HSC Research Report

HSC/03/04

An introduction to simulation of risk processes

Krzysztof Burnecki*
Wolfgang Härdle**
Rafał Weron*

* Hugo Steinhaus Center, Wrocław University of Technology, Poland
** CASE – Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, Germany

Hugo Steinhaus Center
Wrocław University of Technology
Wyb. Wyspiańskiego 27, 50-370 Wrocław, Poland
http://www.im.pwr.wroc.pl/~hugo/

An Introduction to Simulation of Risk Processes

Krzysztof Burnecki[1], Wolfgang Härdle[2] and Rafał Weron[1]

1 Introduction

The standard model for insurance risk is defined as follows [8, 9]. If $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space carrying (i) a *point process* $\{N_t\}_{t\geq 0}$, i.e. an integer valued stochastic process with $N_0 = 0$ a.s., $N_t < \infty$ for each $t < \infty$ and nondecreasing realizations and (ii) an independent sequence $\{X_k\}_{k=1}^{\infty}$ of positive i.i.d. random variables with common mean μ, 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 nonnegative constant u stands for the *initial capital* of the insurance company. To cover its liabilities, the company sells insurance policies and receives a premium according to a *premium function* $c(t)$. Liabilities result from claims covered by the previously sold insurance policies and are represented in the above formula by the *aggregated claim process* $\sum_{i=1}^{N_t} X_i$. The claim severities are described by the random sequence X_k and the number of claims in the interval $(0, t]$ is modeled by the point process N_t, often called the *claim arrival process*.

The simulation of the aggregated claim process reduces to modeling the point process N_t and the claim size sequence X_k. Both processes are assumed to be independent, hence can be simulated independently of each other. Here we will focus on simulating the point process. We will discuss five prominent examples of N_t, namely the classical (homogeneous) Poisson process, the non-homogeneous Poisson process, the mixed Poisson process, the Cox process (also called the doubly stochastic Poisson process) and the renewal process.

2 Claim arrival process

Simulation of a point process typically reduces to modeling 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.

2.1 Homogeneous Poisson process

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 point process, and (ii) the times between events are independent and identically distributed with an exponential(λ) distribution, i.e. exponential with mean $1/\lambda$. Therefore, successive arrival times $T_1, T_2, \ldots, T_n$ of the Poisson process can be generated by the following algorithm:

Step 1: set $T_0 = 0$

Step 2: for $i = 1, 2, \ldots, n$ do

Step 2a: generate an exponential random variable E with intensity λ

[1]Hugo Steinhaus Center for Stochastic Methods, Wrocław University of Technology, 50-370 Wrocław, Poland.
[2]CASE – Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, 10099 Berlin, Germany.

Step 2b: set $T_i = T_{i-1} + E$

To generate an exponential random variable E with intensity λ we can use the inverse transform method, which reduces to taking a random number U distributed uniformly on $(0,1)$ and setting $E = F^{-1}(U)$, where $F^{-1}(x) = (-\log(1-x))/\lambda$ is the inverse of the exponential cumulative distribution function. In fact we can just as well set $E = (-\log U)/\lambda$ since $1-U$ has the same distribution as U.

Since for the homogeneous Poisson process the expected value $\mathbb{E}N_t = \lambda t$, it is natural to define the premium function in this case as $c(t) = ct$, where $c = (1+\theta)\mu\lambda$, $\mu = \mathbb{E}X_k$ and $\theta > 0$ is the *relative safety loading* which "guarantees" survival of the insurance company. With such a choice of the risk function we obtain the classical form of the risk process [8, 9].

2.2 Non-homogeneous Poisson process

One can think of various generalizations of the homogeneous Poisson process in order to obtain a more reasonable description of reality. Note that the choice of such a process implies that the size of the portfolio cannot increase or decrease. In addition, there are situations, like in motor insurance, where claim occurrence epochs are likely to depend on the time of the year or of the week [9]. For modeling such phenomena the non-homogeneous Poisson process (NHPP) is suited much better than the homogeneous one. The NHPP can be thought of as a Poisson process with a variable intensity defined by the deterministic intensity (rate) function $\lambda(t)$. Note that the increments of a NHPP do not have to be stationary. In the special case when $\lambda(t)$ takes the constant value λ, 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 in the homogeneous one. The first approach is based on the observation [9] 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 $\widetilde{\lambda} = \int_s^t \lambda(u)du$. Hence, the cumulative distribution function F_s of the waiting time W_s is given by

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

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

Step 1: set $T_0 = 0$

Step 2: for $i = 1, 2, \ldots, n$ do

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

Step 2b: set $T_i = T_{i-1} + F_s^{-1}(U)$

The second approach, known as the "thinning" or "rejection method", is based on the following observation [3, 12]. Suppose that there exists a constant $\overline{\lambda}$ such that $\lambda(t) \le \overline{\lambda}$ 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 with probability $\lambda(T_i^*)/\overline{\lambda}$, independently of all other arrivals, then the sequence $T_1, T_2, \ldots$ of the accepted arrival times (in ascending order) forms a sequence of the arrival times of a non-homogeneous Poisson process with rate function $\lambda(t)$. The resulting algorithm reads as follows:

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

Step 2: for $i = 1, 2, \dots, n$ do

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

Step 2b: set $T^* = T^* + E$

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

Step 2d: if $U > \lambda(T^*)/\overline{\lambda}$ then return to step 2a ($\rightarrow$ reject the arrival time) else set $T_i = T^*$ ($\rightarrow$ accept the arrival time)

As mentioned in the previous section, the inter-arrival times of a homogeneous Poisson process have an exponential distribution. Therefore steps 2a–2b generate the next arrival time of a homogeneous Poisson process with intensity $\overline{\lambda}$. Steps 2c–2d amount to rejecting (hence the name of the method) or accepting a particular arrival as part of the thinned process (hence the alternative name).

We finally note that since in the non-homogeneous case the expected value $\mathbb{E}N_t = \int_0^t \lambda(s)ds$, it is natural to define the premium function as $c(t) = (1+\theta)\mu \int_0^t \lambda(s)ds$.

2.3 Mixed Poisson process

The very high volatility of risk processes, for example expressed in terms of the *index of dispersion* $\mathrm{Var}(N_t)/\mathbb{E}(N_t)$ being greater than 1 – a value obtained for the homogeneous and the non-homogeneous cases, led to the introduction of the mixed Poisson process [2, 11]. 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.

In the mixed Poisson process the distribution of N_t is given by a mixture of Poisson processes. This means that, conditioning on an extrinsic random variable Λ (called a *s*tructure variable), the process N_t behaves like a homogeneous Poisson process. The process can be generated in the following way: first a realization of a non-negative random variable Λ is generated and, conditioned upon its realization, N_t as a homogeneous Poisson process with that realization as its intensity is constructed. Making the algorithm more formal we can write:

Step 1: generate a realization λ of the random intensity Λ

Step 2: set $T_0 = 0$

Step 3: for $i = 1, 2, \dots, n$ do

Step 3a: generate an exponential random variable E with intensity λ

Step 3b: set $T_i = T_{i-1} + E$

Since for each t the claim numbers N_t up to time t are Poisson with intensity Λt, in the mixed case it is reasonable to consider the premium function of the form $c(t) = (1+\theta)\mu\Lambda t$.

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. Cox processes seem to form a natural class for modeling risk and size fluctuations. 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.

This definition suggests that the Cox process 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. Making the algorithm more formal we can write:

Step 1: generate a realization $\lambda(t)$ of the intensity process $\Lambda(t)$ for a sufficiently large time period

Step 2: set $\overline{\lambda} = \max\{\lambda(t)\}$

Step 3: set $T_0 = 0$ and $T^* = 0$

Step 4: for $i = 1, 2, \ldots, n$ do

- **Step 4a:** generate an exponential random variable E with intensity $\overline{\lambda}$
- **Step 4b:** set $T^* = T^* + E$
- **Step 4c:** generate a random variable U distributed uniformly on $(0, 1)$
- **Step 4d:** if $U > \lambda(T^*)/\overline{\lambda}$ then return to step 4a ($\rightarrow$ reject the arrival time) else set $T_i = T^*$ ($\rightarrow$ accept the arrival time)

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 $c(t) = (1 + \theta)\mu \int_0^t \Lambda(s)ds$.

2.5 Renewal process

Generalizing the point process we come to the position where we can make a variety of different distributional assumptions on the sequence of waiting times $\{W_1, W_2, \ldots\}$. In some particular cases it might be useful to assume that the sequence is generated by a *renewal process* of claim arrival epochs, i.e. the random variables W_i are i.i.d. and nonnegative. 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 for a renewal process:

Step 1: set $T_0 = 0$

Step 2: for $i = 1, 2, \ldots, n$ do

- **Step 2a:** generate a random variable X with an assumed distribution function F
- **Step 2b:** set $T_i = T_{i-1} + X$

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 total claims. For renewal claim arrival processes a constant premium rate allows for a constant safety loading [7]. Let N_t be a renewal process and assume that W_k has finite mean $1/\lambda$. Then the premium function is defined in a natural way as $c(t) = (1 + \theta)\mu\lambda t$, like in the homogeneous Poisson process case.

3 Simulation of risk processes

In this section we will illustrate some of the models described earlier. We will conduct the analysis on the PCS (Property Claim Services [13]) dataset covering losses resulting from catastrophic events in USA that occurred between 1990 and 1999. The data includes market's loss amounts in USD adjusted for inflation. Only natural perils which caused damages exceeding 5 million dollars were taken into consideration. Two largest losses in this period were caused by Hurricane Andrew (24 August 1992) and the Northridge Earthquake (17 January 1994).

The claim arrival process was analyzed by Burnecki et al. [6]. They fitted exponential, lognormal, Pareto, Burr and gamma distributions to the waiting time data and tested the fit with the χ^2, Kolmogorov-Smirnov, Cramer-von Mises and Anderson-Darling test statistics, see [1, 5]. The χ^2 test favored the exponential distribution with $\lambda_w = 30.97$, justifying application of the homogeneous Poisson process. However, other tests suggested that the distribution is rather lognormal with $\mu_w = -3.88$ and $\sigma_w = 0.86$ leading to a renewal process. Since none of the analyzed distributions was an unanimous winner Burnecki et al. [6] suggested to fit the rate function $\lambda(t) = 35.32 + 2.32 \cdot 2\pi \cdot \sin[2\pi(t - 0.20)]$ and treat the claim arrival process as a non-homogeneous Poisson process.

The claim severity distribution was studied by Burnecki and Kukla [4]. They fitted lognormal, Pareto, Burr and gamma distributions and tested the fit with various non-parametric tests. The lognormal distribution with $\mu_s = 18.44$ and $\sigma_s = 1.13$ passed all tests and yielded smallest errors. The Pareto distribution with $\alpha_s = 2.39$ and $\lambda_s = 3.03 \cdot 10^8$ came in second.

The simulation results are presented in Figure 1. We consider a hypothetical scenario where the insurance company insures losses resulting from catastrophic events in the United States. The company's initial capital is assumed to be $u =$ USD 100 billion and the relative safety loading used is $\theta = 0.5$. We choose four models of the risk process whose application is most justified by the statistical results described above: a homogeneous Poisson process with lognormal claim sizes, a non-homogeneous Poisson process with lognormal claim sizes, a non-homogeneous Poisson process with Pareto claim sizes, and a renewal process with Pareto claim sizes and lognormal waiting times.

In all subplots of Figure 1 the thick solid line is the "real" risk process, i.e. a trajectory constructed from the historical arrival times and values of the losses. The thin solid line is a sample trajectory. 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 20000 trajectories of the risk process. Recall that the function $\hat{x}_p(t)$ is called a sample p-quantile line if for each $t \in [t_0, T]$, $\hat{x}_p(t)$ is the sample p-quantile, i.e. if it satisfies $F_n(x_p-) \leq p \leq F_n(x_p)$, where F_n is the sample distribution function. Quantile lines are a very helpful tool in the analysis of stochastic processes. For example, they can provide a simple justification of the stationarity (or the lack of it) of a process, see [10]. In Figure 1 they visualize the evolution of the density of the risk process. Clearly, if claim severities are Pareto distributed then extreme events are more probable to happen than in the lognormal case.

References

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

[2] H. Ammeter (1948), *A generalization of the collective theory of risk in regard to fluctuating basic probabilities*, Skand. Aktuar Tidskr., 171-198.

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

[4] K. Burnecki, G. Kukla (2003), *Pricing of zero-coupon and coupon CAT bonds*, Applicationes Mathematicae 28.3.

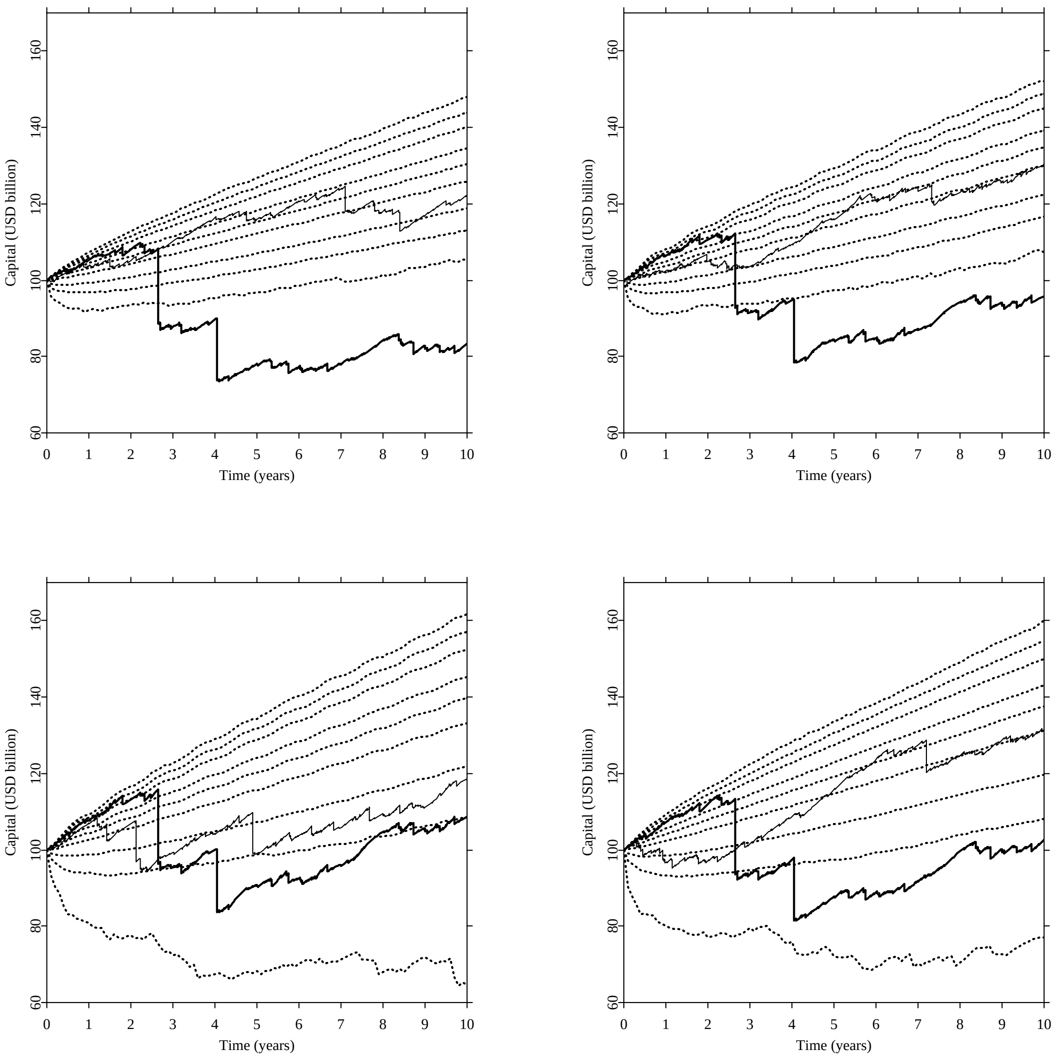

Figure 1: Simulation results for a homogeneous Poisson process with lognormal claim sizes (top left), a non-homogeneous Poisson process with lognormal claim sizes (top right), a non-homogeneous Poisson process with Pareto claim sizes (bottom left), and a renewal process with Pareto claim sizes and lognormal waiting times (bottom right). Figures were created with the Insurance library of XploRe [14].

[5] K. Burnecki, G. Kukla, R. Weron (2000), *Property insurance loss distributions*, Physica A 287, 269–278.

[6] K. Burnecki, P. Miśta, R. Weron (2004), *Visualization of risk processes*, in P. Cizek, W. Härdle, R. Weron (eds.), Statistical Tools for Finance and Insurance, Springer, Berlin.

[7] P. Embrechts and C. Klüppelberg (1993), *Some aspects of insurance mathematic*, Theory Probab. Appl. 38(2), 262-295.

[8] P. Embrechts, C. Klüppelberg and T. Mikosch (1997), Modelling Extremal Events, Springer, Berlin.

[9] J. Grandell (1991), Aspects of Risk Theory, Springer, New York.

[10] A. Janicki, A. Weron (1994), Simulation and Chaotic Behavior of α-Stable Stochastic Processes, Marcel Dekker, New York.

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

[12] S. Ross (2001), Simulation, 3rd ed., Academic Press, San Diego.

[13] See Insurance Services Office Inc. (ISO) web site: www.iso.com/products/2800/prod2801.html.

[14] See XploRe web site: www.xplore-stat.de.

HSC Research Report Series 2003

For a complete list please visit http://ideas.repec.org/s/wuu/wpaper.html

01 *Modeling electricity prices: Jump diffusion and regime switching* by Rafał Weron, Michael Bierbrauer and Stefan Trück

02 *Methods for determining the presence of periodic correlation based on the bootstrap methodology* by Ewa Broszkiewicz-Suwaj

03 *On ARMA(1,q) models with bounded and periodically correlated solutions* by Aleksander Weron and Agnieszka Wyłomańska

04 *An introduction to simulation of risk processes* by Krzysztof Burnecki, Wolfgang Härdle and Rafał Weron