MPRA
Munich Personal RePEc Archive

Simulation of Risk Processes

Burnecki, Krzysztof and Weron, Rafal

Wrocław University of Technology, Poland

2010

Online at https://mpra.ub.uni-muenchen.de/25444/
MPRA Paper No. 25444, posted 27 Sep 2010 03:22 UTC

Simulation of Risk Processes[1]

Krzysztof Burnecki[a], Rafał Weron[b]

[a]Hugo Steinhaus Center, Wrocław University of Technology, 50-370 Wrocław, Poland
[b]Institute of Organization and Management, Wrocław University of Technology, 50-370 Wrocław, Poland

Abstract

This paper is intended as a guide to simulation of risk processes. 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. 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. In this paper we present efficient simulation algorithms for several classes of claim arrival processes.

Keywords: Risk process, Claim arrival process, Homogeneous Poisson process (HPP), Non-homogeneous Poisson process (NHPP), Mixed Poisson process, Cox process, Renewal process.

1. Introduction

A loss model or actuarial risk model 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.

In the collective risk model the aggregate loss has 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. Apart from health insurance, the collective risk model can also be used for modeling 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.

In classical form the collective risk model 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

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

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 *aggregate*

[1]This is a revised version of a paper originally published as Chapter 14 *Modeling of the Risk Process* in Cižek, Härdle, and Weron (2005). Among other changes, the revision includes new figures (thanks go to Joanna Janczura) and simulation algorithms. Also the empirical study of Section 3 concerns a twice longer Danish fire losses dataset. This revised text will form the backbone of the Chapter *Building Loss Models* in Cižek, Härdle, and Weron (2011).

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 simplicity of the risk process defined in eqn. (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 15 in Cižek, Härdle, and Weron (2005). 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).

In Section 2 we present efficient simulation algorithms for five classes of the claim arrival process $\{N_t\}$. For a concise treatment of the modeling of claim severities and testing the goodness-of-fit we refer to Chapter 13 of Cižek, Härdle, and Weron (2005). In Section 3 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.

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 ith 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. 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 five examples of $\{N_t\}$: the classical (homogeneous) Poisson process, the non-homogeneous Poisson process, the mixed Poisson process, the Cox process and the renewal process.

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 λ, i.e. with mean $1/\lambda$. 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 λ

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 Poisson processes are plotted in the top panels of Figure 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 λt

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 $(T_1, T_2, \ldots, T_N) = t \cdot \mathrm{sort}\{U_1, U_2, \ldots, U_N\}$

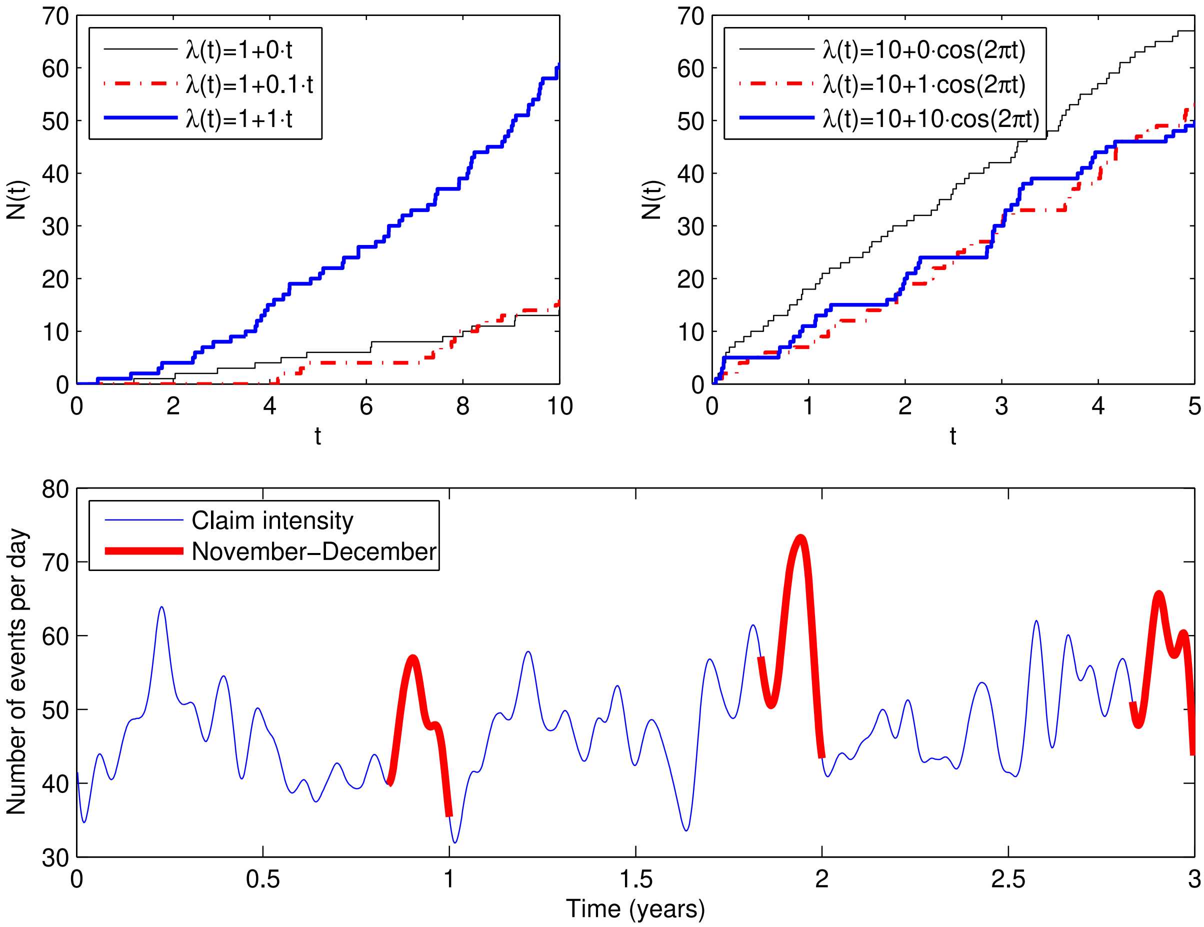

Figure 1: *Top left panel*: Sample trajectories of a NHPP with linear intensity $\lambda(t) = a + b \cdot t$. Note that the first process (with $b = 0$) is in fact a HPP. *Top right panel*: Sample trajectories of a NHPP with periodic intensity $\lambda(t) = a + b \cdot \cos(2\pi t)$. Again, the first 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.

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

$$N = \min\{n : U_1 \cdot \ldots \cdot U_n < \exp(-\lambda)\} - 1, \tag{2}$$

which is a consequence of the properties of the HPP (see above). However, for large λ, this method can become slow as the expected run time is proportional to λ. 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 λ'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 (2) for $\lambda < 15$ and Ahrens' and Dieter's method for larger values of λ.

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:

$$R_t = u + (1+\theta)\mu\lambda t - \sum_{i=1}^{N_t} X_i. \quad (3)$$

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) 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 λ.

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 ith 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. 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) than the HPP.

Finally, we note that since in the non-homogeneous case the expected value of the process at time t is $E(N_t) = \int_0^t \lambda(s)ds$, 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:

$$R_t = u + (1 + \theta)\mu \int_0^t \lambda(s)ds - \sum_{i=1}^{N_t} X_i. \tag{4}$$

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 $\text{Var}(N_t)/E(N_t)$, 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 Λ (called a structure variable), the random variable $\{N_t\}$ has a Poisson distribution. Typical examples for Λ 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

Λt, 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:

$$R_t = u + (1+\theta)\mu\Lambda t - \sum_{i=1}^{N_t} X_i. \tag{5}$$

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 Λ is generated and, conditioned upon its realization, N_t is simulated according to the Poisson law with parameter Λt. 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 Λt

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 $(T_1, T_2, \ldots, T_N) = t \cdot \mathrm{sort}\{U_1, U_2, \ldots, U_N\}$

2.4. Cox process

The Cox process, or doubly stochastic Poisson process, provides flexibility by letting the intensity not only depend on time but also by allowing it to be a stochastic process. Therefore, the doubly stochastic Poisson process can be viewed as a two-step randomization procedure. An intensity process $\{\Lambda(t)\}$ is used to generate another process $\{N_t\}$ by acting as its intensity. That is, $\{N_t\}$ is a Poisson process conditional on $\{\Lambda(t)\}$ which itself is a stochastic process. If $\{\Lambda(t)\}$ is deterministic, then $\{N_t\}$ is a non-homogeneous Poisson process. If $\Lambda(t) = \Lambda$ for some positive random variable Λ, then $\{N_t\}$ is a mixed Poisson process. In the doubly stochastic case the premium function is a generalization of the former functions, in line with the generalization of the claim arrival process. Hence, it takes the form $c(t) = (1+\theta)\mu \int_0^t \Lambda(s)ds$.

The definition of the Cox process suggests that it can be generated in the following way: first a realization of a non-negative stochastic process $\{\Lambda(t)\}$ is generated and, conditioned upon its realization, $\{N_t\}$ as a non-homogeneous Poisson process with that realization as its intensity is constructed. Out of the three methods of generating a non-homogeneous Poisson process the NHPP1 algorithm is the most general and, hence, the most suitable for adaptation. We can write:

Algorithm CP1 (Thinning)

Step 0: generate a realization $\lambda(t)$ of the intensity process $\{\Lambda(t)\}$ for a sufficiently large time period; set $\overline{\lambda} = \max\{\lambda(t)\}$

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

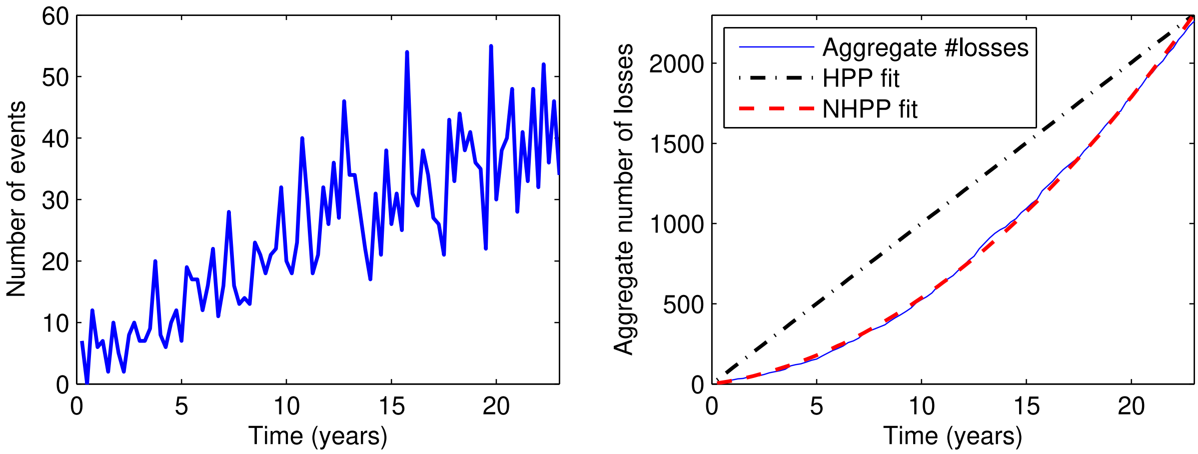

Figure 2: *Left panel*: The quarterly number of losses for the Danish fire data. *Right panel*: The aggregate number of losses and the mean value function $\mathrm{E}(N_t)$ of the calibrated HPP and NHPP. Clearly the latter model gives a better fit to the data.

2.5. Renewal Process

Generalizing the homogeneous Poisson process we come to the point where instead of making λ 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 (3).

3. Applications

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 λ_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 2, 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 2. This leaves us with the NHPP. We tested different exponential and polynomial functional forms, but a simple linear

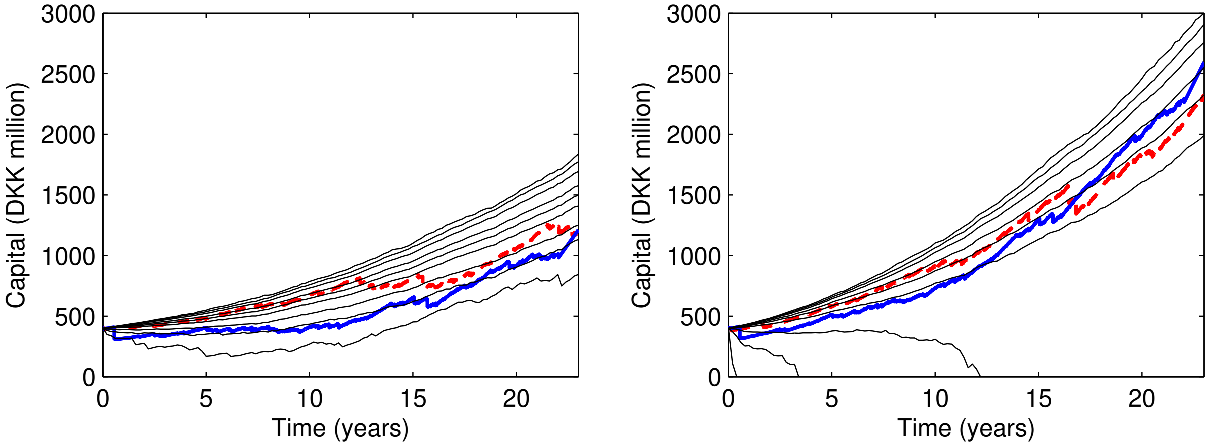

Figure 3: The Danish fire data simulation results for a NHPP with log-normal (*left panel*) and Burr claim sizes (*right panel*). The dotted lines are the sample 0.001, 0.01, 0.05, 0.25, 0.50, 0.75, 0.95, 0.99, 0.999-quantile lines based on 3000 trajectories of the risk process.

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, λ_1 and $\lambda_2(s)$, are illustrated in the right panel of Figure 2, where the accumulated number of fire losses and mean value functions for all 23 years of data are depicted.

After describing the claim arrival process we have to find an appropriate model for the loss amounts. Out of the typical loss distributions (Hogg and Klugman, 1984), the log-normal law 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. It is interesting to note, that in recent years outlier-resistant or so-called robust estimates of parameters are becoming more wide-spread 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 paper we only present the results of the latter.

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 of Figure 3 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 extreme events are more likely to happen. This is manifested by wider quantile lines. Since in the log-normal case 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, we suggest to use this specification for risk process modeling.

References

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.

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.

Cižek, P., Härdle, W., and Weron, R., eds. (2005). *Statistical Tools for Finance and Insurance*, Springer-Verlag, Berlin.

Cižek, P., Härdle, W., and Weron, R., eds. (2011). *Statistical Tools for Finance and Insurance*, 2nd ed., Springer-Verlag, Berlin.

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.

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.

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.

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.

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