated according to the formula $\widehat{n} = \Lambda(T;\widehat{\alpha},\beta) = \widehat{\alpha}T^{\beta}$.

$\widehat{n} = 50$									
T/β	0.2	0.5	0.8	1.	1.5	2.	3.	4.	5.
1.5	46.105	40.825	36.149	33.333	27.217	22.222	14.815	9.877	6.584
2	43.528	35.355	28.717	25.000	17.678	12.500	6.250	3.125	1.562
3	40.137	28.868	20.762	16.667	9.623	5.556	1.852	0.617	0.206
4	37.893	25.000	16.494	12.500	6.250	3.125	0.781	0.195	0.049
5	36.239	22.361	13.797	10.000	4.472	2.000	0.400	0.080	0.016
10	31.548	15.811	7.924	5.000	1.581	0.500	0.050	0.005	0.0005
$\widehat{n} = 100$									
T/β	0.2	0.5	0.8	1.	1.5	2.	3.	4.	5.
1.5	92.211	81.650	72.298	66.667	54.433	44.444	29.630	19.753	13.169
2	87.055	70.711	57.435	50.000	35.355	25.000	12.500	6.250	3.125
3	80.274	57.735	41.524	33.333	19.245	11.111	3.704	1.235	0.412
4	75.786	50.000	32.988	25.000	12.500	6.250	1.562	0.391	0.098
5	72.478	44.721	27.595	20.000	8.944	4.000	0.800	0.160	0.032
10	63.096	31.623	15.849	10.000	3.162	1.000	0.100	0.010	0.001
$\widehat{n} = 200$									
T/β	0.2	0.5	0.8	1.	1.5	2.	3.	4.	5.
1.5	184.422	163.299	144.596	133.333	108.866	88.889	59.259	39.506	26.337
2	174.110	141.421	114.870	100.000	70.711	50.000	25.000	12.500	6.250
3	160.548	115.470	83.049	66.667	38.490	22.222	7.407	2.469	0.823
4	151.572	100.000	65.975	50.000	25.000	12.500	3.125	0.781	0.195
5	144.956	89.443	55.189	40.000	17.889	8.000	1.600	0.320	0.064
10	126.191	63.246	31.698	20.000	6.325	2.000	0.200	0.020	0.002

Table 4.1: Estimates of α for given values of β and the expected number of failures $\widehat{n}$ at the termination time T for the PLP(α,β).

Chapter 5

Estimation of Parameters for Trend-renewal Processes

5.1 Introduction

In this chapter, methods of estimating unknown parameters of a trend function for trend-renewal processes (TRP's) are investigated in the case when the renewal distribution function F is unknown. If the renewal distribution is unknown, then the likelihood function of the trend-renewal process is unknown and consequently the maximum likelihood method cannot be used. In such situation, three other methods of estimating the trend parameters are presented. The methods considered can also be used to predict future occurrence times. The performance of the estimators based on these methods is illustrated numerically for some TRP's for which the statistical inference is analytically intractable.

Parametric inference on the parameters of the TRP was considered in the paper of Lindqvist et al. (2003), where the authors also proposed corresponding models, called heterogeneous TRP's, that extend the TRP to cases involving unobserved heterogeneity. Nonparametric ML estimation of the trend function of a TRP$(F, \lambda(\cdot))$ under the often natural condition that $\lambda(\cdot)$ is monotone was considered by Heggland and Lindqvist (2007).

Peña and Hollander (2004) presented a general class of models that allows the researcher to incorporate the effect of interventions performed on a unit after each event occurrence, the impact of accumulating events on a unit, the effect of unobservable random effects of frailties, and the effect of covariates that could be time-dependent. The ML estimators of this general models parameters were presented, and their finite and asymptotic properties were ascertained by Stocker and Peña (2007).

From practical point of view, the problem of estimating trend parameters of the TRP with unknown renewal distribution may be of interest in the situation when we observe several systems, of the same kind, working in different environments and we are interesting in examining and comparing their trend functions, whatsoever their renewal distribution is.

In Section 5.2 the form of likelihood function for a TRP is presented. The likelihood function and the likelihood equations for estimating the parameters of the Weibull-Power Law TRP are given in Section 5.3. The likelihood equations are also presented in the form which is used in simulation study to obtain the ML estimators of the TRP parameters. In Section 5.4, three methods of estimating the trend parameters in the case when the ML methods can not be used. The estimation problem of the trend parameters in some special case of the TRP is considered in Section 5.5. In Section 5.6 the estimators proposed are examined and compared with the ML estimators (obtained under the additional assumption that the renewal distribution has a known parametric form) through a computer simulation study. Some real data are examined in Section 5.6.3. Section 5.7 contains conclusions and some prospects.

5.2 Likelihood function for a TRP

For a point process $N(t)$ observed in the interval time $[0,\tau]$ with the realizations $t_1, t_2, \ldots, t_{N(\tau)}$ of the jump (failure) times $T_1, T_2, \ldots, T_{N(\tau)}$ and conditional intensity function $\gamma(t)$, the likelihood function is of the form

$$L(\tau) = \Big[\prod_{i=1}^{N(\tau)} \gamma(t_i)\Big] \exp\left(-\int_0^\tau \gamma(u)du\right) \tag{5.1}$$

(Andersen et al. (1993)). The conditional intensity function $\gamma(t)$ is defined by (3.1).

Taking into account formula (3.6), the likelihood function of (5.1) takes the following form for the TRP$(F, \lambda(\cdot))$:

$$\begin{aligned} L(\tau) \quad = \quad & \left[\prod_{i=1}^{N(\tau)} \rho(\Lambda(t_i) - \Lambda(t_{i-1}))\lambda(t_i) \cdot \exp\left(-\int_{t_{i-1}}^{t_i} \rho(\Lambda(u) - \Lambda(t_{i-1}))\lambda(u)du\right)\right] \\ & \cdot \exp\left(-\int_{t_{N(\tau)}}^{\tau} \rho(\Lambda(u) - \Lambda(t_{N(\tau)}))\lambda(u)du\right). \end{aligned}$$

Consequently, for a TRP$(F, \lambda(\cdot))$ observed in the time interval $[0,\tau]$, by applying the substitution $v = \Lambda(u) - \Lambda(T_{i-1})$, the **likelihood function** takes the form

$$\begin{aligned} L(\tau) &= \left[\prod_{i=1}^{N(\tau)} z\big(\Lambda(t_i) - \Lambda(t_{i-1})\big)\lambda(t_i) \exp\left(- \int_0^{\Lambda(t_i)-\Lambda(t_{i-1})} \rho(v)dv \right) \right] \\ &\quad \cdot \exp\left(- \int_0^{\Lambda(\tau)-\Lambda(t_{N(\tau)})} \rho(v)dv \right) \end{aligned} \tag{5.2}$$

(Linqvist et al. (2003, formula (2, ch.2)) and the **log-likelihood function** is defined by

$$\begin{aligned} \ell(\tau) &:= \log L(\tau) \\ &= \sum_{i=1}^{N(\tau)} \left[\log\big(\rho(\Lambda(t_i) - \Lambda(t_{i-1}))\big) + \log\big(\lambda(t_i)\big) - \int_0^{\Lambda(t_i)-\Lambda(t_{i-1})} \rho(v)dv \right] \\ &\quad - \int_0^{\Lambda(\tau)-\Lambda(t_{N(\tau)})} \rho(v)dv. \end{aligned} \tag{5.3}$$

5.3 Estimation in the Weibull-Power Law TRP

5.3.1 Likelihood function

For the WPLP(α, β, γ) the **likelihood function** defined by (5.2) takes the form

$$\begin{aligned} L(\tau) = L(\tau; \vartheta) &= \prod_{i=1}^{N(\tau)} \varphi\beta\gamma t_i^{\beta-1} (t_i^\beta - t_{i-1}^\beta)^{\gamma-1} \\ &\quad \exp\left[- \sum_{i=1}^{N(\tau)} \varphi(t_i^\beta - t_{i-1}^\beta)^\gamma - \varphi(\tau^\beta - t_{N(\tau)}^\beta)^\gamma \right], \end{aligned}$$

where $\vartheta = (\varphi, \beta, \gamma)$ and

$$\varphi = \varphi(\alpha, \gamma) = [\alpha\Gamma(1 + 1/\gamma)]^\gamma. \tag{5.4}$$

The **log-likelihood function** for the WPLP(α, β, γ) is

$$\begin{aligned} \ell(\tau; \vartheta) &:= \log L(\tau; \vartheta) \\ &= N(\tau)(\ln\varphi + \ln\beta + \ln\gamma) + (\beta - 1) \sum_{i=1}^{N(\tau)} \ln t_i + (\gamma - 1) \sum_{i=1}^{N(\tau)} \ln(t_i^\beta - t_{i-1}^\beta) \\ &\quad -\varphi \Big[\sum_{i=1}^{N(\tau)} (t_i^\beta - t_{i-1}^\beta)^\gamma + (\tau^\beta - t_{N(\tau)}^\beta)^\gamma \Big]. \end{aligned}$$

In the inverse estimation plan the **likelihood function** is given by

$$\widetilde{L}(n;\vartheta) = (\varphi\beta\gamma)^n \prod_{i=1}^{n} t_i^{\beta-1}[t_i^\beta - t_{i-1}^\beta]^{\gamma-1} \exp\left\{-\varphi\sum_{i=1}^{n}[t_i^\beta - t_{i-1}^\beta]^\gamma\right\},$$

and the **log-likelihood function** is

$$\begin{aligned}\widetilde{\ell}(n;\vartheta) &= n(\ln\varphi + \ln\beta + \ln\gamma)\\ &\quad + \sum_{i=1}^{n}\Big[(\beta-1)\ln t_i + (\gamma-1)\ln(t_i^\beta - t_{i-1}^\beta) - \varphi[t_i^\beta - t_{i-1}^\beta]^\gamma\Big].\end{aligned}$$

5.3.2 ML estimators

The solution to the equation $\partial\ell/\partial\varphi = 0$ with respect to φ is

$$\widetilde{\varphi} = \widetilde{\varphi}(\beta,\gamma) = \frac{N(\tau)}{\sum_{i=1}^{N(\tau)}(t_i^\beta - t_{i-1}^\beta)^\gamma + (\tau^\beta - t_{N(\tau)}^\beta)^\gamma}. \tag{5.5}$$

Performing the likelihood equations for the parameters β and γ we have the following fact.

Fact 5.3.1 The **ML estimators** $\widehat{\varphi}_{ML}$, $\widehat{\beta}_{ML}$ and $\widehat{\gamma}_{ML}$ of the parameters φ, β and γ, based on the observation up to any stopping time τ, are determined as follows:

$$\widehat{\varphi}_{ML} = \frac{N(\tau)}{\sum_{i=1}^{N(\tau)}\left(t_i^{\widehat{\beta}_{ML}} - t_{i-1}^{\widehat{\beta}_{ML}}\right)^{\widehat{\gamma}_{ML}} + \left(\tau^{\widehat{\beta}_{ML}} - t_{N(\tau)}^{\widehat{\beta}_{ML}}\right)^{\widehat{\gamma}_{ML}}},$$

where $\widehat{\beta}_{ML}$ and $\widehat{\gamma}_{ML}$ are the solutions of the following system of likelihood equations

$$\begin{aligned}\frac{N(\tau)}{\beta} + \sum_{i=1}^{N(\tau)}\left\{\left(t_i^\beta \ln t_i - t_{i-1}^\beta \ln t_{i-1}\right)\left[\frac{\gamma-1}{t_i^\beta - t_{i-1}^\beta} - \widetilde{\varphi}\gamma(t_i^\beta - t_{i-1}^\beta)^{\gamma-1}\right] + \ln t_i\right\}&\\ -\widetilde{\varphi}\gamma\left(\tau^\beta - t_{N(\tau)}^\beta\right)^{\gamma-1}\left(\tau^\beta\ln\tau - t_{N(\tau)}^\beta \ln t_{N(\tau)}\right) &= 0,\\ \frac{N(\tau)}{\gamma} + \sum_{i=1}^{N(\tau)}\ln(t_i^\beta - t_{i-1}^\beta)\left[1 - \widetilde{\varphi}(t_i^\beta - t_{i-1}^\beta)^\gamma\right] - \widetilde{\varphi}\left(\tau^\beta - t_{N(\tau)}^\beta\right)^\gamma \ln\left(\tau^\beta - t_{N(\tau)}^\beta\right) &= 0,\end{aligned}$$

where $\widetilde{\varphi} = \widetilde{\varphi}(\beta,\gamma)$ is defined by (5.5). □

In particular, for the WPLP$(\alpha,\beta,1)$, i.e. for the PLP(α,β), we obtain the explicit form of the ML estimators of α and β, which are given in Fact 4.2.1 (formulae (4.4) and (4.5)).

In the inverse sequential estimation plan, the solution to the equation $\partial\widetilde{\ell}/\partial\varphi = 0$ with respect to φ is

$$\widetilde{\varphi} = \widetilde{\varphi}(\beta,\gamma) = \frac{n}{\sum_{i=1}^{n}(t_i^{\beta} - t_{i-1}^{\beta})^{\gamma}}, \tag{5.6}$$

and we have the following special case of Fact 5.3.1.

Fact 5.3.2 The **ML estimators** $\widehat{\varphi}_{ML}$, $\widehat{\beta}_{ML}$ and $\widehat{\gamma}_{ML}$ of the parameters φ, β and γ in the inverse estimation plan are determined as follows:

$$\widehat{\varphi}_{ML} = \frac{n}{\sum_{i=1}^{n}[t_i^{\widehat{\beta}_{ML}} - t_{i-1}^{\widehat{\beta}_{ML}}]^{\widehat{\gamma}_{ML}}}, \tag{5.7}$$

where $\widehat{\beta}_{ML}$ and $\widehat{\gamma}_{ML}$ are the solutions of the following system of likelihood equations

$$\begin{aligned} \frac{n}{\beta} + \sum_{i=1}^{n}\left\{[t_i^{\beta}\ln t_i - t_{i-1}^{\beta}\ln t_{i-1}]\Big[\frac{\gamma-1}{t_i^{\beta}-t_{i-1}^{\beta}} - \widetilde{\varphi}\gamma(t_i^{\beta}-t_{i-1}^{\beta})^{\gamma-1}\Big] + \ln t_i\right\} &= 0, \\ \frac{n}{\gamma} + \sum_{i=1}^{n}\ln(t_i^{\beta}-t_{i-1}^{\beta})\big[1-\widetilde{\varphi}(t_i^{\beta}-t_{i-1}^{\beta})^{\gamma}\big] &= 0, \end{aligned} \tag{5.8}$$

where $\widetilde{\varphi} = \widetilde{\varphi}(\beta,\gamma)$ is defined by (5.6). □

The estimator $\widehat{\alpha}$ of α is evaluated according to the formula

$$\widehat{\alpha} = \frac{\widehat{\varphi}^{1/\widehat{\gamma}}}{\Gamma(1+1/\widehat{\gamma})}, \tag{5.9}$$

where $\widehat{\varphi}$ and $\widehat{\gamma}$ are estimators of φ and γ.

Regarding that $t_0 = 0$, to avoid indeterminate expressions $0\cdot(-\infty)$ in the numerical evaluations we express the formula for the log-likelihood function in the following form

$$\begin{aligned} \widetilde{\ell}(n;\vartheta) &= n(\ln\varphi + \ln\beta + \ln\gamma) + (\beta-1)\ln t_1 + (\gamma-1)\ln t_1^{\beta} - \varphi t_1^{\beta\gamma} \\ &\quad + \sum_{i=2}^{n}\big[(\beta-1)\ln t_i + (\gamma-1)\ln(t_i^{\beta}-t_{i-1}^{\beta}) - \varphi(t_i^{\beta}-t_{i-1}^{\beta})^{\gamma}\big]. \end{aligned} \tag{5.10}$$

The derivative $\partial\widetilde{\ell}/\partial\beta$ is

$$\begin{aligned} \frac{\partial\widetilde{\ell}}{\partial\beta} &= \frac{n}{\beta} + \gamma(1-\varphi t_1^{\beta\gamma})\ln t_1 \\ &\quad + \sum_{i=2}^{n}\left\{[t_i^{\beta}\ln t_i - t_{i-1}^{\beta}\ln t_{i-1}]\Big[\frac{\gamma-1}{t_i^{\beta}-t_{i-1}^{\beta}} - \varphi\gamma(t_i^{\beta}-t_{i-1}^{\beta})^{\gamma-1}\Big] + \ln t_i\right\} \end{aligned}$$

and in the numerical computation we use the likelihood equation

$$
\begin{aligned}
&\frac{n}{\beta} + \gamma\Big(1 - \varphi(\beta,\gamma) t_1^{\beta\gamma}\Big) \ln t_1 \\
&+ \sum_{i=2}^{n} \left\{ [t_i^{\beta} \ln t_i - t_{i-1}^{\beta} \ln t_{i-1}] \Big[\frac{\gamma - 1}{t_i^{\beta} - t_{i-1}^{\beta}} - \varphi(\beta,\gamma)\gamma (t_i^{\beta} - t_{i-1}^{\beta})^{\gamma-1} \Big] + \ln t_i \right\} = 0
\end{aligned}
$$

instead of equation (5.8).

5.4 The alternative methods of estimating trend parameters of a TRP

In the case when both the form of the renewal distribution function F and the form of the trend function $\lambda(\cdot)$ of the TRP$(F, \lambda(\cdot))$ are known one can estimate unknown parameters of this process using the maximum likelihood (ML) method. The problem is to find the estimators of unknown parameters of F and λ for which the likelihood function defined by (5.2) or the log-likelihood defined by (5.3) takes its maximum.

In the case when the form of F is unknown, we present in Sections 5.4.1, 5.4.2 and 5.4.3 the three methods for estimating unknown parameters of the trend function of a TRP$(F, \lambda(\cdot))$, where $\lambda(\cdot) = \lambda(t; \vartheta)$ and ϑ is a vector of unknown parameters. The problem of estimating trend parameters of the TRP with unknown renewal distribution may be of interest in the situation when we observe several systems, of the same kind, working in different environments and we are interesting in examining and comparing their trend functions, whatsoever their renewal distribution is. Moreover, the following limit results for the TRP hold

$$
\frac{N(t)}{\Lambda(t)} \to 1, \ a.s., \qquad \frac{V(t)}{\Lambda(t)} \to 1 \ \text{ as } \ t \to \infty,
$$

where $V(t) = E(N(t))$ (see Lindqvist et al. (2003)). For NHPP$(\lambda(t))$ the equality $V(t) = \Lambda(t)$ holds for every t. Thus we may, at least asymptotically, think of $\Lambda(t)$ as the expected number of failures until time t. Therefore, we can use $\widehat{\Lambda}(t_0) = \Lambda(t_0; \widehat{\vartheta})$ as an estimator of $V(t_0)$, for some t_0 large enough, whatsoever renewal distribution of the TRP is.

5.4.1 The least squares (LS) method

The least squares (LS) method consists in determining the value of $\hat{\vartheta}_{LS}$ that minimizes the quantity

$$S^2_{LS}(\vartheta) = \sum_{i=1}^{N(\tau)} [\Lambda(t_i;\vartheta) - \Lambda(t_{i-1};\vartheta) - 1]^2, \tag{5.11}$$

where t_i are the realizations of random variables T_i, $i = 1, \ldots, N(\tau)$, and $\Lambda(t_0) := 0$. Let us note that the transformed inter-arrival times

$$W_i = \Lambda(T_i) - \Lambda(T_{i-1})$$

are the observations from the distribution with the expected value 1 (this is assumed for the uniqueness of the representation of a TRP). Thus the LS method consists in deriving such estimate of the unknown parameter ϑ (of the trend function) which minimizes the sum of squares of deviations of the random variables W_i from the expected value 1 (i.e. minimizes the sample variance).

5.4.2 The constrained least squares (CLS) method

The constrained least squares (CLS) method consists in determining the value of ϑ that minimizes the quantity $S^2_{LS}(\vartheta)$ defined by (5.11) subject to the constraint

$$\frac{1}{N(\tau)} \sum_{i=1}^{N(\tau)} [\Lambda(t_i;\vartheta) - \Lambda(t_{i-1};\vartheta)] = 1,$$

i.e., under the condition

$$\Lambda\left(t_{N(\tau)};\vartheta\right) = N(\tau). \tag{5.12}$$

Thus in the CLS method we assume additionally that the sample mean

$$\overline{W} = \frac{1}{N(\tau)} \sum_{i=1}^{N(\tau)} W_i$$

is equal to the theoretical expected value 1 of the distribution defined by F.

5.4.3 The method of moments (M)

If the value of the variance of the renewal distribution F is known, say s, then we can state the following condition on the sample variance:

$$\frac{1}{N(\tau)-1}\sum_{i=1}^{N(\tau)}[\Lambda(t_i;\vartheta)-\Lambda(t_{i-1};\vartheta)-1]^2=s.$$

Taking into account (5.12) we have the following first two sample moment conditions:

$$\begin{cases} \Lambda\left(t_{N(\tau)};\vartheta\right) = N(\tau), \\ \sum_{i=1}^{N(\tau)}[\Lambda(t_i;\vartheta)-\Lambda(t_{i-1};\vartheta)]^2 = (s+1)N(\tau)-s. \end{cases} \tag{5.13}$$

If $\vartheta=(\vartheta_1,\vartheta_2)$, then the method of moments (M method) consists in determining any solution $\hat{\vartheta}_M$ to the system of equations (5.13).

5.4.4 Some remarks

Remark 5.4.1 The LS, CLS and M methods can be useful when we do not know the form of the cumulative distribution function F (the renewal distribution), and consequently, when we do not know the likelihood function of the TRP$(F,\lambda(\cdot))$. □

Remark 5.4.2 The LS, CLS and M methods can be used to predict the next failure time. For example, we have $\widehat{T}_{N(\tau)+1}=\widehat{\Lambda}^{-1}[\widehat{\Lambda}(T_{N(\tau)})+1]$, where $\widehat{\Lambda}(t)=\Lambda(t;\widehat{\vartheta})$. □

Remark 5.4.3 The LS, CLS and M methods can be the alternative methods of obtaining the estimators of an unknown parameter ϑ in the case when the maximum likelihood estimator does not exist. For example, in the case of k-stage Erlangian NHPP, which was first mentioned in Khoshgoftaar (1988), the maximum likelihood estimator exists and is unique if and only if some condition concerning the realizations of the process is satisfied (see Zhao and Xie (1996), Theorem 2.1 (ii)). □

Remark 5.4.4 In the NHPP models for which the number of failures is bounded there are no consistent estimators of the unknown parameters (see Nayak et al. (2008), Theorem 1). Thus, in these cases of the TRP, the estimators obtained by the ML, LS, CLS or M method are not consistent. □

5.5 Estimation of trend parameters in special models of the TRP

Consider a TRP$(F, \lambda(\cdot))$, where $\lambda(t) = \alpha\beta t^{\beta-1}$, $\alpha > 0$, $\beta > 0$. If the renewal distribution function F is not specified, we will call this process the **Power Law TRP**$(F, \lambda(\cdot))$ and denote it by **PTRP**(α, β).

5.5.1 The LS method

Using the LS method we denote

$$S^2_{LS}(\alpha,\beta) = \sum_{i=1}^{N(\tau)} [\Lambda(t_i;\alpha,\beta) - \Lambda(t_{i-1};\alpha,\beta) - 1]^2,$$

and the optimization problem considered is to find

$$(\widehat{\alpha}_{LS}, \widehat{\beta}_{LS}) = \arg \min_{(\alpha,\beta)\in\mathbf{R}_+\times\mathbf{R}_+} S^2_{LS}(\alpha,\beta).$$

For the PTRP(α, β) considered, the equality $\sum_{i=1}^{N(\tau)}(t_i^\beta - t_{i-1}^\beta) = t_{N(\tau)}^\beta$ holds, and consequently

$$S^2_{LS}(\alpha,\beta) = \alpha^2 \sum_{i=1}^{N(\tau)} (t_i^\beta - t_{i-1}^\beta)^2 - 2\alpha t_{N(\tau)}^\beta + N(\tau). \tag{5.14}$$

Substituting the value

$$\alpha = \alpha_{LS}(\beta) = \frac{t_{N(\tau)}^\beta}{\sum_{i=1}^{N(\tau)}(t_i^\beta - t_{i-1}^\beta)^2}, \tag{5.15}$$

which minimizes the trinomial $S^2_{LS}(\alpha,\beta)$, into formula (5.14) we have

$$S^2_{LS}(\alpha_{LS}(\beta),\beta) = N(\tau) - \frac{t_{N(\tau)}^{2\beta}}{\sum_{i=1}^{N(\tau)}(t_i^\beta - t_{i-1}^\beta)^2},$$

and the optimization problem reduces to the problem of finding

$$\widehat{\beta}_{LS} = \arg \min_{\beta\in\mathbf{R}_+} \widetilde{S}^2_{LS}(\beta), \tag{5.16}$$

where

$$\widetilde{S}^2_{LS}(\beta) = -\frac{t_{N(\tau)}^{2\beta}}{\sum_{i=1}^{N(\tau)}(t_i^\beta - t_{i-1}^\beta)^2}.$$

For numerical reasons (to avoid $\ln 0$ in evaluating the estimator $\widehat{\beta}_{LS}$), formula (5.14) is expressed in the form

$$S^2_{LS}(\alpha,\beta)=\alpha^2 t_1^{2\beta}+\alpha^2\sum_{i=2}^{N(\tau)}(t_i^\beta-t_{i-1}^\beta)^2-2\alpha t_{N(\tau)}^\beta+N(\tau).$$

The condition $\partial S^2_{LS}(\alpha,\beta)/\partial\beta=0$ leads to the equation

$$2\alpha\Big[\alpha t_1^{2\beta}\ln t_1-t_{N(\tau)}^\beta\ln t_{N(\tau)}+\alpha\sum_{i=2}^{N(\tau)}\left(t_i^\beta-t_{i-1}^\beta\right)\left(t_i^\beta\ln t_i-t_{i-1}^\beta\ln t_{i-1}\right)\Big]=0.$$

Taking into account formula (5.15) gives

$$t_1^{2\beta}\ln t_1-\ln t_{N(\tau)}\sum_{i=1}^{N(\tau)}\left(t_i^\beta-t_{t-i}^\beta\right)^2+\sum_{i=2}^{N(\tau)}\left(t_i^\beta-t_{i-1}^\beta\right)\left(t_i^\beta\ln t_i-t_{i-1}^\beta\ln t_{i-1}\right)=0,$$

which can be rewritten in the form

$$t_1^{2\beta}\ln\frac{t_1}{t_{N(\tau)}}+\sum_{i=2}^{N(\tau)}\left(t_i^\beta-t_{i-1}^\beta\right)\left(t_i^\beta\ln\frac{t_i}{t_{N(\tau)}}-t_{i-1}^\beta\ln\frac{t_{i-1}}{t_{N(\tau)}}\right)=0. \tag{5.17}$$

Consequently, we have

Proposition 5.5.1 The **LS estimators** $\widehat{\alpha}_{LS}$ and $\widehat{\beta}_{LS}$ of α and β are determined by

$$\widehat{\alpha}_{LS}=\frac{t_{N(\tau)}^{\widehat{\beta}_{LS}}}{\sum_{i=1}^{N(\tau)}\left(t_i^{\widehat{\beta}_{LS}}-t_{i-1}^{\widehat{\beta}_{LS}}\right)^2} \tag{5.18}$$

and the $\widehat{\beta}_{LS}$ which is the solution to equation (5.17). □

5.5.2 The CLS method

Using the CLS method we denote

$$C(\tau)=\Big\{(\alpha,\beta):\Lambda(t_{N(\tau)};\alpha,\beta)=N(\tau)\Big\}, \tag{5.19}$$

and the optimization problem considered is to find

$$(\widehat{\alpha}_{CLS},\widehat{\beta}_{CLS})=\arg\min_{(\alpha,\beta)\in C(\tau)}S^2_{LS}(\alpha,\beta).$$

For the PTRP(α, β) considered the restriction set defined by (5.19) takes the form

$$C(\tau) = \Big\{(\alpha, \beta) : \alpha t^{\beta}_{N(\tau)} = N(\tau)\Big\}.$$

Denote

$$\alpha_{CLS} = \alpha_{CLS}(\beta) = \frac{N(\tau)}{t^{\beta}_{N(\tau)}}$$

and

$$S^2_{CLS}(\beta) = S^2_{LS}(\alpha_{CLS}(\beta), \beta).$$

Thus, under the CLS criterion the optimization problem reduces to the problem of finding

$$\widehat{\beta}_{CLS} = \arg \min_{\beta \in \mathbf{R}_+} S^2_{CLS}(\beta),$$

where

$$\begin{aligned} S^2_{CLS}(\beta) &= \frac{N^2(\tau)}{t^{2\beta}_{N(\tau)}} \sum_{i=1}^{N(\tau)} (t^{\beta}_i - t^{\beta}_{i-1})^2 - 2\frac{N(\tau)}{t^{\beta}_{N(\tau)}} t^{\beta}_{N(\tau)} + N(\tau) \\ &= N(\tau) \left(\frac{N(\tau)}{t^{2\beta}_{N(\tau)}} \sum_{i=1}^{N(\tau)} (t^{\beta}_i - t^{\beta}_{i-1})^2 - 1 \right). \end{aligned}$$

Hence, the problem of finding the estimator $\widehat{\beta}_{CLS}$ is equivalent to the problem of finding

$$\widehat{\beta}_{CLS} = \arg \min_{\beta \in \mathbf{R}_+} \widetilde{S}^2_{CLS}(\beta), \tag{5.20}$$

where

$$\widetilde{S}^2_{CLS}(\beta) = \frac{1}{t^{2\beta}_{N(\tau)}} \sum_{i=1}^{N(\tau)} (t^{\beta}_i - t^{\beta}_{i-1})^2.$$

Observe that

$$\widetilde{S}^2_{CLS}(\beta) = -\left[\widetilde{S}^2_{LS}(\beta)\right]^{-1}$$

and the extrema appear at the same points as in the LS method, so $\widehat{\beta}_{CLS} = \widehat{\beta}_{LS}$.

The condition $\partial \widetilde{S}^2_{CLS}(\beta)/\partial\beta = 0$ leads to the equation

$$t_1^{2\beta} \ln \frac{t_1}{t_{N(\tau)}} + \sum_{i=2}^{N(\tau)} \left(t_i^\beta - t_{i-1}^\beta\right)\left(t_i^\beta \ln \frac{t_i}{t_{N(\tau)}} - t_{i-1}^\beta \ln \frac{t_{i-1}}{t_{N(\tau)}}\right) = 0, \tag{5.21}$$

which has the same form as that one defined by (5.17) for deriving $\widehat{\beta}_{LS}$ in the LS method.

Proposition 5.5.2 The **CLS estimators** $\widehat{\alpha}_{CLS}$ and $\widehat{\beta}_{CLS}$ of α and β are determined by

$$\widehat{\alpha}_{CLS} = \frac{N(\tau)}{t_{N(\tau)}^{\widehat{\beta}_{CLS}}} \tag{5.22}$$

and the $\widehat{\beta}_{CLS}$ which is the solution to equation (5.21). □

5.5.3 The M method

For the PTRP(α, β) considered, the system of equations of (5.13) takes the form

$$\begin{cases} \alpha t_{N(\tau)}^\beta & = & N(\tau), \\ \alpha^2 \sum_{i=1}^{N(\tau)} (t_i^\beta - t_{i-1}^\beta)^2 & = & (s+1)N(\tau) - s. \end{cases}$$

Thus we have the following

Proposition 5.5.3 The **M estimators** $\widehat{\alpha}_M$ and $\widehat{\beta}_M$ of α and β are determined by

$$\widehat{\alpha}_M = \frac{N(\tau)}{t_{N(\tau)}^{\widehat{\beta}_M}} \tag{5.23}$$

and the $\widehat{\beta}_M$ which is the solution to the equation

$$\frac{N^2(\tau)}{t_{N(\tau)}^{2\beta}} \sum_{i=1}^{N(\tau)} (t_i^\beta - t_{i-1}^\beta)^2 - (s+1)N(\tau) + s = 0. \tag{5.24}$$

□

For numerical computation reasons the following equivalent form of equation (5.24)

$$N^2(\tau)\left\{\left(\frac{t_1}{t_{N(\tau)}}\right)^{2\beta}+\sum_{i=2}^{N(\tau)}\left[\left(\frac{t_i}{t_{N(\tau)}}\right)^{\beta}-\left(\frac{t_{i-1}}{t_{N(\tau)}}\right)^{\beta}\right]^2\right\} \\ -(s+1)N(\tau)+s=0 \qquad (5.25)$$

is more useful.

Recall here that for the WPLP(α,β,γ) the variance s of the renewal distribution defined by F is given by formula (3.10). In particular, $s=1$ for the PLP(α,β).

5.6 Numerical results

In this section some numerical results are presented to illustrate the accuracy of the LS, CLS and M estimators proposed in the PTRP(α,β) model (with F unspecified) and in the WPLP(α,β,γ). The samples of the PLP(α,β) and the WPLP(α,β,γ) were generated up to a fixed number n of jumps is reached and for $k=500$ samples for each chosen combination of the parameters α,β and γ. The estimates of the unknown parameters α,β and γ are evaluated as the means of the estimates derived on the basis of individual realizations of the process considered. The variability of an estimator $\widehat{\eta}$ of an unknown parameter η was measured by the root mean squared error (RMSE) which is expressed by RMSE$(\widehat{\eta})=\sqrt{(sd(\widehat{\eta}))^2+(mean(\widehat{\eta})-\eta)^2}$, where sd stands for the standard deviation. In the tables the abbreviation se$(\widehat{\eta})$ is used for this error.

In constructing the executable computer programs, procedures of the package Mathematica 8.0 were used.

5.6.1 The estimates in a PLP

The values of the estimators of α and β were evaluated numerically using two numerical methods: constrained local optimization through solving equations (CLOSE method) and constrained global optimization (CGO method).

The CLOSE method in obtaining ML estimators relies on using the explicit formulae for $\widehat{\alpha}_{ML}$ and $\widehat{\beta}_{ML}$ given by (4.4) and (4.5), respectively, in Fact 4.2.1.

The CLOSE method in evaluating the LS estimators relies on Proposition 5.5.1, i.e. on solving numerically equation (5.17) with respect to $\beta>0$ and then substituting the solution for the estimator $\widehat{\beta}_{LS}$ into formula (5.18) for the estimator $\widehat{\alpha}_{LS}$.

The CLOSE method in evaluating the CLS estimators relies on Proposition 5.5.2, i.e. on solving numerically equation (5.21) with respect to $\beta > 0$ and then substituting the solution for the estimator $\widehat{\beta}_{CLS}$ into formula (5.22) for the estimator $\widehat{\alpha}_{CLS}$.

The M estimators were obtained by Proposition 5.5.3, i.e. by solving numerically equation (5.25) (for $s = 1$) with respect to $\beta > 0$ and then substituting the solution for the estimator $\widehat{\beta}_M$ into formula (5.23) for the estimator $\widehat{\alpha}_M$.

To investigate the numerical results for those processes for which the optimization problems can not be even partially solved explicitly (in contrast to the PLP(α, β)), an analogous numerical investigation is conducted by using CGO method. The CGO method in evaluating ML estimators relies on solving the problem $(\widehat{\alpha}_{ML}, \widehat{\beta}_{ML}) = \arg \max_{(\alpha,\beta)\in \mathbf{R}_+\times\mathbf{R}_+} L(\tau; \alpha, \beta)$ or equivalently $(\widehat{\alpha}_{ML}, \widehat{\beta}_{ML}) = \arg \max_{(\alpha,\beta)\in \mathbf{R}_+\times\mathbf{R}_+} \ell(\tau; \alpha, \beta)$ by using a constrained global optimization procedure with respect to both variables α and β.

The CGO method in evaluating LS and CLS estimators relies on solving the problems defined by (5.16) and (5.20), respectively, by using constrained global optimization procedures with respect to the variable β, and then substituting the solutions into formulas (5.18) and (5.22), respectively. The results carried out by the CGO numerical method have had the same accuracy as those carried out by the CLOSE numerical method, and the latter are not presented in the paper.

The estimates $\widehat{\alpha}_{LS}$, $\widehat{\alpha}_{CLS}$, $\widehat{\beta}_{(C)LS}$, $\widehat{\alpha}_M$ and $\widehat{\beta}_M$ proposed in the PTRP(α, β) are evaluated on the basis of the realizations (samples) of the generated PLP(α, β) and compared with the ML estimates $\widehat{\alpha}_{ML}$ and $\widehat{\beta}_{ML}$ for the latter model. The values of the estimators and their measures of variability are contained in Tables 5.1 – 5.4 for $n = 50$, $n = 100$ and for $k = 500$ samples for each pair (α, β).

5.6.2 The estimates in the Weibull-power law TRP

The ML estimates $\widehat{\beta}_{ML}$ and $\widehat{\gamma}_{ML}$ of β and γ are found by maximizing the log-likelihood function in solving the optimization problem

$$(\widehat{\beta}_{ML}, \widehat{\gamma}_{ML}) = \arg \max_{(\beta,\gamma)} \widetilde{\ell}(n; (\varphi, \beta, \gamma)),$$

by using a constrained global optimization (CGO) procedure, where $\widetilde{\ell}(n; (\varphi, \beta, \gamma))$ is given by (5.10) with $\varphi = \widetilde{\varphi}(\beta, \gamma)$ defined by (5.6). The ML estimate $\widehat{\alpha}_{ML}$ of α is evaluated using formula (5.9) with $\widehat{\varphi}$ defined by (5.7). In the optimization problem the procedure NMaximize of Mathematica package is used.

The tables provide the numerical results for the WPLP(α, β, γ) in comparison to a PTRP(α, β) (the TRP with unknown F). The CLS estimates $\widehat{\alpha}_{CLS}$ and $\widehat{\beta}_{CLS}$

of the parameters α and β are evaluated on the basis of the realizations of the generated WPLP(α, β, γ) supposing that we do know nothing about the renewal distribution function F, i.e., that we observe the PTRP(α, β). The estimates $\widehat{\alpha}_{CLS}$ and $\widehat{\beta}_{CLS}$ are evaluated using Proposition 5.5.2 and the CGO method. The estimates $\widehat{\alpha}_M$ and $\widehat{\beta}_M$ were obtained by Proposition 5.5.3 for s evaluated according to formula (3.10).

The values of the estimators and their measures of variability are contained in Tables 5.5 – 5.12. We assumed $n = 50$ and $n = 100$, and used $k = 500$ simulated realizations for every combination of the three parameters α, β and γ.

5.6.3 Some real data

Let us take into account some real data of failure times, namely the data set contained in the paper of Lindqvist et. al. (2003), given in Table 5.13. These data contain 41 failure times of a gas compressor with time censoring at time 7571 (days).

Supposing that the set of failure times of Table 5.13 forms a TRP belonging to the class of WPLP(α, β, γ), the ML estimates $\widehat{\alpha}_{ML}, \widehat{\beta}_{ML}$ and $\widehat{\gamma}_{ML}$ of α, β and γ have been evaluated and presented in Table 5.14. On the other hand, if no assumptions are made on the renewal distribution function F, the estimates $\widehat{\alpha}_{CLS}$ and $\widehat{\beta}_{CLS}$ of α and β are given as the parameters of the PTRP(α, β).

As the results of Table 5.14 show, the real data of failure times considered can be recognized as the WPLP$(0.048, 0.763, 0.842)$ or the PTRP$(0.028, 0.823)$. In both cases, the estimates of β are almost the same. In Table 5.14 the relative errors $re(\widehat{\alpha}_{CLS}) = |\widehat{\alpha}_{CLS} - \widehat{\alpha}_{ML}|/\widehat{\alpha}_{ML}$ and $re(\widehat{\beta}_{CLS}) = |\widehat{\beta}_{CLS} - \widehat{\beta}_{ML}|/\widehat{\beta}_{ML}$ are given too. For comparison, in Table 5.14 there are also given the sum of squares $SS_{CLS} := S^2_{LS}(\widehat{\alpha}_{CLS}, \widehat{\beta}_{CLS})$ and $SS_{ML} := S^2_{LS}(\widehat{\alpha}_{ML}, \widehat{\beta}_{ML})$, where $S^2_{LS}(\vartheta)$ is defined by (5.11). Note that the sum of squares SS_{CLS} is somewhat smaller then SS_{ML}.

Let us denote by $ENF_{ML}(t) = \Lambda(t; \widehat{\alpha}_{ML}, \widehat{\beta}_{ML}) = \widehat{\alpha}_{ML} t^{\widehat{\beta}_{ML}}$ the estimated number of failures up to time t evaluated on the basis of the ML estimators, and analogously by $ENF_{CLS}(t) = \Lambda(t; \widehat{\alpha}_{CLS}, \widehat{\beta}_{CLS}) = \widehat{\alpha}_{CLS} t^{\widehat{\beta}_{CLS}}$ the estimated number of failures up to time t evaluated on the basis of the CLS estimators. In Table 5.15 we compare the estimated numbers of failures with the observed number of failures $ONF(t)$ for some chosen values of t. The CLS method provides satisfactory estimates of the number of failures.

5.7 Concluding remarks

A good performance of the estimators $\widehat{\alpha}_{CLS}$ and $\widehat{\beta}_{CLS}$ is observed, which are obtained by the CLS method. This method leads to satisfactory accuracy of these estimators in the $\mathrm{TRP}(F, \lambda(\cdot))$ model considered with unspecified F in comparison to the ML estimators $\widehat{\alpha}_{ML}$ and $\widehat{\beta}_{ML}$ for this model with specified F.

The CLS method leads in average to more accurate estimators than the LS and M methods.

The LS method considerably underestimates the parameter α. In most cases considered we have $\mathrm{RMSE}(\widehat{\alpha}_{CLS}) < \mathrm{RMSE}(\widehat{\alpha}_{LS})$, and in all cases, $\mathrm{RMSE}(\widehat{\alpha}_{CLS}) < \mathrm{RMSE}(\widehat{\alpha}_{M})$.

The $\mathrm{RMSE}(\widehat{\alpha}_{M})$ is about two, or even more, times greater than the $\mathrm{RMSE}(\widehat{\alpha}_{CLS})$. Similar remark concerns the $\mathrm{RMSE}(\widehat{\beta}_{M})$ and $\mathrm{RMSE}(\widehat{\beta}_{CLS})$.

In some cases the $\mathrm{RMSE}(\widehat{\alpha}_{CLS})$ is even less than the $\mathrm{RMSE}(\widehat{\alpha}_{ML})$, and the $\mathrm{RMSE}(\widehat{\beta}_{CLS})$ is less than the $\mathrm{RMSE}(\widehat{\beta}_{ML})$.

For a given number n of failures, the RMSE's of all the estimators in the $\mathrm{WPLP}(\alpha, \beta, \gamma)$ become significantly smaller as the parameter γ increases. Remark that, according to formula (3.10), the variance s of the renewal distribution F decreases evidently as γ increases. For example, $s = 1$ for $\gamma = 1$, $s = 0.2732$ for $\gamma = 2$, $s = 0.0787$ for $\gamma = 4$. A smaller value of γ (a larger value of the variance s) causes larger variability of the estimators (recall that the RMSE determines the mean squared deviation of the estimate from the true value of the parameter – the risk). In the $\mathrm{WPLP}(\alpha, \beta, 1)$, i.e. in the $\mathrm{PLP}(\alpha, \beta)$, the variance of F is equal to 1 and constitutes a great value in reference to the same value of the expectation of the renewal distribution as well as in reference to the assumed value 1 of the sample mean of the transformed working times $W_i = \Lambda(T_i) - \Lambda(T_{i-1}), i = 1, \ldots, N(\tau)$.

A great value, such as 1, of the variance of the renewal distribution causes larger variability and instability of the RMSE's of the LS, CLS and M estimators in the case of relatively small sample sizes n. It may then happen that in some cases for $\gamma = 1$ the $\mathrm{RMSE}(\widehat{\alpha}_{LS})$, $\mathrm{RMSE}(\widehat{\alpha}_{CLS})$ and/or $\mathrm{RMSE}(\widehat{\alpha}_{M})$ increase as n increases. Further simulation study for $\gamma = 1$ and much greater than $n = 100$ numbers of failures shows that these RMSE's decrease as n increases. They decrease much more slowly than the RMSE's of the parameter β.

For $\gamma \geq 2$, the $\mathrm{RMSE}(\widehat{\alpha}_{CLS})$ decreases as n increases.

If the number of jumps n increases then all the RMSE's of the parameter β, i.e. the $\mathrm{RMSE}(\widehat{\beta}_{ML})$, $\mathrm{RMSE}(\widehat{\beta}_{CLS})$ and $\mathrm{RMSE}(\widehat{\beta}_{M})$ decrease.

If the renewal distribution function is unknown, then using the CLS method is recommended to obtain the estimators of the unknown parameters of the trend function or the expected number of failures.

The CLS method can also be used to predict the next failure time. Examination of asymptotic properties of the CLS estimators would be desirable, among others, for constructing the confidence intervals for unknown parameters or for the expected number of failures.

5.8 The tables

No.	α	β	$\overline{T}_n$	$\widehat{\alpha}_{ML}$	$\widehat{\beta}_{ML}$	$\widehat{\alpha}_{LS}$	$\widehat{\alpha}_{CLS}$	$\widehat{\beta}_{(C)LS}$	$\widehat{\alpha}_{M}$	$\widehat{\beta}_{M}$
1	20	0.8	3.20	19.7126	0.8305	11.2212	21.0444	0.7766	20.6414	0.8309
2	15	1	3.31	14.6992	1.0545	8.6936	16.3002	0.9714	15.8796	1.0459
3	5	2	3.15	4.7882	2.1119	3.0020	5.6651	1.9511	6.1284	2.0992
4	1	3	3.68	0.9896	3.1501	0.6486	1.2292	2.9107	1.9286	3.0972
5	0.5	4	3.15	0.5461	4.1386	0.3245	0.6139	3.9135	1.1481	4.0914
6	0.2	5	3.02	0.2144	5.2257	0.1276	0.2429	4.8894	0.5600	5.2572
7	5	1	9.98	4.8411	1.0507	3.1241	5.8862	0.9779	6.4392	1.0340
8	1	2	7.04	1.0137	2.0918	0.6966	1.3172	1.9453	1.8174	2.0593
9	0.5	3	4.64	0.4993	3.1460	0.3283	0.6205	2.9369	1.0099	3.1879
10	0.2	4	3.98	0.2279	4.1454	0.1310	0.2504	3.9148	0.5561	4.1892

Table 5.1: The ML estimates of α and β in the PLP(α, β) and the LS, CLS and M estimates of α and β in the PTRP(α, β). The number of jumps $n = 50$.

No.	α	β	se$(\widehat{\alpha}_{ML})$	se$(\widehat{\beta}_{ML})$	se$(\widehat{\alpha}_{LS})$	se$(\widehat{\alpha}_{CLS})$	se$(\widehat{\beta}_{(C)LS})$	se$(\widehat{\alpha}_{M})$	se$(\widehat{\beta}_{M})$
1	20	0.8	4.0017	0.1252	9.2031	4.6789	0.1547	6.7959	0.3138
2	15	1	3.3799	0.1700	6.6997	4.2260	0.1902	6.3619	0.3737
3	5	2	1.8023	0.3391	2.2577	1.9989	0.2823	4.2556	0.7833
4	1	3	0.6163	0.4766	0.4315	0.5171	0.3212	2.1844	1.2008
5	0.5	4	0.3527	0.6276	0.2203	0.2621	0.3520	1.4657	1.5090
6	0.2	5	0.1785	0.8114	0.0867	0.0975	0.3565	0.8084	2.0052
7	5	1	1.7517	0.1666	2.3094	2.6501	0.1955	4.5391	0.4048
8	1	2	0.6019	0.3200	0.4780	0.7487	0.2928	2.0740	0.7657
9	0.5	3	0.3317	0.4674	0.2288	0.3041	0.3247	1.3625	1.1364
10	0.2	4	0.1889	0.6537	0.0901	0.1183	0.3500	0.8735	1.5046

Table 5.2: The measures of variability of the ML, LS, CLS and M estimates of α and β. The number of jumps $n = 50$.

No.	α	β	$\overline{T}_n$	$\widehat{\alpha}_{ML}$	$\widehat{\beta}_{ML}$	$\widehat{\alpha}_{LS}$	$\widehat{\alpha}_{CLS}$	$\widehat{\beta}_{(C)LS}$	$\widehat{\alpha}_{M}$	$\widehat{\beta}_{M}$
1	20	0.8	7.57	19.6564	0.8168	10.8982	21.0258	0.7890	21.4116	0.8138
2	15	1	6.64	14.8118	1.0243	8.4449	16.3125	0.9797	16.6752	1.0127
3	5	2	4.46	4.8790	2.0593	2.9003	5.6046	1.9758	6.4385	2.0319
4	1	3	4.63	0.9943	3.0943	0.6258	1.2199	2.9398	1.7474	3.0440
5	0.5	4	3.75	0.5424	4.0651	0.3144	0.6066	3.9344	1.1275	3.9990
6	0.2	5	3.47	0.2049	5.1247	0.1228	0.2384	4.9218	0.5083	5.0422
7	5	1	20.00	4.9431	1.0223	2.9833	5.7906	0.9820	6.5638	1.0126
8	1	2	9.93	1.0050	2.0583	0.6507	1.2635	1.9767	1.7931	2.0102
9	0.5	3	5.86	0.5024	3.0737	0.3193	0.6181	2.9465	0.9616	3.0427
10	0.2	4	4.73	0.2142	4.0854	0.1252	0.2430	3.9479	0.5168	4.0244

Table 5.3: The ML estimates of α and β in the PLP(α, β) and the LS, CLS and M estimates of α and β in the PTRP(α, β). The number of jumps $n = 100$.

No.	α	β	se$(\widehat{\alpha}_{ML})$	se$(\widehat{\beta}_{ML})$	se$(\widehat{\alpha}_{LS})$	se$(\widehat{\alpha}_{CLS})$	se$(\widehat{\beta}_{(C)LS})$	se$(\widehat{\alpha}_{M})$	se$(\widehat{\beta}_{M})$
1	20	0.8	3.9049	0.0862	9.5037	5.0847	0.1182	8.9502	0.2311
2	15	1	3.2128	0.1090	6.9290	4.4487	0.1401	7.6363	0.2744
3	5	2	1.6100	0.2225	2.3639	2.1384	0.2453	4.5891	0.5753
4	1	3	0.5071	0.3460	0.4576	0.5521	0.2964	1.8783	0.8647
5	0.5	4	0.2913	0.4330	0.2266	0.2662	0.3281	1.4226	1.1547
6	0.2	5	0.1251	0.5209	0.0904	0.0980	0.3408	0.6981	1.4549
7	5	1	1.5318	0.1046	2.3453	2.4555	0.1438	4.8208	0.2934
8	1	2	0.4922	0.2230	0.5025	0.7488	0.2518	1.9405	0.5515
9	0.5	3	0.2680	0.3088	0.2341	0.3074	0.2909	1.1825	0.8043
10	0.2	4	0.1360	0.4353	0.0945	0.1174	0.3188	0.7747	1.1057

Table 5.4: The measures of variability of the ML, LS, CLS and M estimates of α and β. The number of jumps $n = 100$.

No.	α	β	γ	$\overline{T}_n$	$\widehat{\alpha}_{ML}$	$\widehat{\beta}_{ML}$	$\widehat{\gamma}_{ML}$
1	15	1	1	3.32962	15.1619	1.0368	1.0885
2	5	2	1	3.17269	5.3976	1.9964	1.0771
3	1	3	1	3.66570	1.3147	2.9155	1.0592
4	0.5	4	1	3.15200	0.7100	3.8705	1.0442
5	15	1	2	3.33842	15.0668	0.9969	2.2657
6	5	2	2	3.16118	5.2531	1.9725	2.1992
7	1	3	2	3.68059	1.1986	2.9013	2.1112
8	0.5	4	2	3.15758	0.6307	3.8560	2.0721
9	15	1	4	3.33383	14.8703	0.9926	4.8756
10	5	2	4	3.15973	5.1408	1.9724	4.5728
11	1	3	4	3.68595	1.0875	2.9454	4.1946
12	0.5	4	4	3.16234	0.5643	3.9147	4.0732

Table 5.5: The ML estimates of α, β and γ in the WPLP(α, β, γ). The number of jumps $n = 50$.

No.	α	β	γ	$\widehat{\alpha}_{LS}$	$\widehat{\alpha}_{CLS}$	$\widehat{\beta}_{(C)LS}$	$\widehat{\alpha}_M$	$\widehat{\beta}_M$
1	15	1	1	8.4899	15.5415	1.0099	15.9159	1.0466
2	5	2	1	2.9514	5.4646	1.9763	6.6879	1.9894
3	1	3	1	0.6510	1.2038	2.9419	1.8761	3.0830
4	0.5	4	1	0.3055	0.5710	3.9710	1.1167	4.1872
5	15	1	2	12.0570	14.9190	1.0128	15.1474	1.0136
6	5	2	2	4.0270	4.9722	2.0237	5.2634	2.0276
7	1	3	2	0.8286	1.0247	3.0201	1.2190	3.0156
8	0.5	4	2	0.4107	0.5075	4.0335	0.6213	4.0406
9	15	1	4	13.9616	14.7163	1.0182	15.0521	1.0038
10	5	2	4	4.5967	4.8461	2.0332	5.0424	2.0149
11	1	3	4	0.8883	0.9368	3.0594	1.0349	3.0255
12	0.5	4	4	0.4392	0.4625	4.0821	0.5365	4.0183

Table 5.6: The LS, CLS and M estimates of α and β in the PTRP(α, β). The number of jumps $n = 50$.

No.	α	β	γ	se($\widehat{\alpha}_{ML}$)	se($\widehat{\beta}_{ML}$)	se($\widehat{\gamma}_{ML}$)
1	15	1	1	3.33457	0.15557	0.16199
2	5	2	1	1.78142	0.27868	0.93242
3	1	3	1	0.73484	0.39724	1.94492
4	0.5	4	1	0.45766	0.53774	2.95805
5	15	1	2	1.80166	0.07315	1.30066
6	5	2	2	0.97649	0.15192	0.34360
7	1	3	2	0.39435	0.23876	0.92567
8	0.5	4	2	0.23856	0.30699	1.94106
9	15	1	4	0.94914	0.04010	3.93850
10	5	2	4	0.50364	0.07720	2.63491
11	1	3	4	0.18483	0.12200	1.29288
12	0.5	4	4	0.11374	0.16595	0.46185

Table 5.7: The measures of variability of the ML estimates of α, β and γ in the WPLP(α, β, γ). The number of jumps $n = 50$.

No.	α	β	γ	se($\widehat{\alpha}_{LS}$)	se($\widehat{\alpha}_{CLS}$)	se($\widehat{\beta}_{(C)LS}$)	se($\widehat{\alpha}_{M}$)	se($\widehat{\beta}_{M}$)
1	15	1	1	6.94033	4.13278	0.19882	6.72476	0.39719
2	5	2	1	2.30323	1.96049	0.29032	4.59692	0.73614
3	1	3	1	0.44004	0.52805	0.32762	2.22384	1.12291
4	0.5	4	1	0.22956	0.23550	0.34672	1.45936	1.57426
5	15	1	2	3.33391	1.96074	0.08493	3.31304	0.17297
6	5	2	2	1.25120	0.98412	0.16400	2.06921	0.35251
7	1	3	2	0.29908	0.31353	0.22921	0.78828	0.53535
8	0.5	4	2	0.15346	0.15749	0.25820	0.45190	0.66774
9	15	1	4	1.36704	1.02954	0.04660	1.85616	0.09534
10	5	2	4	0.60101	0.51363	0.08390	1.07829	0.18743
11	1	3	4	0.18418	0.17018	0.13490	0.38390	0.28681
12	0.5	4	4	0.09863	0.09214	0.17502	0.22913	0.37755

Table 5.8: The measures of variability of the LS, CLS and M estimates of α and β in the PTRP(α, β). The number of jumps $n = 50$.

No.	α	β	γ	$\overline{T}_n$	$\widehat{\alpha}_{ML}$	$\widehat{\beta}_{ML}$	$\widehat{\gamma}_{ML}$
1	15	1	1	6.65959	15.2537	1.0106	1.0432
2	5	2	1	4.49654	5.3817	1.9808	1.0361
3	1	3	1	4.62962	1.2442	2.9300	1.0246
4	0.5	4	1	3.75459	0.6478	3.9004	1.0193
5	15	1	2	6.67911	15.1817	0.9957	2.1258
6	5	2	2	4.47066	5.2200	1.9805	2.0860
7	1	3	2	4.64077	1.1413	2.9350	2.0476
8	0.5	4	2	3.75763	0.5914	3.9068	2.0315
9	15	1	4	6.66444	15.0487	0.9947	4.3797
10	5	2	4	4.47187	5.1345	1.9817	4.2772
11	1	3	4	4.64240	1.0568	2.9697	4.0841
12	0.5	4	4	3.76311	0.5447	3.9437	4.0555

Table 5.9: The ML estimates of α, β and γ in the WPLP(α, β, γ). The number of jumps $n = 100$.

No.	α	β	γ	$\widehat{\alpha}_{LS}$	$\widehat{\alpha}_{CLS}$	$\widehat{\beta}_{(C)LS}$	$\widehat{\alpha}_{M}$	$\widehat{\beta}_{M}$
1	15	1	1	8.2955	15.8372	0.9962	16.6449	1.0165
2	5	2	1	2.8901	5.5234	1.9765	6.5350	2.0163
3	1	3	1	0.6340	1.2202	2.9356	1.7638	3.0461
4	0.5	4	1	0.3055	0.5885	3.9480	1.0603	4.0136
5	15	1	2	11.9492	14.9728	1.0050	15.5671	0.9975
6	5	2	2	3.9958	5.0149	2.0086	5.3982	2.0018
7	1	3	2	0.8237	1.0336	3.0002	1.2009	2.9886
8	0.5	4	2	0.4120	0.5169	4.0073	0.5902	4.0298
9	15	1	4	13.8816	14.8064	1.0082	15.0318	1.0047
10	5	2	4	4.6145	4.9207	2.0135	5.0642	2.0089
11	1	3	4	0.9025	0.9628	3.0300	1.0275	3.0194
12	0.5	4	4	0.4549	0.4850	4.0290	0.5416	4.0005

Table 5.10: The LS, CLS and M estimates of α and β in the PTRP(α, β). The number of jumps $n = 100$.

No.	α	β	γ	se($\widehat{\alpha}_{ML}$)	se($\widehat{\beta}_{ML}$)	se($\widehat{\gamma}_{ML}$)
1	15	1	1	3.14694	0.09949	0.09575
2	5	2	1	1.62619	0.19181	0.96731
3	1	3	1	0.59630	0.28980	1.97709
4	0.5	4	1	0.33198	0.37474	2.98187
5	15	1	2	1.77660	0.05361	1.14043
6	5	2	2	0.84431	0.10308	0.19245
7	1	3	2	0.29881	0.16427	0.96563
8	0.5	4	2	0.18406	0.22068	1.97425
9	15	1	4	0.90038	0.02865	3.40073
10	5	2	4	0.45525	0.05646	2.30391
11	1	3	4	0.14197	0.08341	1.13197
12	0.5	4	4	0.08742	0.11630	0.33783

Table 5.11: The measures of variability of the ML estimates of α, β and γ in the WPLP(α, β, γ). The number of jumps $n = 100$.

No.	α	β	γ	se($\widehat{\alpha}_{LS}$)	se($\widehat{\alpha}_{CLS}$)	se($\widehat{\beta}_{(C)LS}$)	se($\widehat{\alpha}_{M}$)	se($\widehat{\beta}_{M}$)
1	15	1	1	7.11759	4.51018	0.14632	7.98056	0.28053
2	5	2	1	2.37319	2.12398	0.24919	4.68815	0.58681
3	1	3	1	0.44665	0.53891	0.28675	1.99083	0.85777
4	0.5	4	1	0.22869	0.24739	0.31908	1.25473	1.13285
5	15	1	2	3.40104	1.90465	0.05921	3.88282	0.13443
6	5	2	2	1.20606	0.84533	0.11109	2.09510	0.26953
7	1	3	2	0.26997	0.26255	0.16401	0.69686	0.38839
8	0.5	4	2	0.14504	0.14744	0.21073	0.37910	0.50564
9	15	1	4	1.40465	0.94442	0.03051	2.17527	0.07461
10	5	2	4	0.55396	0.44468	0.05755	1.13060	0.15219
11	1	3	4	0.15325	0.13267	0.08847	0.34732	0.22260
12	0.5	4	4	0.07923	0.07255	0.11161	0.22158	0.31295

Table 5.12: The measures of variability of the LS, CLS and M estimates of α and β in the PTRP(α, β). The number of jumps $n = 100$.

1	4	305	330	651	856	996	1016	1155	1520	1597	1729
1758	1852	2070	2073	2093	2213	3197	3555	3558	3724	3768	4103
4124	4170	4270	4336	4416	4492	4534	4578	4762	5474	5573	5577
5715	6424	6692	6830	6999							

Table 5.13: The real data

$\widehat{\alpha}_{ML}$	$\widehat{\beta}_{ML}$	$\widehat{\gamma}_{ML}$	$\widehat{\alpha}_{CLS}$	$\widehat{\beta}_{CLS}$	re($\widehat{\alpha}_{CLS}$)	re($\widehat{\beta}_{CLS}$)	SS_{ML}	SS_{CLS}
0.047985	0.763104	0.842064	0.027980	0.823383	0.41690	0.078993	58.08	56.8

Table 5.14: The ML and CLS estimates applied to the real data of Table 5.13.

t	$ENF_{ML}(t)$	$ENF_{CLS}(t)$	$ONF(t)$
1000	9.341	8.260	7
2000	15.854	14.617	14
3000	21.603	20.410	18
4000	26.906	25.866	23
5000	31.901	31.083	33
6000	36.663	36.117	37
7000	41.240	41.005	41

Table 5.15: Comparisons of estimated numbers of failures with the observed number of failures for the real data.

Chapter 6

Parameter Estimation in Non-homogeneous Poisson Process Models for Software Reliability

6.1 The software reliability model

In this chapter a subclass of non-homogeneous Poisson processes is considered, which besides of its theoretically interesting structure it can be used to model software reliability. As alternative to the ML method, two other methods are proposed for estimating parameters in the process models. It is demonstrated that in certain cases the ML estimators do not exist despite the fact that we have sufficient information in the form of a large number of faults observed. The methods proposed yield satisfactory estimates of unknown parameters and can be also applied in some process models in which the ML estimators do not exist.

Let $\{N(t), t \geq 0\}$ be a NHPP with intensity function $\lambda(t)$ and the mean value function $\Lambda(t) = \int_0^t \lambda(u)du$ (cumulative intensity) having the parametric form defined by (3.4). Since the model defined by (3.4) has bounded mean value function, this is the reason for which this model is more appropriate as a software reliability model than the NHPP's with unbounded mean value function, because a software system contains only a finite number of faults.

As alternative to the ML method, for estimating parameters α and β of the NHPP with the mean value function $\Lambda(t; \alpha, \beta)$ one can use the LS method or the CLS method proposed in Section 5.4. The LS and/or the CLS method can be applied to the models for which the ML method fails. The LS and CLS methods can also be used to predict the next failure time.

The methods proposed will be applied to the general **NHPP software reliability model** defined by (3.4) and, in particular, to the NHPP software reliability model defined by the following cumulative intensity

$$\Lambda(t;\alpha,\beta) = \alpha\Big[1 - \exp(-t/\beta)\sum_{j=0}^{k}\frac{(t/\beta)^j}{j!}\Big], \quad \alpha,\beta > 0, \tag{6.1}$$

or equivalently, by

$$\lambda(t;\alpha,\beta) = \frac{\alpha(t/\beta)^k}{\beta k!}\exp(-t/\beta). \tag{6.2}$$

The model defined by (6.1) is a special case of the model (3.4) and was first mentioned in the paper of Khoshgoftaar (1988). It is called the **k-stage Erlangian NHPP software reliability model** .

In Section 6.6 some numerical results illustrating the accuracy of the proposed LS and CLS estimators with comparison to the ML estimators are presented for a special case of the Erlangian NHPP software reliability model.

6.2 ML method for the software reliability model

In particular, for the NHPP with the mean value function $\Lambda(t;\vartheta)$, $\vartheta = (\alpha,\beta)$, defined by (3.4), the **likelihood function** based on the observed arrival times $t_1, t_2, \dots, t_{N(T)}$ and $N(T)$ takes the form

$$L(\alpha,\beta) \propto \exp\left[-\alpha F(T/\beta)\right](\alpha/\beta)^{N(T)}\prod_{i=1}^{N(T)} f(t_i/\beta),$$

and the **log-likelihood function** is

$$\log L(\alpha,\beta) \propto -\alpha F(T/\beta) + N(T)\log(\alpha/\beta) + \sum_{i=1}^{N(T)}\log f(t_i/\beta). \tag{6.3}$$

The value of α maximizing the log likelihood function is

$$\alpha = \frac{N(T)}{F(T/\beta)} =: \alpha_{ML}(\beta). \tag{6.4}$$

Substituting this value into formula (6.3) yields

$$\log L(\alpha_{ML}(\beta),\beta) \propto -N(T) + N(T)\log\left[\frac{N(T)}{\beta F(T/\beta)}\right] + \sum_{i=1}^{N(T)}\log f(t_i/\beta).$$

Fact 6.2.1 The **ML estimators** $\widehat{\alpha}_{ML}$ and $\widehat{\beta}_{ML}$ of α and β are determined by

$$\widehat{\alpha}_{ML} = \frac{N(T)}{F(T/\widehat{\beta}_{ML})} \tag{6.5}$$

and the $\widehat{\beta}_{ML}$ which maximizes

$$\frac{\widetilde{L}(\beta) := N(T)\log\left[\frac{N(T)}{\beta F(T/\beta)}\right] + \sum_{i=1}^{N(T)} \log f(t_i}{\beta)} \tag{6.6}$$

with respect to β. $\square$

In particular, for the k-stage Erlangian NHPP software reliability model defined by (6.1), formulae (6.5) and (6.6) take the following form

$$\widehat{\alpha}_{ML} = \frac{N(T)}{1 - \exp(-T/\widehat{\beta}_{ML})\sum_{j=0}^{k} \frac{(T/\widehat{\beta}_{ML})^j}{j!}} \tag{6.7}$$

and

$$\widetilde{L}(\beta) := N(T)\log\left[\frac{N(T)}{\beta[1 - \exp(-T/\beta)]\sum_{j=0}^{k}\frac{(T/\beta)^j}{j!}}\right] + \sum_{i=1}^{N(T)} \log\frac{(t_i/\beta)^k}{k!} - \sum_{i=1}^{N(T)}\frac{t_i}{\beta}. \tag{6.8}$$

For the model with $k = 0$, formulae (6.7) and (6.8) have the following simple form

$$\widehat{\alpha}_{ML} = \frac{N(T)}{1 - \exp(-T/\widehat{\beta}_{ML})} \tag{6.9}$$

and

$$\widetilde{L}(\beta) := N(T)\log\left[\frac{N(T)}{\beta[1 - \exp(-T/\beta)]}\right] - \sum_{i=1}^{N(T)}\frac{t_i}{\beta}. \tag{6.10}$$

6.3 The LS and CLS methods as alternatives to the ML method

The ML estimators of the parameters α and β of the model (3.4) do not always exist. In particular, it follows from Theorem 2.1 of Zhao and Xie (1996) that for the model defined by (6.1) the ML estimators do not exist with the probability $P\left(\frac{1}{N(T)}\sum_{i=1}^{N(T)} t_i \geq \frac{k+1}{k+2}T\right)$, where $N(T)$ is the number of arrives up to time T and $t_1, \ldots, t_{N(T)}$ are the arrival times observed.

6.3.1 The LS method for the software reliability model

For the NHPP with the cumulative intensity function defined by (3.4) the sum of squares of (5.11) takes the following form

$$\begin{aligned} S^2_{LS}(\alpha,\beta) &= \sum_{i=1}^{N(T)} [\Lambda(t_i;\alpha,\beta) - \Lambda(t_{i-1};\alpha,\beta) - 1]^2 \\ &= \sum_{i=1}^{N(T)} [\alpha F(t_i/\beta) - \alpha F(t_{i-1}/\beta) - 1]^2 \\ &= \alpha^2 \sum_{i=1}^{N(T)} [F(t_i/\beta) - F(t_{i-1}/\beta)]^2 - 2\alpha F(t_{N(T)}/\beta) + N(T) \end{aligned} \tag{6.11}$$

Expression (6.11) regarded as a trinomial with respect to α is minimized by

$$\alpha = \alpha_{LS}(\beta) = \frac{F(t_{N(T)}/\beta)}{\sum_{i=1}^{N(T)} [F(t_i/\beta) - F(t_{i-1}/\beta)]^2}.$$

Substituting this value into formula (6.11) yields

$$S^2_{LS}(\alpha_{LS}(\beta),\beta) = N(T) - \frac{F^2(t_{N(T)}/\beta)}{\sum_{i=1}^{N(T)} [F(t_i/\beta) - F(t_{i-1}/\beta)]^2}.$$

Thus we have the following

Proposition 6.3.1 The **LS estimators** $\widehat{\alpha}_{LS}$ and $\widehat{\beta}_{LS}$ of α and β are determined by

$$\widehat{\alpha}_{LS} = \frac{F(t_{N(T)}/\widehat{\beta}_{LS})}{\sum_{i=1}^{N(T)} [F(t_i/\widehat{\beta}_{LS}) - F(t_{i-1}/\widehat{\beta}_{LS})]^2} \tag{6.12}$$

and the $\widehat{\beta}_{LS}$ which maximizes

$$\widetilde{S}^2_{LS}(\beta) = \frac{F^2(t_{N(T)}/\beta)}{\sum_{i=1}^{N(T)} [F(t_i/\beta) - F(t_{i-1}/\beta)]^2} \tag{6.13}$$

with respect to β. □

The estimator $\widehat{\beta}_{LS}$ of the parameter β is a solution to the equation

$$\begin{aligned} F(t_{N(T)}/\beta) &\sum_{i=1}^{N(T)} [F(t_i/\beta) - F(t_{i-1}/\beta)][f(t_i/\beta)t_i - f(t_{i-1}/\beta)t_{i-1}] \\ &- f(t_{N(T)}/\beta)t_{N(T)} \sum_{i=1}^{N(T)} [F(t_i/\beta) - F(t_{i-1}/\beta)]^2 = 0. \end{aligned}$$

6.3.2 The CLS method for the software reliability model

For the NHPP process considered the constraint given by (5.12) takes the form

$$\frac{1}{N(T)}\sum_{i=1}^{N(T)}[\Lambda(t_i;\alpha,\beta)-\Lambda(t_{i-1};\alpha,\beta)]$$

$$=\frac{\alpha}{N(T)}\sum_{i=1}^{N(T)}[F(t_i/\beta)-F(t_{i-1}/\beta)]=\frac{\alpha}{N(T)}F(t_{N(T)}/\beta)=1.$$

It then follows that

$$\alpha=\frac{N(T)}{F(t_{N(T)}/\beta)}=:\alpha_{CLS}(\beta).$$

Substituting this value into formula (6.11) we obtain

$$S_{LS}^2(\alpha_{CLS}(\beta),\beta)=N(T)\left[\frac{N(T)\sum_{i=1}^{N(T)}[F(t_i/\beta)-F(t_{i-1}/\beta)]^2}{F^2(t_{N(T)}/\beta)}-1\right].$$

Thus we have the following

Proposition 6.3.2 The **CLS estimators** $\widehat{\alpha}_{CLS}$ and $\widehat{\beta}_{CLS}$ of α and β are determined by

$$\widehat{\alpha}_{CLS}=\frac{N(T)}{F(t_{N(T)}/\widehat{\beta}_{CLS})} \tag{6.14}$$

and the $\widehat{\beta}_{CLS}$ which maximizes $\widetilde{S}_{LS}^2(\beta)$ given by (6.13). □

Let us recall that the CLS estimate $\widehat{\beta}_{CLS}$ takes the same values as the estimate $\widehat{\beta}_{LS}$ in the LS method.

In general, the optimization problems consisting in finding the ML, LS and CLS estimators for a NHPP, with any mean value parametric function $\Lambda(t;\alpha,\beta)$, can be defined by

$$(\widehat{\alpha}_{ML},\widehat{\beta}_{ML}) = \arg\max_{(\alpha,\beta)\in\mathbf{R}_+\times\mathbf{R}_+}\log L(\alpha,\beta), \tag{6.15}$$

$$(\widehat{\alpha}_{LS},\widehat{\beta}_{LS}) = \arg\min_{(\alpha,\beta)\in\mathbf{R}_+\times\mathbf{R}_+} S_{LS}^2(\alpha,\beta), \tag{6.16}$$

$$(\widehat{\alpha}_{CLS},\widehat{\beta}_{CLS}) = \arg\min_{(\alpha,\beta)\in C} S_{LS}^2(\alpha,\beta), \tag{6.17}$$

respectively, where $L(\alpha,\beta)$ is the corresponding likelihood function, $S^2_{LS}(\alpha,\beta)$ is defined by (5.11) with $\vartheta=(\alpha,\beta)$, and the restriction set

$$C=\Big\{(\alpha,\beta):\frac{1}{N(T)}\sum_{i=1}^{N(T)}[\Lambda(t_i;\alpha,\beta)-\Lambda(t_{i-1};\alpha,\beta)]=1\Big\}.$$

By Fact 6.2.1 and Propositions 6.3.1 and 6.3.2, for the NHPP defined by (3.4) we have the following

Corollary 6.3.1 In the case of the NHPP defined by (3.4), the optimization problems (6.15), (6.16) and (6.17) reduce to the following ones

$$\widehat{\beta}_{ML} = \arg\max_{\beta\in\mathbf{R}_+}\widetilde{L}(\beta), \tag{6.18}$$

$$\widehat{\beta}_{(C)LS} = \arg\max_{\beta\in\mathbf{R}_+}\widetilde{S}^2_{LS}(\beta), \tag{6.19}$$

where $\widetilde{L}(\beta)$ and $\widetilde{S}^2_{LS}(\beta)$ are defined by (6.6) and (6.13), respectively. The ML, LS and CLS estimators $\widehat{\alpha}_{ML}$, $\widehat{\alpha}_{LS}$ and $\widehat{\alpha}_{CLS}$ of the parameter α are determined by formulas (6.5), (6.12) and (6.14), respectively. □

6.4 Remarks on consistency of the estimators

It follows from Zhao and Xie (1996) that the ML estimators of α and β in the NHPP model defined by (3.5) are not consistent when the fixed total testing time T tends to infinity.

Recently, Nayak, Bose and Kundu (2008) proved that for the NHPP models with the intensity function of the form $\lambda(t)=\mu f_\vartheta(t), 0<\mu<\infty, \vartheta\in\Theta$, and satisfying the condition $\lim_{t\to\infty}\Lambda(t)<\infty$, there is no consistent estimator (not just the ML one) of a parametric function, say $\psi(\mu,\vartheta)$, if $\psi(\mu,\vartheta)$ is not a constant function of μ. It then follows that there is no consistent estimator of α in the NHPP model considered.

6.5 Some special models

6.5.1 The k-stage Erlangian NHPP software reliability model

Let us consider the special case of the model defined by (3.4), where

$$F(t/\beta)=1-\exp(-t/\beta)\sum_{j=0}^{k}\frac{(t/\beta)^j}{j!}, \tag{6.20}$$

and

$$f(t/\beta) = \frac{(t/\beta)^k}{k!} \exp(-t/\beta).$$

Thus this is the k-stage Erlangian NHPP software reliability model with the cumulated intensity function $\Lambda(t;\alpha,\beta)$ and intensity function $\lambda(t;\alpha,\beta)$ defined by (6.1) and (6.2), respectively. The parameter k is usually a small integer and it is assumed to be known.

The k-stage Erlangian software reliability model contains the exponential model proposed by Goel and Okumoto (1979) and the delayed s-shaped model studied by Yamada, Ohba and Osaki (1984) as special cases (with $k = 0$ and $k = 1$). These two models are the most widely used NHPP software reliability models in practice.

Applying Propositions 6.3.1 and 6.3.2 to the k-stage Erlangian NHPP we obtain the following result.

Proposition 6.5.3 For the k-stage Erlangian NHPP defined by (6.1) the **LS** and **CLS estimators** of α and β are determined by

$$\widehat{\alpha}_{LS} = \frac{1 - \exp(-t_{N(T)}/\widehat{\beta}) \sum_{j=0}^{k} \frac{(t_{N(T)}/\widehat{\beta})^j}{j!}}{\sum_{i=1}^{N(T)} \left[\exp(-t_{i-1}/\widehat{\beta}) \sum_{j=0}^{k} \frac{(t_{i-1}/\widehat{\beta})^j}{j!} - \exp(-t_i/\widehat{\beta}) \sum_{j=0}^{k} \frac{(t_i/\widehat{\beta})^j}{j!} \right]^2},$$

$$\widehat{\alpha}_{CLS} = \frac{N(T)}{1 - \exp(-t_{N(T)}/\widehat{\beta}) \sum_{j=0}^{k} \frac{(t_{N(T)}/\widehat{\beta})^j}{j!}},$$

and the $\widehat{\beta}$ which maximizes the function

$$\widetilde{S}^2_{LS}(\beta) = \frac{[1 - \exp(-t_{N(T)}/\beta) \sum_{j=0}^{k} \frac{(t_{N(T)}/\beta)^j}{j!}]^2}{\sum_{i=1}^{N(T)} \left[\exp(-t_{i-1}/\beta) \sum_{j=0}^{k} \frac{(t_{i-1}/\beta)^j}{j!} - \exp(-t_i/\beta) \sum_{j=0}^{k} \frac{(t_i/\beta)^j}{j!} \right]^2}$$

with respect to β. □

6.5.2 The Goel and Okumoto model

For $k = 0$ we have from Proposition 6.5.3 the following

Corollary 6.5.2 For the Goel and Okumoto model the **LS** and **CLS estimators** of α and β are determined by

$$\widehat{\alpha}_{LS} = \frac{1 - \exp(-t_{N(T)}/\widehat{\beta})}{\sum_{i=1}^{N(T)} [\exp(-t_{i-1}/\widehat{\beta}) - \exp(-t_i/\widehat{\beta})]^2}, \tag{6.21}$$

$$\widehat{\alpha}_{CLS} = \frac{N(T)}{1 - \exp(-t_{N(T)}/\widehat{\beta})}, \tag{6.22}$$

and the $\widehat{\beta}$ which maximizes the function

$$\widetilde{S}^2_{LS}(\beta) = \frac{[1 - \exp(-t_{N(T)}/\beta)]^2}{\sum_{i=1}^{N(T)} [\exp(-t_{i-1}/\beta) - \exp(-t_i/\beta)]^2} \tag{6.23}$$

with respect to β. □

6.6 Numerical results

In this section, for a given values of pairs of the parameters α and β in the Goel and Okumoto model we present some numerical results illustrating the accuracy of the proposed LS and CLS estimators of these parameters with comparison to the ML estimators. The numerical results are contained in Tables 6.1 – 6.16 for five observation times $T = 0.5, 1, 2, 5, 10$. The variability of an estimator $\widehat{\eta}$ of an unknown parameter η was measured by the root mean squared error which is defined by $\mathrm{se}(\widehat{\eta}) = \sqrt{(sd(\widehat{\eta}))^2 + (mean(\widehat{\eta}) - \eta)^2}$. The tables contain the numerical results obtained on the basis of 2000 generated random samples (trajectories of the NHPP) for each pair (α, β).

The values of estimators of α and β are evaluated using numerical constrained global optimization procedures to solve the problems (6.18) and (6.19) for the Goel and Okumoto model, i.e. for the functions $\widetilde{L}(\beta)$ and $\widetilde{S}^2_{LS}(\beta)$ defined by (6.10) and (6.23), respectively. The resulting estimates of β have been substituted into formulae (6.9), (6.21) and (6.22) to get estimates of α: $\widehat{\alpha}_{ML}$, $\widehat{\alpha}_{LS}$ and $\widehat{\alpha}_{CLS}$, respectively. In constructing the executable computer program, procedures of the package Mathematica 8.0 were used.

The results given in Tables 6.1, 6.4 and 6.7 demonstrate that for short observation times T the ML estimators as well as the (C)LS estimators do not always exist. For example, for $T = 0.5$ and the pairs of the parameters α and β numbered by No. 8 – 10 there is about 40 percentage of no-existence of the ML estimators as well as about 40 percentage of no-existence of the (C)LS estimators. However, there is about 10 percentage that the ML estimator does not exist, whereas the (C)LS estimator does exist. Note that in this case the ML estimator does not exist despite the fact that we have sufficient information in the form of a large number of faults observed in relatively short observation time. Table 6.16 gives the LS and CLS estimates of α and β for some distinct cases in which the ML estimator does not exist, whereas the (C)LS estimator does exist. The table shows that in some cases when the ML method fails one could apply the CLS method yielding satisfactory estimates.

In general, one observes a good performance of the CLS method. The numerical results show that the CLS method yields the estimates of α and β which are practically so accurate as the ML estimates. The LS method considerably underestimates the parameter α.

Tables 6.1, 6.4, 6.7, 6.10 and 6.13 demonstrate that the percentage of non-existence of the ML as well as the (C)LS estimators of α and β tends to zero as T grows. The likelihood function for this process is analytically tractable and it follows from the results of Zhao and Xie (1996) that the ML estimators exist with probability 1 as $T \to \infty$.

Legend:
MJN – the mean jump (fault) number (the estimate of the mean value of the process at time T),
MLJT – the mean last jump time,
DT – the difference time: T – MLJT,
RDT – the relative difference time: (T – MLJT)/MLJT,
M'S– the percentage of no-existence of the ML estimator and existence of the (C)LS estimator,
MS'- the percentage of existence of the ML estimator and no-existence of the (C)LS estimator,
M' - the percentage of no-existence of the ML estimator,
S' - the percentage of no-existence of the (C)LS estimator.

$T = 0.5$										
No.	α	β	MJN	MLJT	DT	RDT	M'S	MS'	M'	S'
1	100	0.01	99	0.0522	0.4478	857.4703	0	0	0	0
2	100	0.1	99	0.4217	0.0783	18.5647	0	0	0	0
3	100	0.2	91	0.4782	0.0218	4.5510	0	0	0	0
4	100	0.5	63	0.4863	0.0137	2.8233	0.65	4.65	1.20	5.20
5	200	0.5	126	0.4931	0.0069	1.3910	0.05	1.35	0.05	1.35
6	200	1	78	0.4917	0.0083	1.6903	3.95	11.8	12.6	20.5
7	400	2	88	0.4936	0.0064	1.2997	8.45	13.5	29.4	34.5
8	600	3	92	0.4942	0.0058	1.1834	9.20	14.4	36.8	41.9
9	1000	4	117	0.4956	0.0044	0.8848	10.8	12.9	40.5	42.6
10	1000	5	95	0.4946	0.0054	1.0943	11.0	10.6	45.0	44.6

Table 6.1: The overall simulation results.

$T = 0.5$							
No.	α	β	$\widehat{\alpha}_{ML}$	$\widehat{\beta}_{ML}$	$\widehat{\alpha}_{LS}$	$\widehat{\alpha}_{CLS}$	$\widehat{\beta}_{(C)LS}$
1	100	0.01	99.8345	0.0100	52.2729	101.1500	0.0104
2	100	0.1	99.9675	0.1002	52.3981	101.4977	0.1041
3	100	0.2	100.7728	0.2039	53.9772	103.7871	0.2159
4	100	0.5	118.3959	0.6442	72.6766	137.2270	0.7709
5	200	0.5	217.5279	0.5717	123.2118	240.2897	0.6517
6	200	1	261.0310	1.3775	135.0313	258.3311	1.3440
7	400	2	388.8401	1.9358	186.3353	357.3556	1.7416
8	600	3	459.7785	2.2096	221.3808	422.5211	1.9947
9	1000	4	615.6102	2.3567	308.7377	594.5541	2.2343
10	1000	5	495.7357	2.3140	234.5192	451.2913	2.0706

Table 6.2: The ML, LS and CLS estimates of α and β.

$T = 0.5$										
No.	α	$\widehat{\alpha}_{ML}$	$\widehat{\alpha}_{CLS}$	se($\widehat{\alpha}_{ML}$)	se($\widehat{\alpha}_{CLS}$)	β	$\widehat{\beta}_{ML}$	$\widehat{\beta}_{CLS}$	se($\widehat{\beta}_{ML}$)	se($\widehat{\beta}_{(C)LS}$)
1	100	99.8345	101.1500	9.90144	10.03231	0.01	0.0100	0.0104	0.00102	0.00173
2	100	99.9675	101.4977	10.07896	10.39099	0.1	0.1002	0.1041	0.01110	0.01819
3	100	100.7728	103.7871	11.08243	13.84747	0.2	0.2039	0.2159	0.03490	0.06168
4	100	118.3959	137.2270	73.33553	130.02049	0.5	0.6442	0.7709	0.57050	0.97258
5	200	217.5279	240.2897	75.58821	155.99518	0.5	0.5717	0.6517	0.31190	0.61326
6	200	261.0310	258.3311	234.83429	248.02734	1	1.3775	1.3440	1.46389	1.55398
7	400	388.8401	357.3556	315.26841	320.11645	2	1.9358	1.7416	1.77728	1.80022
8	600	459.7785	422.5211	384.00196	406.00755	3	2.2096	1.9947	2.04542	2.20174
9	1000	615.6102	594.5541	595.58671	647.47248	4	2.3567	2.2343	2.51025	2.74133
10	1000	495.7357	451.2913	633.66514	674.42133	5	2.3140	2.0706	3.31122	3.55877

Table 6.3: The ML and CLS estimates of α and β and their measures of variability.

$T = 1$											
No.	α	β	MJN	MLJT	DT	RDT	M'S	MS'	M'	S'
1	100	0.01	100	0.0517	0.9483	1833.8208	0	0	0	0
2	100	0.1	100	0.5147	0.4853	94.2821	0	0	0	0
3	100	0.2	99	0.8416	0.1584	18.8180	0	0	0	0
4	100	0.5	86	0.9655	0.0345	3.5686	0	0.1	0	0.1
5	200	0.5	172	0.9819	0.0181	1.8468	0	0	0	0
6	200	1	126	0.9862	0.0138	1.3983	0.05	1.25	0.2	1.4
7	400	2	157	0.9918	0.0082	0.8252	2.75	10.1	7.45	14.9
8	600	3	170	0.9927	0.0073	0.7310	4.90	13.4	17.7	26.2
9	1000	4	221	0.9949	0.0051	0.5162	7.40	14.6	24.8	31.9
10	1000	5	181	0.9938	0.0062	0.6223	9.10	14.3	33.7	38.9

Table 6.4: The overall simulation results.

$T=1$							
No.	α	β	$\widehat{\alpha}_{ML}$	$\widehat{\beta}_{ML}$	$\widehat{\alpha}_{LS}$	$\widehat{\alpha}_{CLS}$	$\widehat{\beta}_{(C)LS}$
1	100	0.01	100.1640	0.010	52.4544	101.5036	0.0104
2	100	0.1	100.1059	0.0998	52.4574	101.4632	0.1039
3	100	0.2	99.8849	0.2004	52.6806	101.4310	0.2089
4	100	0.5	101.5468	0.5217	55.2204	106.6399	0.5703
5	200	0.5	201.5128	0.5087	104.9520	205.8927	0.5295
6	200	1	214.5388	1.1097	120.0287	233.7313	1.2459
7	400	2	458.9113	2.3749	237.9817	464.1703	2.3770
8	600	3	617.2464	3.0991	287.8927	563.8495	2.7445
9	1000	4	919.1599	3.6178	428.8319	841.9079	3.2419
10	1000	5	785.1667	3.7907	370.1944	722.8201	3.4242

Table 6.5: The ML, LS and CLS estimates of α and β.

$T=1$										
No.	α	$\widehat{\alpha}_{ML}$	$\widehat{\alpha}_{CLS}$	se($\widehat{\alpha}_{ML}$)	se($\widehat{\alpha}_{CLS}$)	β	$\widehat{\beta}_{ML}$	$\widehat{\beta}_{CLS}$	se($\widehat{\beta}_{ML}$)	se($\widehat{\beta}_{(C)LS}$)
1	100	100.1640	101.5036	10.08721	10.35084	0.01	0.010	0.0104	0.00101	0.00164
2	100	100.1059	101.4632	10.10175	10.34173	0.1	0.0998	0.1039	0.01001	0.01670
3	100	99.8849	101.4310	9.84716	10.16721	0.2	0.2004	0.2089	0.02260	0.03708
4	100	101.5468	106.6399	13.83005	22.38979	0.5	0.5217	0.5703	0.12561	0.25610
5	200	201.5128	205.8927	18.16194	23.19862	0.5	0.5087	0.5295	0.07760	0.12606
6	200	214.5388	233.7313	64.26733	110.93369	1	1.1097	1.2459	0.50859	0.86813
7	400	458.9113	464.1703	238.88032	283.68599	2	2.3749	2.3770	1.52034	1.76431
8	600	617.2464	563.8495	309.60911	314.13970	3	3.0991	2.7445	1.82292	1.81713
9	1000	919.1599	841.9079	446.78420	464.90551	4	3.6178	3.2419	2.01064	2.08058
10	1000	785.1667	722.8201	435.79853	477.32375	5	3.7907	3.4242	2.38954	2.63843

Table 6.6: The ML and CLS estimates of α and β and their measures of variability.

$T=2$										
No.	α	β	MJN	MLJT	DT	RDT	M'S	MS'	M'	S'
1	100	0.01	100	0.0517	1.9483	3765.2803	0	0	0	0
2	100	0.1	100	0.5223	1.4777	282.9558	0	0	0	0
3	100	0.2	100	1.0294	0.9706	94.2811	0	0	0	0
4	100	0.5	98	1.8001	0.1999	11.1065	0	0	0	0
5	200	0.5	196	1.8883	0.1117	5.9167	0	0	0	0
6	200	1	172	1.9648	0.0352	1.7904	0	0	0	0
7	400	2	252	1.9857	0.0143	0.7193	0	0.45	0	0.45
8	600	3	291	1.9904	0.0096	0.4847	0.45	4.00	1.15	4.7
9	1000	4	393	1.9934	0.0066	0.3334	1.00	7.75	4.25	11.0
10	1000	5	329	1.9925	0.0075	0.3766	3.90	13.2	13.5	22.8

Table 6.7: The overall simulation results.

$T=2$							
No.	α	β	$\widehat{\alpha}_{ML}$	$\widehat{\beta}_{ML}$	$\widehat{\alpha}_{LS}$	$\widehat{\alpha}_{CLS}$	$\widehat{\beta}_{(C)LS}$
1	100	0.01	100.1210	0.010	52.2545	101.3905	0.0103
2	100	0.1	100.1305	0.100	52.4474	101.4425	0.1039
3	100	0.2	100.3011	0.2003	52.7626	101.6802	0.2090
4	100	0.5	100.0520	0.5021	52.8052	101.7973	0.5232
5	200	0.5	200.6302	0.5019	102.9576	202.2732	0.5146
6	200	1	200.9636	1.0145	104.4107	204.6843	1.0477
7	400	2	413.8203	2.1167	219.6360	432.8315	2.2552
8	600	3	632.6611	3.2300	330.6663	651.9038	3.3414
9	1000	4	1059.0552	4.2896	530.8757	1051.5443	4.2347
10	1000	5	998.1929	4.9716	484.0811	955.2734	4.6826

Table 6.8: The ML, LS and CLS estimates of α and β.

$T=2$										
No.	α	$\widehat{\alpha}_{ML}$	$\widehat{\alpha}_{CLS}$	se($\widehat{\alpha}_{ML}$)	se($\widehat{\alpha}_{CLS}$)	β	$\widehat{\beta}_{ML}$	$\widehat{\beta}_{CLS}$	se($\widehat{\beta}_{ML}$)	se($\widehat{\beta}_{(C)LS}$)
1	100	100.1210	101.3905	9.97175	10.24861	0.01	0.010	0.0103	0.00101	0.00166
2	100	100.1305	101.4425	9.80921	10.03947	0.1	0.100	0.1039	0.01010	0.01714
3	100	100.3011	101.6802	9.96455	10.19913	0.2	0.2003	0.2090	0.01990	0.03450
4	100	100.0520	101.7973	10.13708	10.57259	0.5	0.5021	0.5232	0.06226	0.10135
5	200	200.6302	202.2732	14.40906	14.75109	0.5	0.5019	0.5146	0.04273	0.06367
6	200	200.9636	204.6843	17.96299	22.03652	1	1.0145	1.0477	0.15364	0.23706
7	400	413.8203	432.8315	72.55890	118.92414	2	2.1167	2.2552	0.57770	0.96890
8	600	632.6611	651.9038	166.80740	220.84577	3	3.2300	3.3414	1.15839	1.52343
9	1000	1059.0552	1051.5443	304.00612	352.31364	4	4.2896	4.2347	1.53550	1.78781
10	1000	998.1929	955.2734	291.40354	314.83987	5	4.9716	4.6826	1.75030	1.90834

Table 6.9: The ML and CLS estimates of α and β and their measures of variability.

$T=5$										
No.	α	β	MJN	MLJT	DT	RDT	M'S	MS'	M'	S'
1	100	0.01	99	0.0516	4.9484	9588.3850	0	0	0	0
2	100	0.1	100	0.5149	4.4851	870.9928	0	0	0	0
3	100	0.2	99	1.0307	3.9693	385.1239	0	0	0	0
4	100	0.5	99	2.5602	2.4398	95.2995	0	0	0	0
5	200	0.5	199	2.9163	2.0837	71.4528	0	0	0	0
6	200	1	198	4.5183	0.4817	10.6615	0	0	0	0
7	400	2	366	4.9402	0.0598	1.2115	0	0	0	0
8	600	3	486	4.9743	0.0257	0.5164	0	0	0	0
9	1000	4	713	4.9856	0.0144	0.2881	0	0	0	0
10	1000	5	631	4.9865	0.0135	0.2708	0	0.6	0	0.6

Table 6.10: The overall simulation results.

$T = 5$							
No.	α	β	$\widehat{\alpha}_{ML}$	$\widehat{\beta}_{ML}$	$\widehat{\alpha}_{LS}$	$\widehat{\alpha}_{CLS}$	$\widehat{\beta}_{(C)LS}$
1	100	0.01	99.9015	0.0100	52.6948	101.3996	0.0104
2	100	0.1	100.2555	0.0999	52.5178	101.6082	0.1042
3	100	0.2	99.9290	0.2001	52.3069	101.2768	0.2084
4	100	0.5	99.9255	0.4997	52.4787	101.3558	0.5220
5	200	0.5	199.9839	0.4995	102.5104	201.2169	0.5121
6	200	1	199.7320	1.0001	102.4012	201.0208	1.0169
7	400	2	400.5251	2.0130	204.0943	403.5445	2.0481
8	600	3	601.1071	3.0166	305.5299	606.0451	3.0574
9	1000	4	1007.1716	4.0491	511.2744	1018.5122	4.1218
10	1000	5	1013.6533	5.1109	516.7032	1026.8711	5.2013

Table 6.11: The ML, LS and CLS estimates of α and β.

$T = 5$										
No.	α	$\widehat{\alpha}_{ML}$	$\widehat{\alpha}_{CLS}$	se($\widehat{\alpha}_{ML}$)	se($\widehat{\alpha}_{CLS}$)	β	$\widehat{\beta}_{ML}$	$\widehat{\beta}_{CLS}$	se($\widehat{\beta}_{ML}$)	se($\widehat{\beta}_{(C)LS}$)
1	100	99.9015	101.3996	10.16475	11.76965	0.01	0.0100	0.0104	0.00102	0.00288
2	100	100.2555	101.6082	9.87321	10.15561	0.1	0.0999	0.1042	0.01008	0.01636
3	100	99.9290	101.2768	9.87946	10.10133	0.2	0.2001	0.2084	0.01990	0.03409
4	100	99.9255	101.3558	10.05046	10.28690	0.5	0.4997	0.5220	0.05065	0.08931
5	200	199.9839	201.2169	14.25428	14.40052	0.5	0.4995	0.5121	0.03553	0.05595
6	200	199.7320	201.0208	14.29414	14.36701	1	1.0001	1.0169	0.07732	0.11306
7	400	400.5251	403.5445	22.47987	24.10586	2	2.0130	2.0481	0.17146	0.24931
8	600	601.1071	606.0451	37.38264	46.44106	3	3.0166	3.0574	0.31511	0.45285
9	1000	1007.1716	1018.5122	69.77062	94.41023	4	4.0491	4.1218	0.46060	0.67581
10	1000	1013.6533	1026.8711	102.65836	142.62692	5	5.1109	5.2013	0.79590	1.14531

Table 6.12: The ML and CLS estimates of α and β and their measures of variability.

$T = 10$										
No.	α	β	MJN	MLJT	DT	RDT	M'S	MS'	M'	S'
1	100	0.01	99	0.0517	9.9483	19253.8649	0	0	0	0
2	100	0.1	100	0.5208	9.4792	1820.1137	0	0	0	0
3	100	0.2	100	1.0264	8.9736	874.2933	0	0	0	0
4	100	0.5	99	2.5959	7.4041	285.2206	0	0	0	0
5	200	0.5	200	2.9125	7.0875	243.3421	0	0	0	0
6	200	1	200	5.8037	4.1963	72.3028	0	0	0	0
7	400	2	397	9.4281	0.5719	6.0660	0	0	0	0
8	600	3	578	9.8657	0.1343	1.3610	0	0	0	0
9	1000	4	915	9.9516	0.0484	0.4865	0	0	0	0
10	1000	5	865	9.9630	0.0370	0.3713	0	0	0	0

Table 6.13: The overall simulation results.

$T=10$							
No.	α	β	$\widehat{\alpha}_{ML}$	$\widehat{\beta}_{ML}$	$\widehat{\alpha}_{LS}$	$\widehat{\alpha}_{CLS}$	$\widehat{\beta}_{(C)LS}$
1	100	0.01	99.8560	0.010	52.3971	101.1986	0.0104
2	100	0.1	100.0595	0.1002	52.5124	101.3912	0.1043
3	100	0.2	100.0195	0.1997	52.5651	101.4310	0.2088
4	100	0.5	99.8295	0.4997	52.3013	101.1690	0.5187
5	200	0.5	200.2260	0.4997	102.7633	201.5462	0.5124
6	200	1	200.3899	0.9991	102.7611	201.5969	1.0222
7	400	2	399.8627	2.0059	202.4707	401.1568	2.0282
8	600	3	600.3302	3.0032	302.8656	602.0908	3.0245
9	1000	4	998.4622	4.0045	502.4132	1001.3562	4.0312
10	1000	5	1002.4721	5.0252	505.5559	1006.4256	5.0622

Table 6.14: The ML, LS and CLS estimates of α and β.

$T=10$											
No.	α	$\widehat{\alpha}_{ML}$	$\widehat{\alpha}_{CLS}$	se($\widehat{\alpha}_{ML}$)	se($\widehat{\alpha}_{CLS}$)	β	$\widehat{\beta}_{ML}$	$\widehat{\beta}_{CLS}$	se($\widehat{\beta}_{ML}$)	se($\widehat{\beta}_{(C)LS}$)
1	100	99.8560	101.1986	9.65000	9.85634	0.01	0.010	0.0104	0.00098	0.00163
2	100	100.0595	101.3912	9.73298	9.90838	0.1	0.1002	0.1043	0.01017	0.01701
3	100	100.0195	101.4310	10.08540	10.34112	0.2	0.1997	0.2088	0.02037	0.03460
4	100	99.8295	101.1690	10.16657	10.35191	0.5	0.4997	0.5187	0.04905	0.08528
5	200	200.2260	201.5462	14.44262	14.55868	0.5	0.4997	0.5124	0.03538	0.05697
6	200	200.3899	201.5969	14.24994	14.41365	1	0.9991	1.0222	0.07112	0.10774
7	400	399.8627	401.1568	20.14591	20.17262	2	2.0059	2.0282	0.11269	0.16418
8	600	600.3302	602.0908	25.36019	25.91476	3	3.0032	3.0245	0.16111	0.22880
9	1000	998.4622	1001.3562	35.13743	37.13600	4	4.0045	4.0312	0.20847	0.30171
10	1000	1002.4721	1006.4256	40.17269	46.87337	5	5.0252	5.0622	0.33539	0.47756

Table 6.15: The ML and CLS estimates of α and β and their measures of variability.

No.	α	β	T	M'S	$\widehat{\alpha}_{LS}$	$\widehat{\alpha}_{CLS}$	$\widehat{\beta}_{(C)LS}$
7	400	2	0.5	8.45	306.8854	592.5294	3.0851
8	600	3	0.5	9.20	344.1557	651.6280	3.2350
9	1000	4	0.5	10.8	495.1580	946.3645	3.6824
10	1000	5	0.5	11.0	365.4143	708.0149	3.4497
8	600	3	1	4.90	414.1261	814.9304	4.2234
9	1000	4	1	7.40	590.0903	1167.8799	4.7378
10	1000	5	1	9.10	492.6215	969.8215	4.8369
10	1000	5	2	3.90	618.7331	1230.9129	6.4051

Table 6.16: The LS and CLS estimates of α and β when the ML estimator does not exist, whereas the (C)LS estimator does exist.

Bibliography

[1] Aalen O. O. and Husebye E. (1991). Statistical analysis of repeated events forming renewal processes. *Statistics in Medicine*, 10:1227–1240.

[2] Andersen P. K., Bentzon M. W. and Klein J. P. (1996). Estimating the survival function in the proportional hazards regression model: a study of the small sample size properties. *Scand. J. Statist.*, 23(1):1–12.

[3] Andersen P. K., Borgan Ø., Gill R. D. and Keiding N. (1988). Censoring, truncation and filtering in statistical models based on counting processes. In *Statistical inference from stochastic processes (Ithaca, NY, 1987)*, volume 80 of *Contemp. Math.*, pages 19–60. Amer. Math. Soc., Providence, RI.

[4] Andersen P. K., Borgan Ø., Gill R. D. and Keiding N. (1993). *Statistical Models Based on Counting Processes*. Springer Series in Statistics. Springer-Verlag, New York.

[5] Andersen P. K., Hansen L. S. and Keiding N. (1991). Non- and semi-parametric estimation of transition probabilities from censored observation of a nonhomogeneous Markov process. *Scand. J. Statist.*, 18(2):153–167.

[6] Arjas E. (2002). Predictive inference and discontinuities. *J. Nonparametr. Stat.*, 14:31–42.

[7] Ascher H. and Feingold H. (1984). *Repairable Systems Reliability: Modeling, Inference, Misconceptions and their Causes*. Marcel Decker, New York.

[8] Bae J. and Lee E. Y. (2001). A repair policy with limited number of minimal repairs. *J. Nonparametr. Stat.*, 13:153163.

[9] Bagai I. and Jain K. (1994). Improvement, deterioration, and optimal replacement under agereplacement with minimal repair. *IEEE Trans. Reliab.*, 43:156–162.

[10] Barlow R. E. (2003). Mathematical reliability theory: from the beginning to the present time. In *Mathematical and statistical methods in reliability (Trondheim, 2002)*, volume 7 of *Ser. Qual. Reliab. Eng. Stat.*, pages 3–13. World Sci. Publ., River Edge, NJ.

[11] Barlow R. E. and Proschan F. (1975). *Statistical theory of reliability and life testing*. Holt, Rinehart and Winston, Inc., New York.

[12] Barlow R. E. and Proschan F. (1996). *Mathematical theory of reliability*, volume 17 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

[13] Bedford T. and Lindqvist B. H. (2004). The identifiability problem for repairable systems subject to competing risks. *Adv. in Appl. Probab.*, 36(3):774–790.

[14] Berman M. (1981a). Inhomogeneous and modulated gamma processes. *Biometrika*, 68(1):143–152.

[15] Berman M. (1981b). The maximum likelihood estimators of the parameters of the gamma distribution are always positively biased. *Comm. Statist. A— Theory Methods*, 10(7):693–697.

[16] Bharucha-Reid A. T. (1997). *Elements of the Theory of Markov Processes and Their Applications*. Dover Publications, Inc. (reprint of the original: McGraw-Hill Series in Probability and Statistics, McGraw-Hill Book Co., Inc., New York-Toronto-London (1960))., Mineola, NY.

[17] Bunea C., Cooke R. and Lindqvist B. H. (2003). Competing risk perspective on reliability databases. In *Mathematical and statistical methods in reliability (Trondheim, 2002)*, volume 7 of *Ser. Qual. Reliab. Eng. Stat.*, pages 355–370. World Sci. Publ., River Edge, NJ.

[18] Calabria R., Guida M. and Pulcini G. (1992). Power bounds for a test of equality of trends in k independent power law processes. *Comm. Statist. Theory Methods*, 21(11):3275–3290.

[19] Calabria R. and Pulcini G. (1996a). Maximum likelihood and Bayes prediction of current system lifetime. *Comm. Statist. Theory Methods*, 25(10):2297–2309.

[20] Calabria R. and Pulcini G. (1996b). Point estimation under asymmetric loss functions for left-truncated exponential samples. *Comm. Statist. Theory Methods*, 25(3):585–600.

[21] Chen Y. and Singpurwalla N. D. (1997). Unification of software reliability models by self-exciting point processes. *Advances in Applied Probability*, 29(2):337–352.

[22] Cox D. R. (1972). The statistical analysis of dependencies in point processes. In Lewis P., editor, *Stochastic Point Processes*, pages 155–166. Wiley, New York.

[23] Dalal S. R. and Mallows C. L. (1988). When should one stop testing software? *JASA*, 403:872–879.

[24] Deshpande J. V. and Purohit S. G. (2005). *Life Time Data: Statistical Models and Methods*. Series on Quality, Reliability and Engineering Statistics. Vol. 11. World Scientific.

[25] Engelhardt M. and Bain L. J. (1992). Statistical analysis of a Weibull process with left-censored data. In *Survival analysis: state of the art (Columbus, OH, 1991)*, volume 211 of *NATO Adv. Sci. Inst. Ser. E Appl. Sci.*, pages 173–195. Kluwer Acad. Publ., Dordrecht.

[26] Franz J. (1999). On repair models and estimators of wear-out limits. *Economic Quality Control*, 14(3-4):135–151.

[27] Goel A. L. and Okumoto K. (1979). Time-dependent error-detection rate model for software reliability and other performance measures. *IEEE Trans. Reliab.*, 28:206–211.

[28] Heggland K. and Lindqvist B. H. (2007). A non-parametric monotone maximum likelihood estimator of time trend for repairable system data. *Reliability Engineering and System Safety,*, 92:575–584.

[29] Huang C.-Y., Lyu M. R. and Kuo S.-Y. (2003). A unified scheme of some nonhomogeneous Poisson process models for software reliability estimation. *IEEE transactions on Software Engineering*, 29(3):261–269.

[30] Ibragimov I. A. (2001). Statistical problems in the theory of stochastic processes. In Hazewinkel M., editor, *Encyclopaedia of Mathematics*. Kluwer Academic Publishers, Dordrecht.

[31] Jeske D. and Pham H. (2001). On the maximum estimates for the Goel-Okumoto software reliability model. *The American Statistician*, 55(3):219–222.

[32] Jokiel-Rokita A. and Magiera R. (2010). Parameter estimation in nonhomogeneous Poisson process models for software reliability. Technical report, Wroclaw University of Technology, Institute of Mathematics and Computer Science.

[33] Jokiel-Rokita A. and Magiera R. (2011). Estimation of parameters for trend-renewal processes. *Statistics and Computing.*

[34] Karlin S. and Taylor H. M. (1975). *A First Course in Stochastic Processes.* Elsevier Science.

[35] Keiding N. and Andersen P. K. (1989). Nonparametric estimation of transition intensities and transition probabilities: a case study of a two-state Markov process. *J. Roy. Statist. Soc. Ser. C*, 38(2):319–329.

[36] Khoshgoftaar T. M. (1988). Nonhomogeneous Poisson processes for software reliability growth. In *COMSTAT'88*, pages 13–14, Copenhagen, Denmark.

[37] Kijima M. (1989). Some results for repairable systems with general repair. *Journal of Applied Probability*, 26(1):89–102.

[38] Küchler U. and Sørensen M. (1997). *Exponential Families of Stochastic Processes.* Springer Series in Statistics. Springer Verlag, New York.

[39] Kvaløy J. T. and Lindqvist B. H. (2003a). A class of tests for renewal process versus monotonic and nonmonotonic trend in repairable systems data. In *Mathematical and statistical methods in reliability (Trondheim, 2002)*, volume 7 of *Ser. Qual. Reliab. Eng. Stat.*, pages 401–414. World Sci. Publ., River Edge, NJ.

[40] Kvaløy J. T. and Lindqvist B. H. (2003b). Estimation and inference in nonparametric Cox-models: time transformation methods. *Comput. Statist.*, 18(2):205–221.

[41] Kvaløy J. T. and Lindqvist B. H. (2004). The covariate order method for nonparametric exponential regression and some applications in other lifetime models. In *Parametric and semiparametric models with applications to reliability, survival analysis, and quality of life*, Stat. Ind. Technol., pages 221–237. Birkhäuser Boston, Boston, MA.

[42] Lai C.-D. and Xie M. (2006). *Stochastic Ageing and Dependence for Reliability.* Springer, New York.

[43] Lakey M. J. and Rigdon S. E. (1992). The modulated power-law process. In *Proceedings of the 45th Annual Quality Congress*, pages 559–563.

[44] Langaas M., Lindqvist B. H. and Ferkingstad E. (2005). Estimating the proportion of true null hypotheses, with application to DNA microarray data. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 67(4):555–572.

[45] Langseth H. and Lindqvist B. H. (2003). A maintenance model for components exposed to several failure mechanisms and imperfect repair. In *Mathematical and statistical methods in reliability (Trondheim, 2002)*, volume 7 of *Ser. Qual. Reliab. Eng. Stat.*, pages 415–430. World Sci. Publ., River Edge, NJ.

[46] Langseth H. and Lindqvist B. H. (2006). Competing risks for repairable systems: a data study. *J. Statist. Plann. Inference*, 136(5):1687–1700.

[47] Lawless J. F. and Thiagarajah K. (1996). A point-process model incorporating renewals and time trends with application to repairable systems. *Technometrics*, 38(2):131–138.

[48] Lindqvist B. H. (1993). The trend-renewal process, a useful model for repairable systems. Malmö, Sweden. Society in Reliability Engineers, Scandinavian Chapter, Annual Conference.

[49] Lindqvist B. H. (2006). On the statistical modeling and analysis of repairable systems. *Statist. Sci.*, 21(4):532–551.

[50] Lindqvist B. H. and Doksum K. A., editors (2003). *Mathematical and statistical methods in reliability*, volume 7 of *Series on Quality, Reliability & Engineering Statistics.* World Scientific Publishing Co. Inc., River Edge, NJ.

[51] Lindqvist B. H., Elvebakk G. and Heggland K. (2003). The trend-renewal process for statistical analysis of repairable systems. *Technometrics*, 45(1):31–44.

[52] Lindqvist B. H., Kjønstad G. A. and Meland N. (1994). Testing for trend in repairable system data. In *Proceedings of ESREL'94*, La Boule, France.

[53] Lindqvist B. H. and Langseth H. (2005). Statistical modeling and inference for component failure times under preventive maintenance and independent

censoring. In *Modern statistical and mathematical methods in reliability*, volume 10 of *Ser. Qual. Reliab. Eng. Stat.*, pages 323–337. World Sci. Publ., Singapore.

[54] Lindqvist B. H., Støve B. and Langseth H. (2006). Modelling of dependence between critical failure and preventive maintenance: the repair alert model. *J. Statist. Plann. Inference*, 136(5):1701–1717.

[55] Lindqvist B. H. and Taraldsen G. (2005). Monte Carlo conditioning on a sufficient statistic. *Biometrika*, 92(2):451–464.

[56] Lindqvist B. H. and Taraldsen G. (2007). Conditional Monte Carlo based on sufficient statistics with applications. In *Advances in statistical modeling and inference*, volume 3 of *Ser. Biostat.*, pages 545–561. World Sci. Publ., Hackensack, NJ.

[57] Muralidharan K. (2008). A review of repairable systems and point process models. *ProbStat Forum*, 1(July):26–49.

[58] Nakagawa T. and Kowada M. (1983). Analysis of a system with minimal repair and its application to replacement policy. *Eur. J. Oper. Res.*, 12:176–182.

[59] Nayak T. K., Bose S. and Kundu S. (2008). On inconsistency of estimators of parameters of non-homogeneous Poisson process models for software reliability. *Statistics & Probability Letters*, 78:2217–2221.

[60] Nielsen G. G., Gill R. D., Andersen P. K. and Sørensen T. I. A. (1992). A counting process approach to maximum likelihood estimation in frailty models. *Scand. J. Statist.*, 19(1):25–43.

[61] Osborne J. A. and Severini T. A. (2000). Inference for exponential order statistic models based on an integrated likelihood function. *JASA*, 92(452):1220–1228.

[62] Peña E. A. and Hollander M. (2004). Models for Recurrent Events in Reliability and Survival Analysis. In Soyer R., Mazzuchi T. and Singpurwalla N., editors, *Mathematical Reliability: An Expository Perspective*, pages 105–118. Kluwer.

[63] Raftery A. E. (1988). Analysis of simple debugging model. *Applied Statistics*, 37(1):12–22.

[64] Rigdon S. E. and Basu A. P. (2000). *Statistical Methods for the Reliability of Repairable Systems.* Wiley.

[65] Ruggeri F. and Pievatolo A. (2002). On the reliability of repairable systems: methods and applications. In *Atti della XLI Riunione Scientifica della Società Italiana di Statistica - Sessioni plenarie e specializzate*, pages 241–250.

[66] Sheu S.-H. (1990). Periodic replacement when minimal repair costs depend on the age and the number of minimal repairs for a multi-unit system. *Microelectron. Reliab.*, 30:713–718.

[67] Singpurwalla N. D. and Wilson S. P. (1994). Software reliability modelling. *International Statistocal Review*, 62(3):289–317.

[68] Stocker R. S. and Peña E. A. (2007). A general class of parametric models for recurrent event data. *Technometrics*, 49(2):210–220.

[69] Thompson W. A. J. (1988). *Point process models with applications to safety and reliability.* Chapman & Hall, New York.

[70] Wang R. (2005). A mixture and self-exciting model for software reliability. *Statistics & Probability Letters*, 72:187–194.

[71] Yamada S., Ohba M. and Osaki S. (1984). S-shaped software reliability growth models and their applications. *IEEE Trans. Reliab.*, 33:289–292.

[72] Yamada S. and Osaki S. (1985). Software reliability growth modelling: models and applications. *IEEE Trans. Software Engrg*, 11:1431–1437.

[73] Zhao M. and Xie M. (1996). On maximum likelihood estimation for a general non-homogeneous Poisson process. *Scandinavian Journal of Statistics*, 23(4):597–607.

Index