SFB 649 Discussion Paper 2010-048

# Building Loss Models

Krzysztof Burnecki*
Joanna Janczura*
Rafał Weron**

* Hugo Steinhaus Center, Wrocław University of Technology, Poland
** Institute of Organization and Management, Wrocław University of Technology, Poland

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin



# Building Loss Models[1]

Krzysztof Burnecki[2], Joanna Janczura[2], and Rafał Weron[3]

**Abstract:** This paper is intended as a guide to building insurance risk (loss) models. A typical model for insurance risk, the so-called collective risk model, treats the aggregate loss as having a compound distribution with two main components: one characterizing the arrival of claims and another describing the severity (or size) of loss resulting from the occurrence of a claim. In this paper we first present efficient simulation algorithms for several classes of claim arrival processes. Then we review a collection of loss distributions and present methods that can be used to assess the goodness-of-fit of the claim size distribution. The collective risk model is often used in health insurance and in general insurance, whenever the main risk components are the number of insurance claims and the amount of the claims. It can also be used for modeling other non-insurance product risks, such as credit and operational risk.

**Keywords:** Insurance risk model; Loss distribution; Claim arrival process; Poisson process; Renewal process; Random variable generation; Goodness-of-fit testing

**JEL:** C15, C46, C63, G22, G32

---

[1] Chapter prepared for the 2nd edition of *Statistical Tools for Finance and Insurance*, P.Cizek, W.Härdle, R.Weron (eds.), Springer-Verlag, *forthcoming in 2011*.
Research partially supported by Deutsche Forschungsgemeinschaft via SFB 649 „Ökonomisches Risiko“, Humboldt-Universität zu Berlin.

[2] Hugo Steinhaus Center, Wrocław University of Technology, Poland
[3] Institute of Organization and Management, Wrocław University of Technology, Poland

# 1 Building Loss Models

*Krzysztof Burnecki, Joanna Janczura, and Rafał Weron*

## 1.1 Introduction

A loss model or actuarial risk model is a parsimonious mathematical description of the behavior of a collection of risks constituting an insurance portfolio. It is not intended to replace sound actuarial judgment. In fact, according to Willmot (2001), a well formulated model is consistent with and adds to intuition, but cannot and should not replace experience and insight. Moreover, a properly constructed loss model should reflect a balance between simplicity and conformity to the data since overly complex models may be too complicated to be useful.

A typical model for insurance risk, the so-called collective risk model, treats the aggregate loss as having a compound distribution with two main components: one characterizing the frequency (or incidence) of events and another describing the severity (or size or amount) of gain or loss resulting from the occurrence of an event (Kaas et al., 2008; Klugman, Panjer, and Willmot, 2008; Tse, 2009). The stochastic nature of both components is a fundamental assumption of a realistic risk model. In classical form it is defined as follows. If $\{N_t\}_{t\geq 0}$ is a process *counting* claim occurrences and $\{X_k\}_{k=1}^{\infty}$ is an independent sequence of positive independent and identically distributed (i.i.d.) random variables representing claim sizes, then the *risk process* $\{R_t\}_{t\geq 0}$ is given by

$$R_t = u + c(t) - \sum_{i=1}^{N_t} X_i. \tag{1.1}$$

The non-negative constant $u$ stands for the initial capital of the insurance company and the deterministic or stochastic function of time $c(t)$ for the premium from sold insurance policies. The sum $\{\sum_{i=1}^{N_t} X_i\}$ is the so-called *ag-*

*gregate claim process*, with the number of claims in the interval $(0, t]$ being modeled by the *counting process* $N_t$. Recall, that the latter is defined as $N_t = \max\{n : \sum_{i=1}^{n} W_i \leq t\}$, where $\{W_i\}_{i=0}^{\infty}$ is a sequence of positive random variables and $\sum_{i=1}^{0} W_i \equiv 0$. In the insurance risk context $N_t$ is also referred to as the *claim arrival process*.

The collective risk model is often used in health insurance and in general insurance, whenever the main risk components are the number of insurance claims and the amount of the claims. It can also be used for modeling other non-insurance product risks, such as credit and operational risk (Chernobai, Rachev, and Fabozzi, 2007; Panjer, 2006). In the former, for example, the main risk components are the number of credit events (either defaults or downgrades), and the amount lost as a result of the credit event.

The simplicity of the risk process defined in eqn. (1.1) is only illusionary. In most cases no analytical conclusions regarding the time evolution of the process can be drawn. However, it is this evolution that is important for practitioners, who have to calculate functionals of the risk process like the expected time to ruin and the ruin probability, see Chapter **??**. The modeling of the aggregate claim process consists of modeling the counting process $\{N_t\}$ and the claim size sequence $\{X_k\}$. Both processes are usually assumed to be independent, hence can be treated independently of each other (Burnecki, Härdle, and Weron, 2004). Modeling of the claim arrival process $\{N_t\}$ is treated in Section 1.2, where we present efficient algorithms for four classes of processes. Modeling of claim severities and testing the goodness-of-fit is covered in Sections 1.3 and 1.4, respectively. Finally, in Section 1.5 we build a model for the Danish fire losses dataset, which concerns major fire losses in profits that occurred between 1980 and 2002 and were recorded by Copenhagen Re.

## 1.2 Claim Arrival Processes

In this section we focus on efficient simulation of the claim arrival process $\{N_t\}$. This process can be simulated either via the arrival times $\{T_i\}$, i.e. moments when the $i$th claim occurs, or the inter-arrival times (or waiting times) $W_i = T_i - T_{i-1}$, i.e. the time periods between successive claims (Burnecki and Weron, 2005). Note that in terms of $W_i$'s the claim arrival process is given by $N_t = \sum_{n=1}^{\infty} I(T_n \leq t)$. In what follows we discuss four examples of $\{N_t\}$, namely the classical (homogeneous) Poisson process, the non-homogeneous Poisson process, the mixed Poisson process and the renewal process.

### 1.2.1 Homogeneous Poisson Process (HPP)

The most common and best known claim arrival process is the homogeneous Poisson process (HPP). It has stationary and independent increments and the number of claims in a given time interval is governed by the Poisson law. While this process is normally appropriate in connection with life insurance modeling, it often suffers from the disadvantage of providing an inadequate fit to insurance data in other coverages with substantial temporal variability.

Formally, a continuous-time stochastic process $\{N_t : t \geq 0\}$ is a (homogeneous) Poisson process with intensity (or rate) $\lambda > 0$ if (i) $\{N_t\}$ is a counting process, and (ii) the waiting times $W_i$ are independent and identically distributed and follow an exponential law with intensity $\lambda$, i.e. with mean $1/\lambda$ (see Section 1.3.2). This definition naturally leads to a simulation scheme for the successive arrival times $T_i$ of the Poisson process on the interval $(0, t]$:

**Algorithm HPP1 (Waiting times)**

**Step 1:** set $T_0 = 0$

**Step 2:** generate an exponential random variable $E$ with intensity $\lambda$

**Step 3:** if $T_{i-1} + E < t$ then set $T_i = T_{i-1} + E$ and return to step 2 else stop

Sample trajectories of homogeneous (and non-homogeneous) Poisson processes are plotted in the top panels of Figure 1.1. The thin solid line is a HPP with intensity $\lambda = 1$ (left) and $\lambda = 10$ (right). Clearly the latter jumps more often.

Alternatively, the homogeneous Poisson process can be simulated by applying the following property (Rolski et al., 1999). Given that $N_t = n$, the $n$ occurrence times $T_1, T_2, \ldots, T_n$ have the same distribution as the order statistics corresponding to $n$ i.i.d. random variables uniformly distributed on the interval $(0, t]$. Hence, the arrival times of the HPP on the interval $(0, t]$ can be generated as follows:

**Algorithm HPP2 (Conditional theorem)**

**Step 1:** generate a Poisson random variable $N$ with intensity $\lambda t$

**Step 2:** generate $N$ random variables $U_i$ distributed uniformly on $(0, 1)$, i.e. $U_i \sim \mathrm{U}(0, 1), i = 1, 2, \ldots, N$

**Step 3:** set $(T_1, T_2, \ldots, T_N) = t \cdot \mathrm{sort}\{U_1, U_2, \ldots, U_N\}$

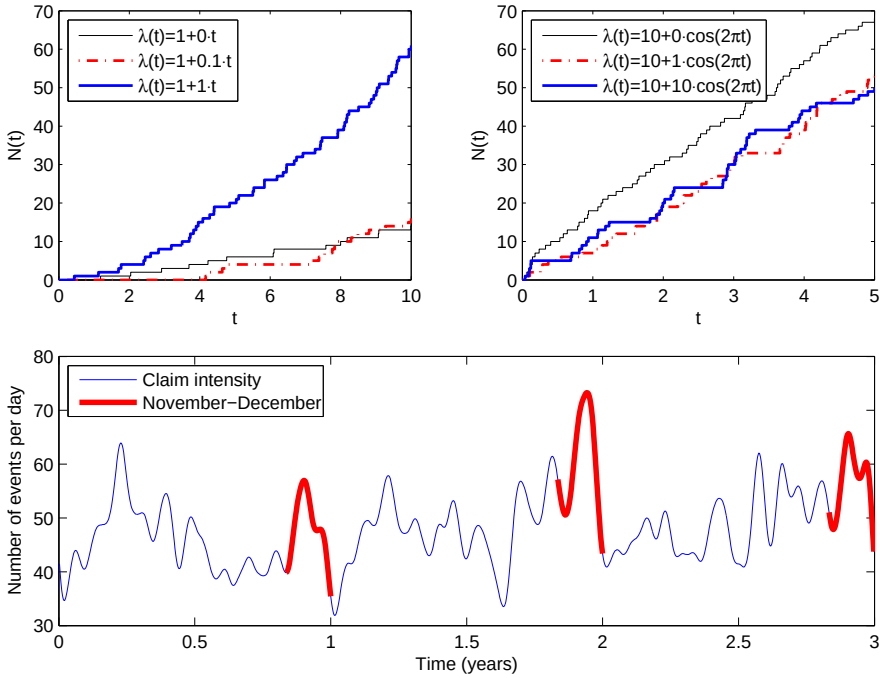


Figure 1.1: *Top left panel*: Sample trajectories of a NHPP with linear intensity $\lambda(t) = a + b \cdot t$. Note that the first process (with $b = 0$) is in fact a HPP. *Top right panel*: Sample trajectories of a NHPP with periodic intensity $\lambda(t) = a + b \cdot \cos(2\pi t)$. Again, the first process is a HPP. *Bottom panel*: Intensity of car accident claims in Greater Wrocław area, Poland, in the period 1998-2000 (data from one of the major insurers in the region). Note, the larger number of accidents in late Fall/early Winter due to worse weather conditions.

`STF2loss01.m`

In general, this algorithm will run faster than HPP1 as it does not involve a loop. The only two inherent numerical difficulties involve generating a Poisson random variable and sorting a vector of occurrence times. Whereas the latter problem can be solved, for instance, via the standard quicksort algorithm implemented in most statistical software packages (like `sortrows.m` in Matlab), the former requires more attention.

A straightforward algorithm to generate a Poisson random variable would take

$$N = \min\{n : U_1 \cdot \ldots \cdot U_n < \exp(-\lambda)\} - 1, \tag{1.2}$$

which is a consequence of the properties of the HPP (see above). However, for large $\lambda$, this method can become slow as the expected run time is proportional to $\lambda$. Faster, but more complicated methods are available. Ahrens and Dieter (1982) suggested a generator which utilizes acceptance-complement with truncated normal variates for $\lambda > 10$ and reverts to table-aided inversion otherwise. Stadlober (1989) adapted the ratio of uniforms method for $\lambda > 5$ and classical inversion for small $\lambda$'s. Hörmann (1993) advocated the transformed rejection method, which is a combination of the inversion and rejection algorithms. Statistical software packages often use variants of these methods. For instance, Matlab's `poissrnd.m` function uses the waiting time method (1.2) for $\lambda < 15$ and Ahrens' and Dieter's method for larger values of $\lambda$.

Finally, since for the HPP the expected value of the process $\mathrm{E}(N_t) = \lambda t$, it is natural to define the premium function as $c(t) = ct$, where $c = (1 + \theta)\mu\lambda$ and $\mu = \mathrm{E}(X_k)$. The parameter $\theta > 0$ is the relative safety loading which "guarantees" survival of the insurance company. With such a choice of the premium function we obtain the classical form of the risk process:

$$R_t = u + (1 + \theta)\mu\lambda t - \sum_{i=1}^{N_t} X_i. \tag{1.3}$$

### 1.2.2 Non-Homogeneous Poisson Process (NHPP)

The choice of a homogeneous Poisson process implies that the size of the portfolio cannot increase or decrease. In addition, it cannot describe situations, like in motor insurance, where claim occurrence epochs are likely to depend on the time of the year (worse weather conditions in Central Europe in late Fall/early Winter lead to more accidents, see the bottom panel in Figure 1.1) or of the week (heavier traffic occurs on Friday afternoons and before holidays). For modeling such phenomena the non-homogeneous Poisson process (NHPP) is much better. The NHPP can be thought of as a Poisson process with a variable (but predictable) intensity defined by the deterministic intensity (or rate) function $\lambda(t)$. Note that the increments of a NHPP do not have to be stationary. In the special case when the intensity takes a constant value $\lambda(t) = \lambda$, the NHPP reduces to the homogeneous Poisson process with intensity $\lambda$.

The simulation of the process in the non-homogeneous case is slightly more complicated than for the HPP. The first approach, known as the thinning or rejection method, is based on the following fact (Bratley, Fox, and Schrage, 1987; Ross, 2002). Suppose that there exists a constant $\overline{\lambda}$ such that $\lambda(t) \leq \overline{\lambda}$ for all $t$. Let $T_1^*, T_2^*, T_3^*, \ldots$ be the successive arrival times of a homogeneous Poisson process with intensity $\overline{\lambda}$. If we accept the $i$th arrival time $T_i^*$ with probability $\lambda(T_i^*)/\overline{\lambda}$, independently of all other arrivals, then the sequence $\{T_i\}_{i=0}^{\infty}$ of the accepted arrival times (in ascending order) forms a sequence of the arrival times of a non-homogeneous Poisson process with the rate function $\lambda(t)$. The resulting simulation algorithm on the interval $(0, t]$ reads as follows:

**Algorithm NHPP1 (Thinning)**

**Step 1:** set $T_0 = 0$ and $T^* = 0$

**Step 2:** generate an exponential random variable $E$ with intensity $\overline{\lambda}$

**Step 3:** if $T^* + E < t$ then set $T^* = T^* + E$ else stop

**Step 4:** generate a random variable $U$ distributed uniformly on $(0, 1)$

**Step 5:** if $U < \lambda(T^*)/\overline{\lambda}$ then set $T_i = T^*$ ($\rightarrow$ accept the arrival time)

**Step 6:** return to step 2

As mentioned in the previous section, the inter-arrival times of a homogeneous Poisson process have an exponential distribution. Therefore steps 2–3 generate the next arrival time of a homogeneous Poisson process with intensity $\overline{\lambda}$. Steps 4–5 amount to rejecting (hence the name of the method) or accepting a particular arrival as part of the thinned process (hence the alternative name). Note, that in this algorithm we generate a HPP with intensity $\overline{\lambda}$ employing the HPP1 algorithm. We can also generate it using the HPP2 algorithm, which in general is much faster.

The second approach is based on the observation that for a NHPP with rate function $\lambda(t)$ the increment $N_t - N_s$, $0 < s < t$, is distributed as a Poisson random variable with intensity $\widetilde{\lambda} = \int_s^t \lambda(u)du$ (Grandell, 1991). Hence, the cumulative distribution function $F_s$ of the waiting time $W_s$ is given by

$$\begin{aligned} F_s(t) &= \mathrm{P}(W_s \leq t) = 1 - \mathrm{P}(W_s > t) = 1 - \mathrm{P}(N_{s+t} - N_s = 0) = \\ &= 1 - \exp\left\{-\int_s^{s+t} \lambda(u)du\right\} = 1 - \exp\left\{-\int_0^t \lambda(s+v)dv\right\}. \end{aligned}$$

If the function $\lambda(t)$ is such that we can find a formula for the inverse $F_s^{-1}$ for each $s$, we can generate a random quantity $X$ with the distribution $F_s$ by using the inverse transform method. The simulation algorithm on the interval $(0,t]$, often called the *integration method*, can be summarized as follows:

**Algorithm NHPP2 (Integration)**

**Step 1:** set $T_0 = 0$

**Step 2:** generate a random variable $U$ distributed uniformly on $(0,1)$

**Step 3:** if $T_{i-1} + F_s^{-1}(U) < t$ set $T_i = T_{i-1} + F_s^{-1}(U)$ and return to step 2 else stop

The third approach utilizes a generalization of the property used in the HPP2 algorithm. Given that $N_t = n$, the $n$ occurrence times $T_1, T_2, \ldots, T_n$ of the non-homogeneous Poisson process have the same distributions as the order statistics corresponding to $n$ independent random variables distributed on the interval $(0,t]$, each with the common density function $f(v) = \lambda(v)/\int_0^t \lambda(u)du$, where $v \in (0,t]$. Hence, the arrival times of the NHPP on the interval $(0,t]$ can be generated as follows:

**Algorithm NHPP3 (Conditional theorem)**

**Step 1:** generate a Poisson random variable $N$ with intensity $\int_0^t \lambda(u)du$

**Step 2:** generate $N$ random variables $V_i, i = 1,2,\ldots N$ with density $f(v) = \lambda(v)/\int_0^t \lambda(u)du$.

**Step 3:** set $(T_1, T_2, \ldots, T_N) = \text{sort}\{V_1, V_2, \ldots, V_N\}$.

The performance of the algorithm is highly dependent on the efficiency of the computer generator of random variables $V_i$. Simulation of $V_i$'s can be done either via the inverse transform method by integrating the density $f(v)$ or via the acceptance-rejection technique using the uniform distribution on the interval $(0,t)$ as the reference distribution. In a sense, the former approach leads to Algorithm NHPP2, whereas the latter one to Algorithm NHPP1.

Sample trajectories of non-homogeneous Poisson processes are plotted in the top panels of Figure 1.1. In the top left panel realizations of a NHPP with linear intensity $\lambda(t) = a + b \cdot t$ are presented for the same value of parameter

$a$. Note, that the higher the value of parameter $b$, the more pronounced is the increase in the intensity of the process. In the top right panel realizations of a NHPP with periodic intensity $\lambda(t) = a + b \cdot \cos(2\pi t)$ are illustrated, again for the same value of parameter $a$. This time, for high values of parameter $b$ the events exhibit a seasonal behavior. The process has periods of high activity (grouped around natural values of $t$) and periods of low activity, where almost no jumps take place. Such a process is much better suited to model the seasonal intensity of car accident claims (see the bottom panel in Figure 1.1) than the HPP.

Finally, we note that since in the non-homogeneous case the expected value of the process at time $t$ is $\mathrm{E}(N_t) = \int_0^t \lambda(s)ds$, it is natural to define the premium function as $c(t) = (1+\theta)\mu \int_0^t \lambda(s)ds$. Then the risk process takes the form:

$$R_t = u + (1+\theta)\mu \int_0^t \lambda(s)ds - \sum_{i=1}^{N_t} X_i. \tag{1.4}$$

### 1.2.3 Mixed Poisson Process

In many situations the portfolio of an insurance company is diversified in the sense that the risks associated with different groups of policy holders are significantly different. For example, in motor insurance we might want to make a difference between male and female drivers or between drivers of different age. We would then assume that the claims come from a heterogeneous group of clients, each one of them generating claims according to a Poisson distribution with the intensity varying from one group to another.

Another practical reason for considering yet another generalization of the classical Poisson process is the following. If we measure the volatility of risk processes, expressed in terms of the index of dispersion $\mathrm{Var}(N_t)/\mathrm{E}(N_t)$, then often we obtain estimates in excess of one – a value obtained for the homogeneous and the non-homogeneous cases. These empirical observations led to the introduction of the mixed Poisson process (MPP), see Rolski et al. (1999).

In the mixed Poisson process the distribution of $\{N_t\}$ is given by a mixed Poisson distribution (Rolski et al., 1999). This means that, conditioning on an extrinsic random variable $\Lambda$ (called a structure variable), the random variable $\{N_t\}$ has a Poisson distribution. Typical examples for $\Lambda$ are two-point, gamma and general inverse Gaussian distributions (Teugels and Vynckier, 1996). Since

for each $t$ the claim numbers $\{N_t\}$ up to time $t$ are Poisson variates with intensity $\Lambda t$, it is now reasonable to consider the premium function of the form $c(t) = (1+\theta)\mu\Lambda t$. This leads to the following representation of the risk process:

$$R_t = u + (1+\theta)\mu\Lambda t - \sum_{i=1}^{N_t} X_i. \tag{1.5}$$

The MPP can be generated using the uniformity property: given that $N_t = n$, the $n$ occurrence times $T_1, T_2, \ldots, T_n$ have the same joint distribution as the order statistics corresponding to $n$ i.i.d. random variables uniformly distributed on the interval $(0, t]$ (Albrecht, 1982). The procedure starts with the simulation of $n$ as a realization of $N_t$ for a given value of $t$. This can be done in the following way: first a realization of a non-negative random variable $\Lambda$ is generated and, conditioned upon its realization, $N_t$ is simulated according to the Poisson law with parameter $\Lambda t$. Then we simulate $n$ uniform random numbers in $(0, t)$. After rearrangement, these values yield the sample $T_1 \leq \ldots \leq T_n$ of occurrence times. The algorithm is summarized below.

**Algorithm MPP1 (Conditional theorem)**

**Step 1:** generate a mixed Poisson random variable $N$ with intensity $\Lambda t$

**Step 2:** generate $N$ random variables $U_i$ distributed uniformly on $(0, 1)$, i.e. $U_i \sim \mathrm{U}(0, 1), i = 1, 2, \ldots, N$

**Step 3:** set $(T_1, T_2, \ldots, T_N) = t \cdot \mathrm{sort}\{U_1, U_2, \ldots, U_N\}$

### 1.2.4 Renewal Process

Generalizing the homogeneous Poisson process we come to the point where instead of making $\lambda$ non-constant, we can make a variety of different distributional assumptions on the sequence of waiting times $\{W_1, W_2, \ldots\}$ of the claim arrival process $\{N_t\}$. In some particular cases it might be useful to assume that the sequence is generated by a renewal process, i.e. the random variables $W_i$ are i.i.d., positive with a distribution function $F$. Note that the homogeneous Poisson process is a renewal process with exponentially distributed inter-arrival times. This observation lets us write the following algorithm for the generation of the arrival times of a renewal process on the interval $(0, t]$:

**Algorithm RP1 (Waiting times)**

**Step 1:** set $T_0 = 0$

**Step 2:** generate an $F$-distributed random variable $X$

**Step 3:** if $T_{i-1} + X < t$ then set $T_i = T_{i-1} + X$ and return to step 2 else stop

An important point in the previous generalizations of the Poisson process was the possibility to compensate risk and size fluctuations by the premiums. Thus, the premium rate had to be constantly adapted to the development of the claims. For renewal claim arrival processes, a constant premium rate allows for a constant safety loading (Embrechts and Klüppelberg, 1993). Let $\{N_t\}$ be a renewal process and assume that $W_1$ has finite mean $1/\lambda$. Then the premium function is defined in a natural way as $c(t) = (1 + \theta)\mu\lambda t$, like for the homogeneous Poisson process, which leads to the risk process of the form (1.3).

## 1.3 Loss Distributions

There are three basic approaches to deriving the loss distribution: empirical, analytical, and moment based. The empirical method, presented in Section 1.3.1, can be used only when large data sets are available. In such cases a sufficiently smooth and accurate estimate of the cumulative distribution function (cdf) is obtained. Sometimes the application of curve fitting techniques – used to smooth the empirical distribution function – can be beneficial. If the curve can be described by a function with a tractable analytical form, then this approach becomes computationally efficient and similar to the second method.

The analytical approach is probably the most often used in practice and certainly the most frequently adopted in the actuarial literature. It reduces to finding a suitable analytical expression which fits the observed data well and which is easy to handle. Basic characteristics and estimation issues for the most popular and useful loss distributions are discussed in Sections 1.3.2-1.3.8. Note, that sometimes it may be helpful to subdivide the range of the claim size distribution into intervals for which different methods are employed. For example, the small and medium size claims could be described by the empirical claim size distribution, while the large claims – for which the scarcity of data eliminates the use of the empirical approach – by an analytical loss distribution.

In some applications the exact shape of the loss distribution is not required. We may then use the moment based approach, which consists of estimating

only the lowest characteristics (moments) of the distribution, like the mean and variance. However, it should be kept in mind that even the lowest three or four moments do not fully define the shape of a distribution, and therefore the fit to the observed data may be poor. Further details on the moment based approach can be found e.g. in Daykin, Pentikainen, and Pesonen (1994).

Having a large collection of distributions to choose from, we need to narrow our selection to a single model and a unique parameter estimate. The type of the objective loss distribution can be easily selected by comparing the shapes of the empirical and theoretical mean excess functions. Goodness-of-fit can be measured by tests based on the empirical distribution function. Finally, the hypothesis that the modeled random event is governed by a certain loss distribution can be statistically tested. In Section 1.4 these statistical issues are thoroughly discussed.

### 1.3.1 Empirical Distribution Function

A natural estimate for the loss distribution is the observed (empirical) claim size distribution. However, if there have been changes in monetary values during the observation period, inflation corrected data should be used. For a sample of observations $\{x_1, \ldots, x_n\}$ the empirical distribution function (edf) is defined as:

$$F_n(x) = \frac{1}{n}\#\{i : x_i \leq x\}, \tag{1.6}$$

i.e. it is a piecewise constant function with jumps of size $1/n$ at points $x_i$. Very often, especially if the sample is large, the edf is approximated by a continuous, piecewise linear function with the "jump points" connected by linear functions, see Figure 1.2.

The empirical distribution function approach is appropriate only when there is a sufficiently large volume of claim data. This is rarely the case for the tail of the distribution, especially in situations where exceptionally large claims are possible. It is often advisable to divide the range of relevant values of claims into two parts, treating the claim sizes up to some limit on a discrete basis, while the tail is replaced by an analytical cdf.

If the claim statistics are too sparse to use the empirical approach it is desirable to find an explicit analytical expression for a loss distribution. It should be stressed, however, that many standard models in statistics – like the Gaussian distribution – are unsuitable for fitting the claim size distribution. The main

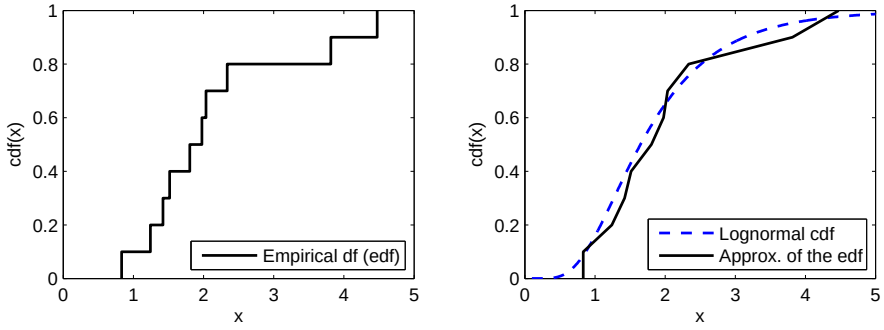


Figure 1.2: *Left panel*: Empirical distribution function (edf) of a 10-element log-normally distributed sample with parameters $\mu = 0.5$ and $\sigma = 0.5$, see Section 1.3.5. *Right panel*: Approximation of the edf by a continuous, piecewise linear function superimposed on the theoretical distribution function.

STF2loss02.m

reason for this is the strongly skewed nature of loss distributions. The log-normal, Pareto, Burr, and Weibull distributions are typical candidates to be considered in applications. However, before we review these probability laws we introduce two very versatile distributions – the exponential and gamma.

### 1.3.2 Exponential Distribution

Consider the random variable with the following density and distribution functions, respectively:

$$f(x) = \beta e^{-\beta x}, \qquad x > 0, \tag{1.7}$$
$$F(x) = 1 - e^{-\beta x}, \qquad x > 0. \tag{1.8}$$

This distribution is called the exponential distribution with parameter (or intensity) $\beta > 0$. The Laplace transform of (1.7) is

$$L(t) \stackrel{\text{def}}{=} \int_0^\infty e^{-tx} f(x) dx = \frac{\beta}{\beta + t}, \qquad t > -\beta, \tag{1.9}$$

yielding the general formula for the $k$-th raw moment

$$m_k \stackrel{\text{def}}{=} (-1)^k \frac{\partial^k L(t)}{\partial t^k}\Big|_{t=0} = \frac{k!}{\beta^k}. \tag{1.10}$$

The mean and variance are thus $\beta^{-1}$ and $\beta^{-2}$, respectively. The maximum likelihood estimator (equal to the method of moments estimator) for $\beta$ is given by:

$$\hat{\beta} = \frac{1}{\hat{m}_1}, \tag{1.11}$$

where

$$\hat{m}_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k, \tag{1.12}$$

is the sample $k$-th raw moment.

To generate an exponential random variable $X$ with intensity $\beta$ we can use the inverse transform (or inversion) method (Devroye, 1986; Ross, 2002). The method consists of taking a random number $U$ distributed uniformly on the interval $(0, 1)$ and setting $X = F^{-1}(U)$, where $F^{-1}(x) = -\frac{1}{\beta}\log(1-x)$ is the inverse of the exponential cdf (1.8). In fact we can set $X = -\frac{1}{\beta}\log U$ since $(1-U)$ has the same distribution as $U$.

The exponential distribution has many interesting features. As we have seen in Section 1.2.1, it arises as the inter-occurrence time of the events in a HPP. It has the memoryless property, i.e. $\mathrm{P}(X > x + y | X > y) = \mathrm{P}(X > x)$. Further, the $n$-th root of the Laplace transform (1.9) is

$$L(t) = \left(\frac{\beta}{\beta + t}\right)^{\frac{1}{n}}, \tag{1.13}$$

which is the Laplace transform of a gamma variate (see Section 1.3.4). Thus the exponential distribution is infinitely divisible.

The exponential distribution is often used in developing models of insurance risks. This usefulness stems in a large part from its many and varied tractable mathematical properties. However, a disadvantage of the exponential distribution is that its density is monotone decreasing (see Figure 1.3), a situation which may not be appropriate in some practical situations.

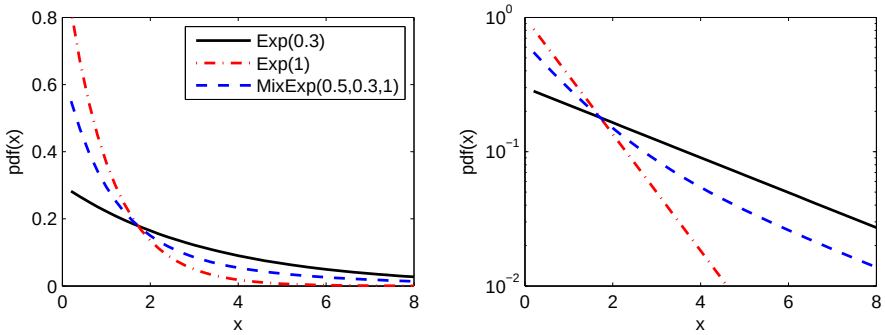


Figure 1.3: *Left panel:* A probability density function (pdf) of a mixture of two exponential distributions with mixing parameter $a = 0.5$ superimposed on the pdfs of the component distributions. *Right panel:* Semi-logarithmic plots of the pdfs displayed in the left panel. The exponential pdfs are now straight lines with slopes $-\beta$. Note, the curvature of the pdf of the mixture of two exponentials.

STF2loss03.m

### 1.3.3 Mixture of Exponential Distributions

Using the technique of mixing one can construct a wide class of distributions. The most commonly used in applications is a mixture of two exponentials, see Chapter **??**. In Figure 1.3 a pdf of a mixture of two exponentials is plotted together with the pdfs of the mixing laws. Note, that the mixing procedure can be applied to arbitrary distributions.

The construction goes as follows. Let $a_1, a_2, \ldots, a_n$ denote a series of non-negative weights satisfying $\sum_{i=1}^{n} a_i = 1$. Let $F_1(x), F_2(x), \ldots, F_n(x)$ denote an arbitrary sequence of exponential distribution functions given by the parameters $\beta_1, \beta_2, \ldots, \beta_n$, respectively. Then, the distribution function:

$$F(x) = \sum_{i=1}^{n} a_i F_i(x) = \sum_{i=1}^{n} a_i \left\{1 - \exp(-\beta_i x)\right\}, \tag{1.14}$$

is called a mixture of $n$ exponential distributions (exponentials). The density

function of the constructed distribution is

$$f(x) = \sum_{i=1}^{n} a_i f_i(x) = \sum_{i=1}^{n} a_i \beta_i \exp(-\beta_i x), \tag{1.15}$$

where $f_1(x), f_2(x), \ldots, f_n(x)$ denote the density functions of the input exponential distributions. The Laplace transform of (1.15) is

$$L(t) = \sum_{i=1}^{n} a_i \frac{\beta_i}{\beta_i + t}, \qquad t > -\min_{i=1\ldots n}\{\beta_i\}, \tag{1.16}$$

yielding the general formula for the $k$-th raw moment

$$m_k = \sum_{i=1}^{n} a_i \frac{k!}{\beta_i^k}. \tag{1.17}$$

The mean is thus $\sum_{i=1}^{n} a_i \beta_i^{-1}$. The maximum likelihood and method of moments estimators for the mixture of $n$ ($n \geq 2$) exponential distributions can only be evaluated numerically.

Simulation of variates defined by (1.14) can be performed using the composition approach (Ross, 2002). First generate a random variable $I$, equal to $i$ with probability $a_i$, $i = 1, ..., n$. Then simulate an exponential variate with intensity $\beta_I$. Note, that the method is general in the sense that it can be used for any set of distributions $F_i$'s.

### 1.3.4 Gamma Distribution

The probability law with density and distribution functions given by:

$$f(x) = \beta(\beta x)^{\alpha-1} \frac{e^{-\beta x}}{\Gamma(\alpha)}, \qquad x > 0, \tag{1.18}$$

$$F(x) = \int_0^x \beta(\beta s)^{\alpha-1} \frac{e^{-\beta s}}{\Gamma(\alpha)} ds, \qquad x > 0, \tag{1.19}$$

where $\alpha$ and $\beta$ are non-negative, is known as the gamma (or Pearson's Type III) distribution. In the above formulas

$$\Gamma(a) \stackrel{\text{def}}{=} \int_0^\infty y^{a-1} e^{-y} dy, \tag{1.20}$$

is the standard gamma function. Moreover, for $\beta = 1$ the integral in (1.19):

$$\Gamma(\alpha, x) \stackrel{\text{def}}{=} \frac{1}{\Gamma(\alpha)} \int_0^x s^{\alpha-1} e^{-s} ds, \tag{1.21}$$

is called the incomplete gamma function. If the shape parameter $\alpha = 1$, the exponential distribution results. If $\alpha$ is a positive integer, the distribution is called an Erlang law. If $\beta = \frac{1}{2}$ and $\alpha = \frac{\nu}{2}$ then it is called a chi-squared ($\chi^2$) distribution with $\nu$ degrees of freedom, see the top panels in Figure 1.4. Moreover, a mixed Poisson distribution with gamma mixing distribution is negative binomial.

The gamma distribution is closed under convolution, i.e. a sum of independent gamma variates with the same parameter $\beta$ is again gamma distributed with this $\beta$. Hence, it is infinitely divisible. Moreover, it is right-skewed and approaches a normal distribution in the limit as $\alpha$ goes to infinity.

The Laplace transform of the gamma distribution is given by:

$$L(t) = \left(\frac{\beta}{\beta + t}\right)^{\alpha}, \qquad t > -\beta. \tag{1.22}$$

The $k$-th raw moment can be easily derived from the Laplace transform:

$$m_k = \frac{\Gamma(\alpha + k)}{\Gamma(\alpha)\beta^k}. \tag{1.23}$$

Hence, the mean and variance are

$$\mathrm{E}(X) = \frac{\alpha}{\beta}, \tag{1.24}$$

$$\mathrm{Var}(X) = \frac{\alpha}{\beta^2}. \tag{1.25}$$

Finally, the method of moments estimators for the gamma distribution parameters have closed form expressions:

$$\hat{\alpha} = \frac{\hat{m}_1^2}{\hat{m}_2 - \hat{m}_1^2}, \tag{1.26}$$

$$\hat{\beta} = \frac{\hat{m}_1}{\hat{m}_2 - \hat{m}_1^2}, \tag{1.27}$$

but maximum likelihood estimators can only be evaluated numerically.

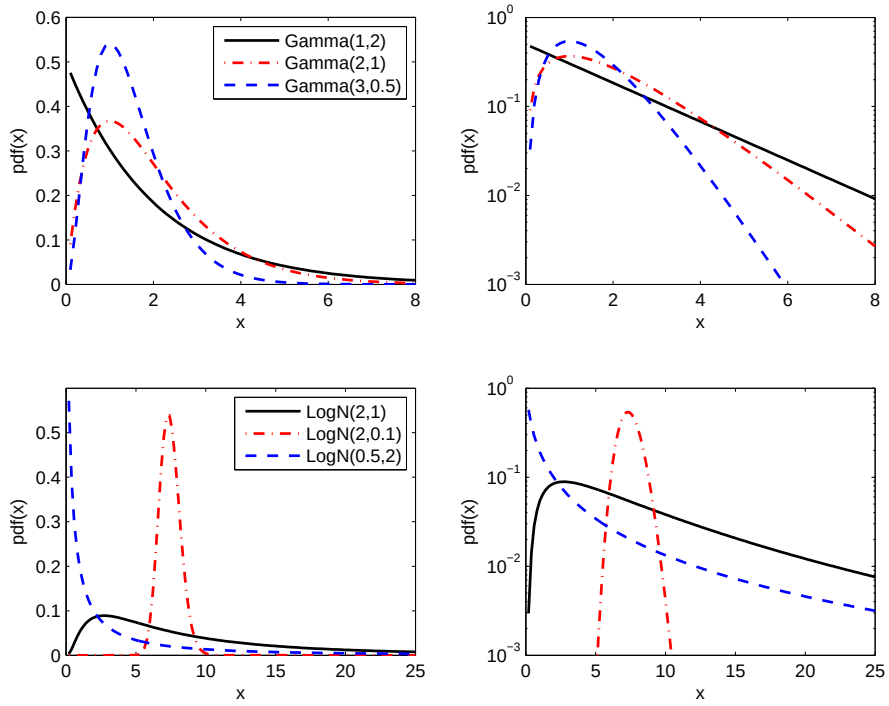


Figure 1.4: *Top panels:* Three sample gamma pdfs, Gamma$(\alpha, \beta)$, on linear and semi-logarithmic plots. Note, that the first one (black solid line) is an exponential law, while the last one (dashed blue line) is a $\chi^2$ distribution with $\nu = 6$ degrees of freedom. *Bottom panels:* Three sample log-normal pdfs, LogN$(\mu, \sigma)$, on linear and semi-logarithmic plots. For small $\sigma$ the log-normal distribution resembles the Gaussian.

STF2loss04.m

Unfortunately, simulation of gamma variates is not straightforward. For $\alpha < 1$ a simple but slow algorithm due to Jöhnk (1964) can be used, while for $\alpha > 1$ the rejection method is more optimal (Bratley, Fox, and Schrage, 1987; Devroye, 1986). The gamma law is one of the most important distributions for modeling because it has very tractable mathematical properties. As we have seen above it is also very useful in creating other distributions, but by itself is rarely a reasonable model for insurance claim sizes.

### 1.3.5 Log-Normal Distribution

Consider a random variable $X$ which has the normal distribution with density

$$f_N(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right\}, \qquad -\infty < x < \infty. \tag{1.28}$$

Let $Y = e^X$ so that $X = \log Y$. Then the probability density function of $Y$ is given by:

$$f(y) = f_N(\log y)\frac{1}{y} = \frac{1}{\sqrt{2\pi}\sigma y} \exp\left\{-\frac{1}{2}\frac{(\log y-\mu)^2}{\sigma^2}\right\}, \qquad y > 0, \tag{1.29}$$

where $\sigma > 0$ is the scale and $-\infty < \mu < \infty$ is the location parameter. The distribution of $Y$ is called log-normal; in econometrics it is also known as the Cobb-Douglas law. The log-normal cdf is given by:

$$F(y) = \Phi\left(\frac{\log y - \mu}{\sigma}\right), \qquad y > 0, \tag{1.30}$$

where $\Phi(\cdot)$ is the standard normal (with mean 0 and variance l) distribution function. The $k$-th raw moment $m_k$ of the log-normal variate can be easily derived using results for normal random variables:

$$m_k = \mathrm{E}\left(Y^k\right) = \mathrm{E}\left(e^{kX}\right) = M_X(k) = \exp\left(\mu k + \frac{\sigma^2 k^2}{2}\right), \tag{1.31}$$

where $M_X(z)$ is the moment generating function of the normal distribution. In particular, the mean and variance are

$$\mathrm{E}(X) = \exp\left(\mu + \frac{\sigma^2}{2}\right), \tag{1.32}$$

$$\mathrm{Var}(X) = \left\{\exp\left(\sigma^2\right) - 1\right\}\exp\left(2\mu + \sigma^2\right), \tag{1.33}$$

respectively. For both standard parameter estimation techniques the estimators are known in closed form. The method of moments estimators are given by:

$$\hat{\mu} = 2\log\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) - \frac{1}{2}\log\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right), \tag{1.34}$$

$$\hat{\sigma}^2 = \log\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right) - 2\log\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right), \tag{1.35}$$

while the maximum likelihood estimators by:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}\log(x_i), \tag{1.36}$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left\{\log(x_i)-\hat{\mu}\right\}^2. \tag{1.37}$$

Finally, the generation of a log-normal variate is straightforward. We simply have to take the exponent of a normal variate.

The log-normal distribution is very useful in modeling of claim sizes. For large $\sigma$ its tail is (semi-)heavy – heavier than the exponential but lighter than power-law, see the bottom panels in Figure 1.4. For small $\sigma$ the log-normal resembles a normal distribution, although this is not always desirable. It is infinitely divisible and closed under scale and power transformations. However, it also suffers from some drawbacks. Most notably, the Laplace transform does not have a closed form representation and the moment generating function does not exist.

### 1.3.6 Pareto Distribution

Suppose that a variate $X$ has (conditional on $\beta$) an exponential distribution with intensity $\beta$ (i.e. with mean $\beta^{-1}$, see Section 1.3.2). Further, suppose that $\beta$ itself has a gamma distribution (see Section 1.3.4). The unconditional distribution of $X$ is a mixture and is called the Pareto distribution. Moreover, it can be shown that if $X$ is an exponential random variable and $Y$ is a gamma random variable, then $X/Y$ is a Pareto random variable.

The density and distribution functions of a Pareto variate are given by:

$$f(x) = \frac{\alpha\lambda^{\alpha}}{(\lambda+x)^{\alpha+1}}, \qquad x>0, \tag{1.38}$$

$$F(x) = 1-\left(\frac{\lambda}{\lambda+x}\right)^{\alpha}, \qquad x>0, \tag{1.39}$$

respectively. Clearly, the shape parameter $\alpha$ and the scale parameter $\lambda$ are both positive. The $k$-th raw moment:

$$m_k = \lambda^k k!\frac{\Gamma(\alpha-k)}{\Gamma(\alpha)}, \tag{1.40}$$

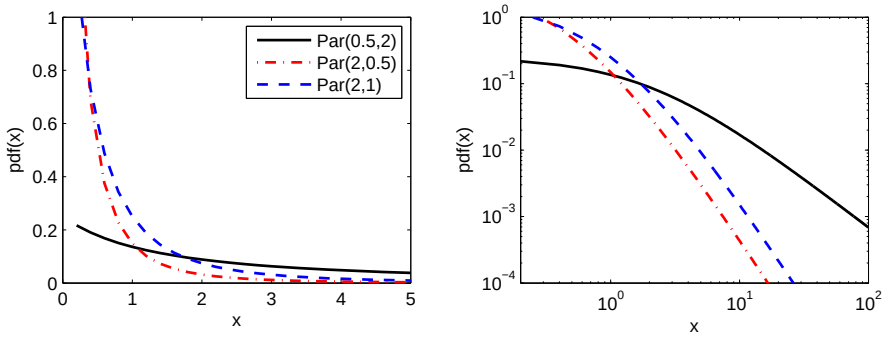


Figure 1.5: Three sample Pareto pdfs, $\mathrm{Par}(\alpha, \lambda)$, on linear and double-logarithmic plots. The thick power-law tails of the Pareto distribution (asymptotically linear in the log-log scale) are clearly visible.

STF2loss05.m

exists only for $k < \alpha$. The mean and variance are thus:

$$\mathrm{E}(X) = \frac{\lambda}{\alpha - 1}, \tag{1.41}$$

$$\mathrm{Var}(X) = \frac{\alpha\lambda^2}{(\alpha-1)^2(\alpha-2)}, \tag{1.42}$$

respectively. Note, that the mean exists only for $\alpha > 1$ and the variance only for $\alpha > 2$. Hence, the Pareto distribution has very thick (or heavy) tails, see Figure 1.5. The method of moments estimators are given by:

$$\hat{\alpha} = 2\frac{\hat{m}_2 - \hat{m}_1^2}{\hat{m}_2 - 2\hat{m}_1^2}, \tag{1.43}$$

$$\hat{\lambda} = \frac{\hat{m}_1\hat{m}_2}{\hat{m}_2 - 2\hat{m}_1^2}, \tag{1.44}$$

where, as before, $\hat{m}_k$ is the sample $k$-th raw moment (1.12). Note, that the estimators are well defined only when $\hat{m}_2 - 2\hat{m}_1^2 > 0$. Unfortunately, there are no closed form expressions for the maximum likelihood estimators and they can only be evaluated numerically.

Like for many other distributions the simulation of a Pareto variate $X$ can be conducted via the inverse transform method. The inverse of the cdf (1.39)

has a simple analytical form $F^{-1}(x) = \lambda\left\{(1-x)^{-1/\alpha} - 1\right\}$. Hence, we can set $X = \lambda\left(U^{-1/\alpha} - 1\right)$, where $U$ is distributed uniformly on the unit interval. We have to be cautious, however, when $\alpha$ is larger but very close to one. The theoretical mean exists, but the right tail is very heavy. The sample mean will, in general, be significantly lower than $\mathrm{E}(X)$.

The Pareto law is very useful in modeling claim sizes in insurance, due in large part to its extremely thick tail. Its main drawback lies in its lack of mathematical tractability in some situations. Like for the log-normal distribution, the Laplace transform does not have a closed form representation and the moment generating function does not exist. Moreover, like the exponential pdf the Pareto density (1.38) is monotone decreasing, which may not be adequate in some practical situations.

### 1.3.7 Burr Distribution

Experience has shown that the Pareto formula is often an appropriate model for the claim size distribution, particularly where exceptionally large claims may occur. However, there is sometimes a need to find heavy-tailed distributions which offer greater flexibility than the Pareto law, including a non-monotone pdf. Such flexibility is provided by the Burr distribution and its additional shape parameter $\tau > 0$. If $Y$ has the Pareto distribution, then the distribution of $X = Y^{1/\tau}$ is known as the Burr distribution, see the top panels in Figure 1.6. Its density and distribution functions are given by:

$$f(x) = \tau\alpha\lambda^{\alpha}\frac{x^{\tau-1}}{(\lambda + x^{\tau})^{\alpha+1}}, \qquad x > 0, \tag{1.45}$$

$$F(x) = 1 - \left(\frac{\lambda}{\lambda + x^{\tau}}\right)^{\alpha}, \qquad x > 0, \tag{1.46}$$

respectively. The $k$-th raw moment

$$m_k = \frac{1}{\Gamma(\alpha)}\lambda^{k/\tau}\Gamma\left(1 + \frac{k}{\tau}\right)\Gamma\left(\alpha - \frac{k}{\tau}\right), \tag{1.47}$$

exists only for $k < \tau\alpha$. Naturally, the Laplace transform does not exist in a closed form and the distribution has no moment generating function as it was the case with the Pareto distribution.

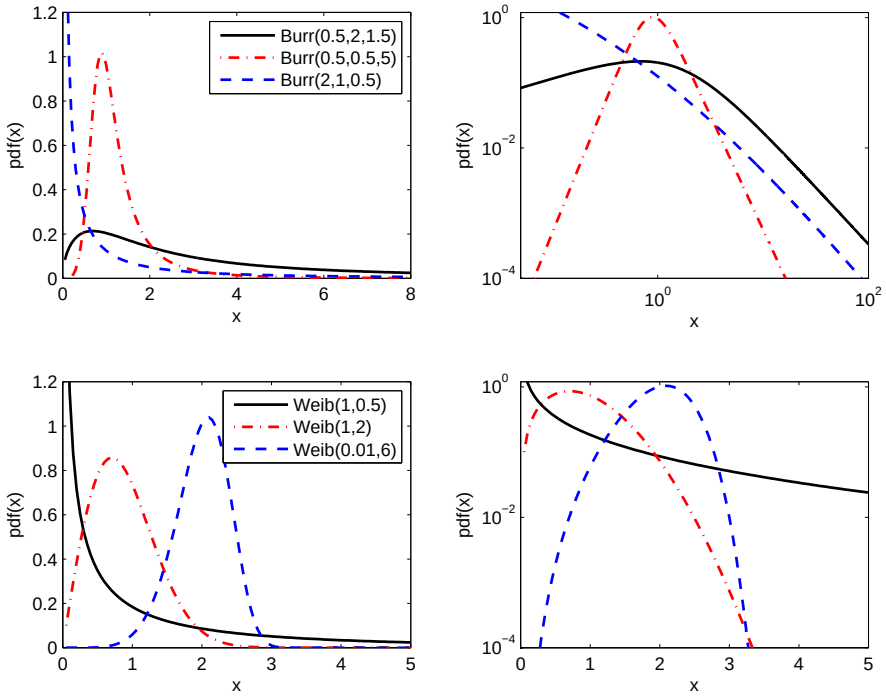


Figure 1.6: *Top panels:* Three sample Burr pdfs, Burr$(\alpha, \lambda, \tau)$, on linear and double-logarithmic plots. Note, the heavy, power-law tails. *Bottom panels:* Three sample Weibull pdfs, Weib$(\beta, \tau)$, on linear and semi-logarithmic plots. We can see that for $\tau < 1$ the tails are much heavier and they look like power-law.

STF2loss06.m

The maximum likelihood and method of moments estimators for the Burr distribution can only be evaluated numerically. A Burr variate $X$ can be generated using the inverse transform method. The inverse of the cdf (1.46) has a simple analytical form $F^{-1}(x) = \left[\lambda\left\{(1-x)^{-1/\alpha} - 1\right\}\right]^{1/\tau}$. Hence, we can set $X = \left\{\lambda\left(U^{-1/\alpha} - 1\right)\right\}^{1/\tau}$, where $U$ is distributed uniformly on the unit interval. Like in the Pareto case, we have to be cautious when $\tau\alpha$ is larger but very close to one. The sample mean will generally be significantly lower than $\mathrm{E}(X)$.

### 1.3.8 Weibull Distribution

If $V$ is an exponential variate, then the distribution of $X = V^{1/\tau}$, $\tau > 0$, is called the Weibull (or Frechet) distribution. Its density and distribution functions are given by:

$$f(x) = \tau\beta x^{\tau-1}e^{-\beta x^{\tau}}, \qquad x > 0, \tag{1.48}$$

$$F(x) = 1 - e^{-\beta x^{\tau}}, \qquad x > 0, \tag{1.49}$$

respectively. For $\tau = 2$ it is known as the Rayleigh distribution. The Weibull law is roughly symmetrical for the shape parameter $\tau \approx 3.6$. When $\tau$ is smaller the distribution is right-skewed, when $\tau$ is larger it is left-skewed, see the bottom panels in Figure 1.6. We can also observe that for $\tau < 1$ the distribution becomes heavy-tailed. In this case, like for the Pareto and Burr distributions, the moment generating function is infinite. The $k$-th raw moment can be shown to be

$$m_k = \beta^{-k/\tau}\Gamma\left(1 + \frac{k}{\tau}\right). \tag{1.50}$$

Like for the Burr distribution, the maximum likelihood and method of moments estimators can only be evaluated numerically. Similarly, Weibull variates can be generated using the inverse transform method.

## 1.4 Statistical Validation Techniques

Having a large collection of distributions to choose from we need to narrow our selection to a single model and a unique parameter estimate. The type of the objective loss distribution can be easily selected by comparing the shapes of the empirical and theoretical mean excess functions. The mean excess function, presented in Section 1.4.1, is based on the idea of conditioning a random variable given that it exceeds a certain level.

Once the distribution class is selected and the parameters are estimated using one of the available methods the goodness-of-fit has to be tested. A standard approach consists of measuring the distance between the empirical and the fitted analytical distribution function. A group of statistics and tests based on this idea is discussed in Section 1.4.2.

### 1.4.1 Mean Excess Function

For a claim amount random variable $X$, the mean excess function or mean residual life function is the expected payment per claim on a policy with a fixed amount deductible of $x$, where claims with amounts less than or equal to $x$ are completely ignored:

$$e(x) = \mathrm{E}(X - x | X > x) = \frac{\int_x^\infty \{1 - F(u)\}\, du}{1 - F(x)}. \tag{1.51}$$

In practice, the mean excess function $e$ is estimated by $\hat{e}_n$ based on a representative sample $x_1, \ldots, x_n$:

$$\hat{e}_n(x) = \frac{\sum_{x_i > x} x_i}{\#\{i : x_i > x\}} - x. \tag{1.52}$$

Note, that in a financial risk management context, switching from the right tail to the left tail, $e(x)$ is referred to as the expected shortfall (Weron, 2004).

When considering the shapes of mean excess functions, the exponential distribution plays a central role. It has the memoryless property, meaning that whether the information $X > x$ is given or not, the expected value of $X - x$ is the same as if one started at $x = 0$ and calculated $\mathrm{E}(X)$. The mean excess function for the exponential distribution is therefore constant. One in fact easily calculates that for this case $e(x) = 1/\beta$ for all $x > 0$.

If the distribution of $X$ is heavier-tailed than the exponential distribution we find that the mean excess function ultimately increases, when it is lighter-tailed $e(x)$ ultimately decreases. Hence, the shape of $e(x)$ provides important information on the sub-exponential or super-exponential nature of the tail of the distribution at hand.

Mean excess functions for the distributions discussed in Section 1.3 are given by the following formulas (and plotted in Figure 1.7):

- exponential:
$$e(x) = \frac{1}{\beta};$$
- mixture of two exponentials:
$$e(x) = \frac{\frac{a}{\beta_1}\exp(-\beta_1 x) + \frac{1-a}{\beta_2}\exp(-\beta_2 x)}{a\exp(-\beta_1 x) + (1-a)\exp(-\beta_2 x)};$$

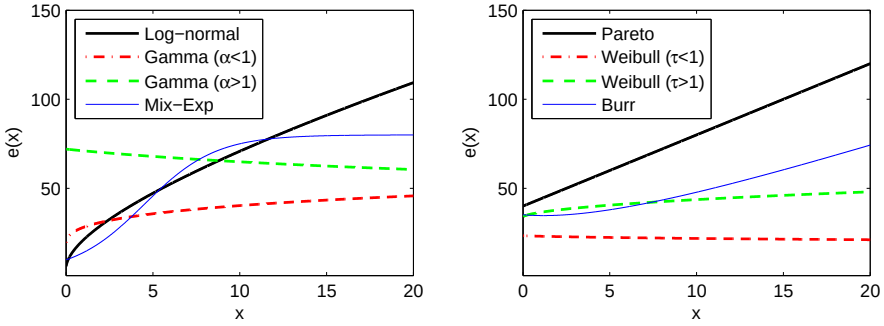


Figure 1.7: *Left panel:* Shapes of the mean excess function $e(x)$ for the log-normal, gamma (with $\alpha < 1$ and $\alpha > 1$) and a mixture of two exponential distributions. *Right panel:* Shapes of the mean excess function $e(x)$ for the Pareto, Weibull (with $\tau < 1$ and $\tau > 1$) and Burr distributions.

STF2loss07.m

- gamma:

$$e(x) \;=\; \frac{\alpha}{\beta}\cdot\frac{1-F\left(x,\alpha+1,\beta\right)}{1-F\left(x,\alpha,\beta\right)}-x \;=\; \beta^{-1}\left\{1+o(1)\right\},$$

  where $F(x,\alpha,\beta)$ is the gamma distribution function (1.19);

- log-normal:

$$\begin{aligned} e(x) \;&=\; \frac{\exp\left(\mu+\frac{\sigma^2}{2}\right)\left\{1-\Phi\left(\frac{\ln x-\mu-\sigma^2}{\sigma}\right)\right\}}{\left\{1-\Phi\left(\frac{\ln x-\mu}{\sigma}\right)\right\}}-x= \\ &=\; \frac{\sigma^2 x}{\ln x-\mu}\left\{1+o(1)\right\}, \end{aligned}$$

  where $o(1)$ stands for a term which tends to zero as $u\to\infty$;

- Pareto:

$$e(x)=\frac{\lambda+x}{\alpha-1}, \qquad \alpha>1;$$

- Burr:

$$\begin{aligned} e(x) &= \frac{\lambda^{1/\tau}\Gamma\left(\alpha-\frac{1}{\tau}\right)\Gamma\left(1+\frac{1}{\tau}\right)}{\Gamma(\alpha)} \cdot \left(\frac{\lambda}{\lambda+x^\tau}\right)^{-\alpha} \cdot \\ &\quad \cdot \left\{1-\mathrm{B}\left(1+\frac{1}{\tau},\alpha-\frac{1}{\tau},\frac{x^\tau}{\lambda+x^\tau}\right)\right\} - x = \\ &= \frac{x}{\alpha\tau-1}\left\{1+o(1)\right\}, \qquad \alpha\tau>1, \end{aligned}$$

  where $\Gamma(\cdot)$ is the standard gamma function (1.20) and

$$\mathrm{B}(a,b,x) \stackrel{\text{def}}{=} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\int_0^x y^{a-1}(1-y)^{b-1}dy, \tag{1.53}$$

  is the beta function;

- Weibull:

$$\begin{aligned} e(x) &= \frac{\Gamma\left(1+1/\tau\right)}{\beta^{1/\tau}}\left\{1-\Gamma\left(1+\frac{1}{\tau},\beta x^\tau\right)\right\}\exp\left(\beta x^\tau\right) - x = \\ &= \frac{x^{1-\tau}}{\beta\tau}\left\{1+o(1)\right\}, \end{aligned}$$

  where $\Gamma(\cdot,\cdot)$ is the incomplete gamma function (1.21).

### 1.4.2 Tests Based on the Empirical Distribution Function

A statistics measuring the difference between the empirical $F_n(x)$ and the fitted $F(x)$ distribution function, called an edf statistic, is based on the vertical difference between the distributions. This distance is usually measured either by a supremum or a quadratic norm (D'Agostino and Stephens, 1986).

The most popular supremum statistic:

$$D = \sup_x \left|F_n(x) - F(x)\right|, \tag{1.54}$$

is known as the Kolmogorov or Kolmogorov-Smirnov statistic. It can also be written in terms of two supremum statistics:

$$D^+ = \sup_x \left\{F_n(x) - F(x)\right\} \quad \text{and} \quad D^- = \sup_x \left\{F(x) - F_n(x)\right\},$$

where the former is the largest vertical difference when $F_n(x)$ is larger than $F(x)$ and the latter is the largest vertical difference when it is smaller. The Kolmogorov statistic is then given by $D = \max(D^+, D^-)$. A closely related statistic proposed by Kuiper is simply a sum of the two differences, i.e. $V = D^+ + D^-$.

The second class of measures of discrepancy is given by the Cramer-von Mises family

$$Q = n \int_{-\infty}^{\infty} \{F_n(x) - F(x)\}^2 \psi(x) \mathrm{d}F(x), \tag{1.55}$$

where $\psi(x)$ is a suitable function which gives weights to the squared difference $\{F_n(x) - F(x)\}^2$. When $\psi(x) = 1$ we obtain the $W^2$ statistic of Cramer-von Mises. When $\psi(x) = [F(x)\{1 - F(x)\}]^{-1}$ formula (1.55) yields the $A^2$ statistic of Anderson and Darling. From the definitions of the statistics given above, suitable computing formulas must be found. This can be done by utilizing the transformation $Z = F(X)$. When $F(x)$ is the true distribution function of $X$, the random variable $Z$ is uniformly distributed on the unit interval.

Suppose that a sample $x_1, \ldots, x_n$ gives values $z_i = F(x_i),\ i = 1, \ldots, n$. It can be easily shown that, for values $z$ and $x$ related by $z = F(x)$, the corresponding vertical differences in the edf diagrams for $X$ and for $Z$ are equal. Consequently, edf statistics calculated from the empirical distribution function of the $z_i$'s compared with the uniform distribution will take the same values as if they were calculated from the empirical distribution function of the $x_i$'s, compared with $F(x)$. This leads to the following formulas given in terms of the order statistics $z_{(1)} < z_{(2)} < \cdots < z_{(n)}$:

$$D^+ = \max_{1 \le i \le n} \left\{ \frac{i}{n} - z_{(i)} \right\}, \tag{1.56}$$

$$D^- = \max_{1 \le i \le n} \left\{ z_{(i)} - \frac{(i-1)}{n} \right\}, \tag{1.57}$$

$$D = \max(D^+, D^-), \tag{1.58}$$

$$V = D^+ + D^-, \tag{1.59}$$

$$W^2 = \sum_{i=1}^{n} \left\{ z_{(i)} - \frac{(2i-1)}{2n} \right\}^2 + \frac{1}{12n}, \tag{1.60}$$

$$A^2 = -n - \frac{1}{n}(2i-1)\sum_{i=1}^{n}\left\{\log z_{(i)} + \log(1 - z_{(n+1-i)})\right\} = \quad (1.61)$$

$$= -n - \frac{1}{n}\sum_{i=1}^{n}\left\{(2i-1)\log z_{(i)} + \right.$$
$$\left. +(2n+1-2i)\log(1 - z_{(i)})\right\}. \quad (1.62)$$

The general test of fit is structured as follows. The null hypothesis is that a specific distribution is acceptable, whereas the alternative is that it is not:

$$H_0: \quad F_n(x) = F(x;\theta),$$
$$H_1: \quad F_n(x) \neq F(x;\theta),$$

where $\theta$ is a vector of known parameters. Small values of the test statistic $T$ are evidence in favor of the null hypothesis, large ones indicate its falsity. To see how unlikely such a large outcome would be if the null hypothesis was true, we calculate the $p$-value by:

$$p\text{-value} = P(T \geq t), \quad (1.63)$$

where $t$ is the test value for a given sample. It is typical to reject the null hypothesis when a small $p$-value is obtained.

However, we are in a situation where we want to test the hypothesis that the sample has a common distribution function $F(x;\theta)$ with unknown $\theta$. To employ any of the edf tests we first need to estimate the parameters. It is important to recognize that when the parameters are estimated from the data, the critical values for the tests of the uniform distribution (or equivalently of a fully specified distribution) must be reduced. In other words, if the value of the test statistics $T$ is $d$, then the $p$-value is overestimated by $P_U(T \geq d)$. Here $P_U$ indicates that the probability is computed under the assumption of a uniformly distributed sample. Hence, if $P_U(T \geq d)$ is small, then the $p$-value will be even smaller and the hypothesis will be rejected. However, if it is large then we have to obtain a more accurate estimate of the $p$-value.

Ross (2002) advocates the use of Monte Carlo simulations in this context. First the parameter vector is estimated for a given sample of size $n$, yielding $\hat{\theta}$, and the edf test statistics is calculated assuming that the sample is distributed according to $F(x;\hat{\theta})$, returning a value of $d$. Next, a sample of size $n$ of $F(x;\hat{\theta})$-distributed variates is generated. The parameter vector is estimated for this simulated sample, yielding $\hat{\theta}_1$, and the edf test statistics is calculated assuming

that the sample is distributed according to $F(x; \hat{\theta}_1)$. The simulation is repeated as many times as required to achieve a certain level of accuracy. The estimate of the $p$-value is obtained as the proportion of times that the test quantity is at least as large as $d$.

An alternative solution to the problem of unknown parameters was proposed by Stephens (1978). The half-sample approach consists of using only half the data to estimate the parameters, but then using the entire data set to conduct the test. In this case, the critical values for the uniform distribution can be applied, at least asymptotically. The quadratic edf tests seem to converge fairly rapidly to their asymptotic distributions (D'Agostino and Stephens, 1986). Although, the method is much faster than the Monte Carlo approach it is not invariant – depending on the choice of the half-samples different test values will be obtained and there is no way of increasing the accuracy.

As a side product, the edf tests supply us with a natural technique of estimating the parameter vector $\theta$. We can simply find such $\hat{\theta}^*$ that minimizes a selected edf statistic. Out of the four presented statistics $A^2$ is the most powerful when the fitted distribution departs from the true distribution in the tails (D'Agostino and Stephens, 1986). Since the fit in the tails is of crucial importance in most actuarial applications $A^2$ is the recommended statistic for the estimation scheme.

## 1.5 Applications

In this section we illustrate some of the methods described earlier in the chapter. We conduct the analysis for the Danish fire losses dataset, which concerns major fire losses in Danish Krone (DKK) that occurred between 1980 and 2002 and were recorded by Copenhagen Re. Here we consider only losses in profits. The Danish fire losses dataset has been adjusted for inflation using the Danish consumer price index.

### 1.5.1 Calibration of Loss Distributions

We first look for the appropriate shape of the distribution. To this end we plot the empirical mean excess function for the analyzed data set, see Figure 1.8. Since the function represents an empirical mean above some threshold level, its values for high $x$'s are not credible, and we do not include them in the plot.

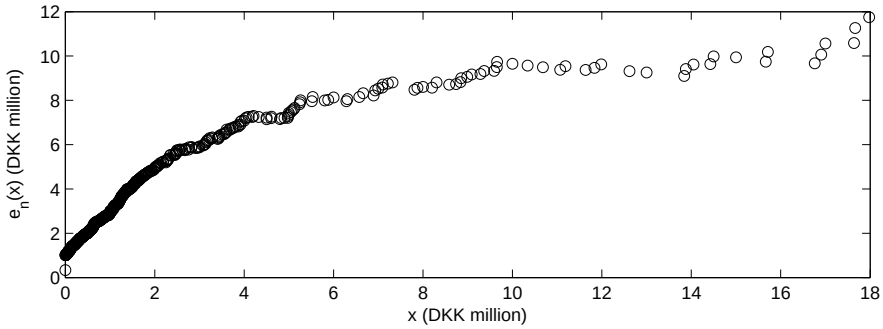


Figure 1.8: The empirical mean excess function $\hat{e}_n(x)$ for the Danish fire data.
STF2loss08.m

Before we continue with calibration let us note, that in recent years outlier-resistant or so-called robust estimates of parameters are becoming more widespread in risk modeling. Such models – called robust (statistics) models – were introduced by P.J. Huber in 1981 and applied to robust regression analysis (Huber, 2004). Under the robust approach the extreme data points are eliminated to avoid a situation when outliers drive future forecasts in an unwanted (such as worst-case scenario) direction. One of the first applications of robust analysis to insurance claim data can be found in Chernobai et al. (2006). In that paper top 1% of the catastrophic losses were treated as outliers and excluded from the analysis. This procedure led to an improvement of the forecasting power of considered models. Also the resulting ruin probabilities were more optimistic than those predicted by the classical model. It is important to note, however, that neither of the two approaches – classical or robust – is preffered over the other. Rather, in the presence of outliers, the robust model can be used to complement to the classical one. Due to space limits, in this chapter we will only present the results of the latter. The robust approach can be easily conducted following the steps detailed in this Section.

The Danish fire losses show a super-exponential pattern suggesting a log-normal, Pareto or Burr distribution as the most adequate for modeling. Hence, in what follows we fit only these three distributions. We apply two estimation schemes: maximum likelihood and $A^2$ statistics minimization. Out of the three fitted distributions only the log-normal has closed form expressions for

Table 1.1: Parameter estimates obtained via the $A^2$ minimization scheme and test statistics for the fire loss amounts. The corresponding $p$-values based on 1000 simulated samples are given in parentheses.

| **Distributions:** | | log-normal | Pareto | Burr |
|---|---|---|---|---|
| **Parameters:** | | $\mu$=12.525 | $\alpha$=1.3127 | $\alpha$=0.9844 |
| | | $\sigma$=1.5384 | $\lambda$=$4.0588 \cdot 10^5$ | $\lambda$=$1.0585 \cdot 10^6$ |
| | | | | $\tau$=1.1096 |
| **Tests:** | $D$ | 0.0180 | 0.0262 | 0.0266 |
| | | (0.020) | (<0.005) | (<0.005) |
| | $V$ | 0.0326 | 0.0516 | 0.0496 |
| | | (0.012) | (<0.005) | (<0.005) |
| | $W^2$ | 0.0932 | 0.2322 | 0.2316 |
| | | (0.068) | (<0.005) | (<0.005) |
| | $A^2$ | 0.9851 | 2.6748 | 1.8894 |
| | | (0.005) | (<0.005) | (<0.005) |

STF2loss08t.m

the maximum likelihood estimators. Parameter calibration for the remaining distributions and the $A^2$ minimization scheme is carried out via a simplex numerical optimization routine. A limited simulation study suggests that the $A^2$ minimization scheme tends to return lower values of all edf test statistics than maximum likelihood estimation. Hence, it is exclusively used for further analysis.

The results of parameter estimation and hypothesis testing for the Danish fire loss amounts are presented in Table 1.1. The log-normal distribution with parameters $\mu = 12.525$ and $\sigma = 1.5384$ returns the best results. It is the only distribution that passes any of the four applied tests ($D$, $V$, $W^2$, and $A^2$) at a reasonable level. The Burr and Pareto laws yield worse fits as the tails of the edf are lighter than power-law tails. As expected, the remaining distributions return even worse fits. Hence, we suggest to use the log-normal distribution as a model for the Danish fire loss amounts.

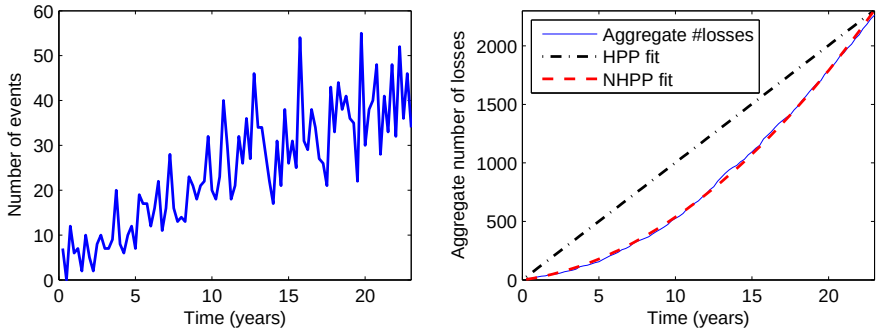


Figure 1.9: *Left panel*: The quarterly number of losses for the Danish fire data. *Right panel*: The aggregate number of losses and the mean value function $\mathrm{E}(N_t)$ of the calibrated HPP and NHPP. Clearly the latter model gives a better fit to the data.

STF2loss09.m

### 1.5.2 Simulation of Risk Processes

We conduct empirical studies for Danish fire losses recorded by Copenhagen Re. The data concerns major Danish fire losses in Danish Krone (DKK), occurred between 1980 and 2002 and adjusted for inflation. Only losses of profits connected with the fires are taken into consideration. We start the analysis with a HPP with a constant intensity $\lambda_1$. Studies of the quarterly numbers of losses and the inter-occurrence times of the fires lead us to the conclusion that the annual intensity of $\lambda_1 = 98.39$ gives the best fitted HPP. However, as we can see in the right panel of Figure 1.9, the fit is not very good suggesting that the HPP is too simplistic. A renewal process would also give unsatisfactory results as the data reveals a clear increasing trend in the number of quarterly losses, see the left panel in Figure 1.9. This leaves us with the NHPP. We tested different exponential and polynomial functional forms, but a simple linear intensity function $\lambda_2(s) = c + ds$ gives the best fit. Applying the least squares procedure we arrive at the following values of the parameters: $c = 17.99$ and $d = 7.15$. Processes with both choices of the intensity function, $\lambda_1$ and $\lambda_2(s)$, are illustrated in the right panel of Figure 1.9, where the accumulated number of fire losses and mean value functions for all 23 years of data are depicted.

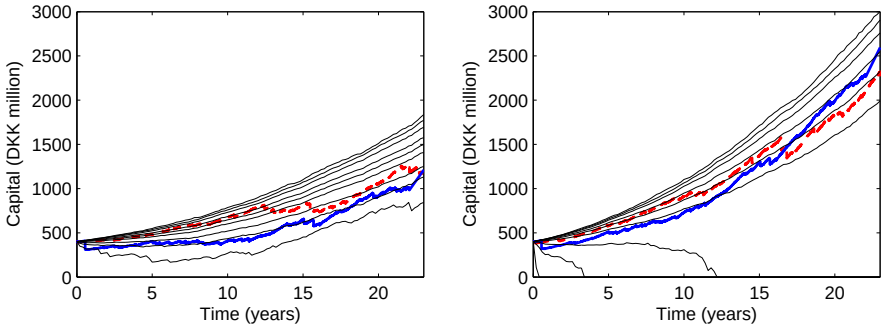


Figure 1.10: The Danish fire data simulation results for a NHPP with log-normal claim sizes (*left panel*) and a NHPP with Burr claim sizes (*right panel*). The dotted lines are the sample 0.001, 0.01, 0.05, 0.25, 0.50, 0.75, 0.95, 0.99, 0.999-quantile lines based on 3000 trajectories of the risk process.

STF2loss10.m

After describing the claim arrival process we have to find an appropriate model for the loss amounts. In Section 1.5.1 a number of distributions were fitted to loss sizes. The log-normal distribution with parameters $\mu = 12.525$ and $\sigma = 1.5384$ produced the best results. The Burr distribution with $\alpha = 0.9844$, $\lambda = 1.0585 \cdot 10^6$, and $\tau = 1.1096$ overestimated the tails of the empirical distribution, nevertheless it gave the next best fit.

The simulation results are presented in Figure 1.10. We consider a hypothetical scenario where the insurance company insures losses resulting from fire damage. The company's initial capital is assumed to be $u = 400$ million DKK and the relative safety loading used is $\theta = 0.5$. We choose two models of the risk process whose application is most justified by the statistical results described above: a NHPP with log-normal claim sizes and a NHPP with Burr claim sizes. In both panels the thick solid blue line is the "real" risk process, i.e. a trajectory constructed from the historical arrival times and values of the losses. The different shapes of the "real" risk process in the two panels are due to the different forms of the premium function $c(t)$ which has to be chosen accordingly to the type of the claim arrival process. The dashed red line is a sample trajectory. The thin solid lines are the sample 0.001, 0.01, 0.05, 0.25,

0.50, 0.75, 0.95, 0.99, 0.999-quantile lines based on 3000 trajectories of the risk process. We assume that if the capital of the insurance company drops bellow zero, the company goes bankrupt, so the capital is set to zero and remains at this level hereafter.

Comparing the log-normal and Burr claim size models, we can conclude that in the latter model extreme events are more likely to happen. This is manifested by wider quantile lines in the right panel of Figure 1.10. Since for log-normal claim sizes the historical trajectory is above the 0.01-quantile line for most of the time, and taking into account that we have followed a non-robust estimation approach of loss severities, we suggest to use this specification for further risk process modeling using the 1980-2002 Danish fire losses dataset.

# Bibliography


Ahrens, J. H. and Dieter, U. (1982). Computer generation of Poisson deviates from modified normal distributions, *ACM Trans. Math. Soft.* **8**: 163–179.

Albrecht, P. (1982). On some statistical methods connected with the mixed Poisson process, *Scandinavian Actuarial Journal*: 1–14.

Bratley, P., Fox, B. L., and Schrage, L. E. (1987). *A Guide to Simulation*, Springer-Verlag, New York.

Burnecki, K., Härdle, W., and Weron, R. (2004). Simulation of Risk Processes, *in* J. Teugels, B. Sundt (eds.) *Encyclopedia of Actuarial Science*, Wiley, Chichester, 1564–1570.

Burnecki, K., Kukla, G., and Weron, R. (2000). Property insurance loss distributions, *Physica A* **287**: 269-278.

K. Burnecki and R. Weron (2005). Modeling of the Risk Process, *in* P. Cizek, W. Härdle, R. Weron (eds.) *Statistical Tools for Finance and Insurance*, Springer-Verlag, Berlin, 319–339.

Chernobai, A., Burnecki, K., Rachev, S. T., Trueck, S. and Weron, R. (2006). Modelling catastrophe claims with left-truncated severity distributions, *Computational Statistics* **21**: 537-555.

Chernobai, A.S., Rachev, S.T., and Fabozzi, F.J. (2007). *Operational Risk: A Guide to Basel II Capital Requirements, Models, and Analysis*, Wiley.

D'Agostino, R. B. and Stephens, M. A. (1986). *Goodness-of-Fit Techniques*, Marcel Dekker, New York.

Daykin, C.D., Pentikainen, T., and Pesonen, M. (1994). *Practical Risk Theory for Actuaries*, Chapman, London.

Devroye, L. (1986). *Non-Uniform Random Variate Generation*, Springer-Verlag, New York.

Embrechts, P. and Klüppelberg, C. (1993). Some aspects of insurance mathematics, *Theory Probab. Appl.* **38**: 262–295.

Grandell, J. (1991). *Aspects of Risk Theory*, Springer, New York.

Hogg, R. and Klugman, S. A. (1984). *Loss Distributions*, Wiley, New York.

Hörmann, W. (1993). The transformed rejection method for generating Poisson random variables, *Insurance: Mathematics and Economics* **12**: 39–45.

Huber, P. J. (2004). *Robust statistics*, Wiley, Hoboken.

Jöhnk, M. D. (1964). Erzeugung von Betaverteilten und Gammaverteilten Zufallszahlen, *Metrika* **8**: 5-15.

Kaas, R., Goovaerts, M., Dhaene, J., and Denuit, M. (2008). *Modern Actuarial Risk Theory: Using R*, Springer.

Klugman, S. A., Panjer, H.H., and Willmot, G.E. (2008). *Loss Models: From Data to Decisions* (3rd ed.), Wiley, New York.

Panjer, H.H. (2006). *Operational Risk : Modeling Analytics*, Wiley.

Panjer, H.H. and Willmot, G.E. (1992). *Insurance Risk Models*, Society of Actuaries, Chicago.

Rolski, T., Schmidli, H., Schmidt, V., and Teugels, J. L. (1999). *Stochastic Processes for Insurance and Finance*, Wiley, Chichester.

Ross, S. (2002). *Simulation*, Academic Press, San Diego.

Stadlober, E. (1989). Sampling from Poisson, binomial and hypergeometric distributions: ratio of uniforms as a simple and fast alternative, *Math. Statist. Sektion* **303**, Forschungsgesellschaft Joanneum Graz.

Stephens, M. A. (1978). On the half-sample method for goodness-of-fit, *Journal of the Royal Statistical Society B* **40**: 64-70.

Teugels, J.L. and Vynckier, P. (1996). The structure distribution in a mixed Poisson process, *J. Appl. Math. & Stoch. Anal.* **9**: 489–496.

Tse, Y.-K. (2009). *Nonlife Actuarial Models: Theory, Methods and Evaluation*, Cambridge University Press, Cambridge.

Weron, R. (2004). Computationally Intensive Value at Risk Calculations, *in* J. E. Gentle, W. Härdle, Y. Mori (eds.) *Handbook of Computational Statistics*, Springer, Berlin, 911–950.

Willmot, G. E. (2001). The nature of modelling insurance losses, *The Munich Re Inaugural Lecture*, December 5, 2001, Toronto.

# SFB 649 Discussion Paper Series 2010

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Volatility Investing with Variance Swaps" by Wolfgang Karl Härdle and Elena Silyakova, January 2010.
002 "Partial Linear Quantile Regression and Bootstrap Confidence Bands" by Wolfgang Karl Härdle, Ya'acov Ritov and Song Song, January 2010.
003 "Uniform confidence bands for pricing kernels" by Wolfgang Karl Härdle, Yarema Okhrin and Weining Wang, January 2010.
004 "Bayesian Inference in a Stochastic Volatility Nelson-Siegel Model" by Nikolaus Hautsch and Fuyu Yang, January 2010.
005 "The Impact of Macroeconomic News on Quote Adjustments, Noise, and Informational Volatility" by Nikolaus Hautsch, Dieter Hess and David Veredas, January 2010.
006 "Bayesian Estimation and Model Selection in the Generalised Stochastic Unit Root Model" by Fuyu Yang and Roberto Leon-Gonzalez, January 2010.
007 "Two-sided Certification: The market for Rating Agencies" by Erik R. Fasten and Dirk Hofmann, January 2010.
008 "Characterising Equilibrium Selection in Global Games with Strategic Complementarities" by Christian Basteck, Tijmen R. Daniels and Frank Heinemann, January 2010.
009 "Predicting extreme VaR: Nonparametric quantile regression with refinements from extreme value theory" by Julia Schaumburg, February 2010.
010 "On Securitization, Market Completion and Equilibrium Risk Transfer" by Ulrich Horst, Traian A. Pirvu and Gonçalo Dos Reis, February 2010.
011 "Illiquidity and Derivative Valuation" by Ulrich Horst and Felix Naujokat, February 2010.
012 "Dynamic Systems of Social Interactions" by Ulrich Horst, February 2010.
013 "The dynamics of hourly electricity prices" by Wolfgang Karl Härdle and Stefan Trück, February 2010.
014 "Crisis? What Crisis? Currency vs. Banking in the Financial Crisis of 1931" by Albrecht Ritschl and Samad Sarferaz, February 2010.
015 "Estimation of the characteristics of a Lévy process observed at arbitrary frequency" by Johanna Kappusl and Markus Reiß, February 2010.
016 "Honey, I'll Be Working Late Tonight. The Effect of Individual Work Routines on Leisure Time Synchronization of Couples" by Juliane Scheffel, February 2010.
017 "The Impact of ICT Investments on the Relative Demand for High-Medium-, and Low-Skilled Workers: Industry versus Country Analysis" by Dorothee Schneider, February 2010.
018 "Time varying Hierarchical Archimedean Copulae" by Wolfgang Karl Härdle, Ostap Okhrin and Yarema Okhrin, February 2010.
019 "Monetary Transmission Right from the Start: The (Dis)Connection Between the Money Market and the ECB's Main Refinancing Rates" by Puriya Abbassi and Dieter Nautz, March 2010.
020 "Aggregate Hazard Function in Price-Setting: A Bayesian Analysis Using Macro Data" by Fang Yao, March 2010.
021 "Nonparametric Estimation of Risk-Neutral Densities" by Maria Grith, Wolfgang Karl Härdle and Melanie Schienle, March 2010.

# SFB 649 Discussion Paper Series 2010

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

022 "Fitting high-dimensional Copulae to Data" by Ostap Okhrin, April 2010.
023 "The (In)stability of Money Demand in the Euro Area: Lessons from a Cross-Country Analysis" by Dieter Nautz and Ulrike Rondorf, April 2010.
024 "The optimal industry structure in a vertically related market" by Raffaele Fiocco, April 2010.
025 "Herding of Institutional Traders" by Stephanie Kremer, April 2010.
026 "Non-Gaussian Component Analysis: New Ideas, New Proofs, New Applications" by Vladimir Panov, May 2010.
027 "Liquidity and Capital Requirements and the Probability of Bank Failure" by Philipp Johann König, May 2010.
028 "Social Relationships and Trust" by Christine Binzel and Dietmar Fehr, May 2010.
029 "Adaptive Interest Rate Modelling" by Mengmeng Guo and Wolfgang Karl Härdle, May 2010.
030 "Can the New Keynesian Phillips Curve Explain Inflation Gap Persistence?" by Fang Yao, June 2010.
031 "Modeling Asset Prices" by James E. Gentle and Wolfgang Karl Härdle, June 2010.
032 "Learning Machines Supporting Bankruptcy Prediction" by Wolfgang Karl Härdle, Rouslan Moro and Linda Hoffmann, June 2010.
033 "Sensitivity of risk measures with respect to the normal approximation of total claim distributions" by Volker Krätschmer and Henryk Zähle, June 2010.
034 "Sociodemographic, Economic, and Psychological Drivers of the Demand for Life Insurance: Evidence from the German Retirement Income Act" by Carolin Hecht and Katja Hanewald, July 2010.
035 "Efficiency and Equilibria in Games of Optimal Derivative Design" by Ulrich Horst and Santiago Moreno-Bromberg, July 2010.
036 "Why Do Financial Market Experts Misperceive Future Monetary Policy Decisions?" by Sandra Schmidt and Dieter Nautz, July 2010.
037 "Dynamical systems forced by shot noise as a new paradigm in the interest rate modeling" by Alexander L. Baranovski, July 2010.
038 "Pre-Averaging Based Estimation of Quadratic Variation in the Presence of Noise and Jumps: Theory, Implementation, and Empirical Evidence" by Nikolaus Hautsch and Mark Podolskij, July 2010.
039 "High Dimensional Nonstationary Time Series Modelling with Generalized Dynamic Semiparametric Factor Model" by Song Song, Wolfgang K. Härdle, and Ya'acov Ritov, July 2010.
040 "Stochastic Mortality, Subjective Survival Expectations, and Individual Saving Behavior" by Thomas Post and Katja Hanewald, July 2010.
041 "Prognose mit nichtparametrischen Verfahren" by Wolfgang Karl Härdle, Rainer Schulz, and Weining Wang, August 2010.
042 "Payroll Taxes, Social Insurance and Business Cycles" by Michael C. Burda and Mark Weder, August 2010.
043 "Meteorological forecasts and the pricing of weather derivatives" by Matthias Ritter, Oliver Mußhoff, and Martin Odening, September 2010.
044 "The High Sensitivity of Employment to Agency Costs: The Relevance of Wage Rigidity" by Atanas Hristov, September 2010.
045 "Parametric estimation of risk neutral density functions" by Maria Grith and Volker Krätschmer, September 2010.

# SFB 649 Discussion Paper Series 2010

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

046 "Mandatory IFRS adoption and accounting comparability" by Stefano Cascino and Joachim Gassen, October 2010.
047 "FX Smile in the Heston Model" by Agnieszka Janek, Tino Kluge, Rafał Weron, and Uwe Wystup, October 2010.
048 "Building Loss Models" by Krzysztof Burnecki, Joanna Janczura, and Rafał Weron, October 2010.