"SIMULATION OF RISK PROCESSES WITH VARIABLE PREMIUN RATE"

Professor Semyen Spivak

Bashkir State University

Frunze srt. 32, Ufa, 450074, Russia

Spivak@bsu.bashedu.ru

Andrei Klimin

Birsk State Pedagogical Institute

Internacionalnaya str. 10, Birsk, Bashkortostan, 452450, Russia

Klimin_a_s@yahoo.com

Doctor Gyuzel Minnulina

Insurance Company "SOGAZ"

Ufa, Russia

Synopsy

The classical Cramer-Lundberg process assumes that the rate of premium income received by the insurance company is a constant. The present paper examines generalization of the classical process in which the rate of premium income varies with time or is stochastic. To construct the generalized process we use the Monte-Carlo simulation. The ruin probability of the insurance company in case of variable premium rate is estimated. The results are illustrated by actual data.

1. INTRODUCTION

The problem of estimating (calculating) the *probability of ruin* has been discussed in many papers on actuarial risk theory. Most of these papers assume that surplus at time t is defined as a sum of initial surplus u, income obtained from premium payments $P(t)$ minus total amount of paid claims $S(t)$:

$$U(t) = u + P(t) - S(t). \qquad \textbf{(1.1)}$$

$S(t)$ is a stochastic process, for example, Poisson process and amount of claims has some distribution function F in these papers. The classical Cramer-Lundberg process assumes that the rate of premium income received by an insurance company is constant and can be written as

$$P(t) = ct,$$

and the risk process in this case can be written by the formula:

$$U(t) = u + ct - \sum_{i=1}^{N_t} Y_i, \qquad \textbf{(1.2)}$$

where $U(t)$ — is surplus at time t, u – initial surplus, c – is a constant premium rate, Y_i – are claim amounts paid out between (i-1)-th and i-th claims, N_t — is the number of claims occurred up to time t. The premium income rate according to the expected value principle is

$$c = (1+\theta)\lambda\mu,$$

where θ — is a strictly positive parameter called *safety loading*, μ — is the mean of claim amounts, λ — is an intensity of Poisson process.

Let us consider the problem of finding ruin probability of an insurance company. We define the time of ruin as

$$T = \inf\{t \mid U(t) < 0\}$$

and define $T = \infty$ if $U(t) \geq 0$ for all t. We define the ruin probability over a finite time $T_{\max}$

$$\psi(u, T_{\max}) = P[T < T_{\max} \mid U(0) = u] \qquad \textbf{(1.3)}$$

and the probability of ultimate ruin as

$$\psi(u) = P[T < \infty \mid U(0) = u]. \qquad \textbf{(1.4)}$$

We can derive explicit formulae **(1.3)** and **(1.4)** for rather a small number of claims distribution.

If we know the so-called *adjustment coefficient* $\kappa > 0$, which is non-zero solution of the equation:

$$1+(1+\theta)\kappa\mu = L_X(-\kappa),$$

and we have a Laplace transform for the distribution of the claim amounts, then we can use a *Cramer's asymptotic ruin formula*:

$$\psi(u) \sim -\frac{\theta\mu e^{-\kappa u}}{L'_X(-\kappa)+\mu(1+\theta)}, \quad u \to \infty.$$

We are going to introduce more general models for this classical Cramer-Lundberg risk process.

$$X(t) = U(t) + W(t),$$

where $U(t)$ – is the classical Cramer-Lundberg process, $W(t)$ – is some stochastic perturbation process, for example, $W(t) = \varepsilon W_B(t)$, where $\varepsilon > 0$ is a constant, $W_B(t)$ — is the standard Brownian motion independent of $U(t)$. The rate at which premiums come in is non-constant and is a function of the current surplus. In this case the accumulated premiums at time t can be written by the following formula

$$P(t) = \int_0^t c\big(U(\tau)\big)d\tau . \qquad \textbf{(1.5)}$$

It is possible to consider generalizations of this model. In particular, when premium income rate is a function of time:

$$c = c(t).$$

Then total premiums up to time t can be written as

$$P(t) = \int_0^t c(\tau)d\tau . \qquad \textbf{(1.6)}$$

The process $P(t)$ can also be defined as a stochastic process.

In this paper we consider the methods of estimating the probability of ruin for variable and stochastic premium rate.

ESTIMATING OF RUIN PROBABILITY BY SIMULATION

Monte-Carlo can be one of the possible methods of estimating ruin probability.

The main idea of this method is multiple simulation of the risk processes and calculation of the relative frequency of processes leading to ruin. But we cannot get the information about the risk processes when the time goes to the infinity by straightforward simulation. Therefore, this method allows us to estimate the ruin probability in finite time $T_{\max}$ not only for the classical Cramer-Lundberg processes but also for ordinary and delayed renewal processes and others.

Consider the following problem. We would like to evaluate ruin probability of an insurance company given the information about all premiums and claims occurring in some time period. If we know a constant premium income rate, initial surplus and claims distribution then we can try to use the results for classical Cramer-Lundberg process. If we can find an adjustment coefficient and Laplace transform for our claims distribution then we can also use the famous asymptotic Cramer-Lundberg formula.

But in real insurance practice the process $P(t)$ cannot be expressed by ct exactly and therefore we can use Monte-Carlo simulation.

We suggest the following method of estimating ruin probability. With the actual data on premiums and claims we shall construct a suitable risk model and then try to estimate ruin probability by simulations.

We denote $Z_i = I(\tau_i < T_{\max})$ as indicator of ruin of i-th process for the time $T_{\max}$. For N realizations of stochastic process we can write a number of ruined processes

$$N_r = \sum_{i=1}^{N} Z_i \, , \qquad \textbf{(2.1)}$$

and we can also calculate approximate values of ruin probability and complimentary 'survival' (non-ruin) probability as

$$\psi_N = \frac{1}{N}(Z_1 + Z_2 + \ldots + Z_N) = \frac{N_r}{N}, \qquad \textbf{(2.2)}$$

$$\delta_N = 1 - \psi_N$$

respectively.

We can write sample variance as

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{N}\left(Z_i - \psi_N\right)^2 =$$
$$= \frac{1}{n-1}\left(N_r\left(1-\psi_N\right)^2 + \left(N - N_r\right)\psi_N^2\right) =$$
$$= \frac{1}{n-1}\left(N\psi_N^2 - 2N_r\psi_N + N_r\right).$$

We will discuss a concrete example.

There are actual data of the premium income and claim amounts for some period. Consequently, we can analyze processes $P(t)$ and $S(t)$ for this period. There are dates of premium income (dates of contract agreements), validity period of policies and cost of these insurance policies (a value of premium). There are claim dates and claim amounts for claims.

Premium income

The period is from May 10, 1994 to November 27, 1997. The number of the records is 32791. In calculating we used the trial-version of the Palisade Corp. BestFit 4.0.4. The fitting to the lognormal distribution with density function

$$f(x) = \frac{1}{x\sqrt{2\pi\sigma^2}}\exp\left(-\frac{\left(\ln x - \mu\right)^2}{2\sigma^2}\right) \qquad \textbf{(2.3)}$$

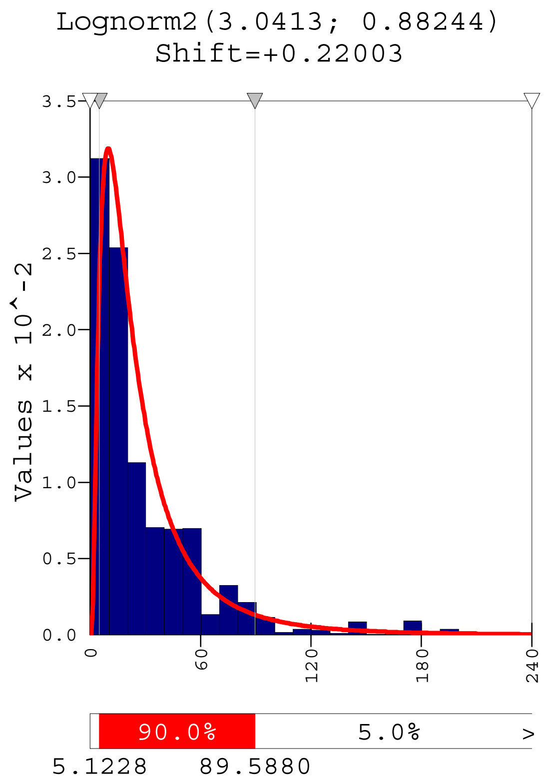

Fig. 1. The fit of Lognorm2 distribution to data of the premium income.

Time intervals between premiums

Intervals between the premiums were from 0 to 4 days.

Claim amounts

The period is from May 10, 1994 to November 27, 1997. The number of the records is 3207. The fitting to the lognormal distribution **(2.3)** gives

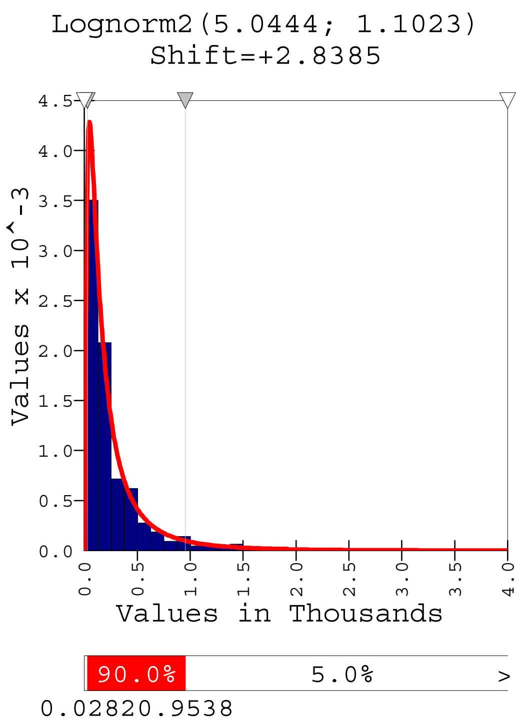

Fig. 3. The fit of Lognorm2 distribution to data of the claim amounts.

Random number generation

Before we start simulations we need to choose a method of generating random numbers. Below we discuss different possible methods.

First of all, we examine the *uniform variates*, which are random numbers that lie within a range from 0 to 1. There is the so-called *linear congruential generator* which generates a sequence of integrals $W_1, W_2, \ldots$. This algorithm starts with explicit setting of the first integer W_1 with the following recursive calculation of the remaining integers

$$W_{i+1} = \left(kW_i + C\right) \bmod p . \qquad \textbf{(2.4)}$$

Here k, C, p — are constants. k is called the *multiplier*, and C and p are integers called the *increment* and *modulus* respectively. A uniform random number from [0;1] is obtained by dividing W_i by modulus

$$U_i = \frac{W_i}{p}. \qquad \textbf{(2.5)}$$

The period of such generator equals $p-1$. The constants k, C, p ought to be chosen very carefully. For example, the following values would be suitable: $p = 2^{32} - 1 = 2147483647$, $C = 0$, $k = 16807$, $W_1 = 12345$.

There are many other uniform random number generators with other periods and other CPU timing, for example, nonlinear congruential generators, generators with shuffle, generators based on data encryption and others.

Modeling the standard Poisson process

For modeling the *standard Poisson process* with the rate λ, note that the interarrival times $\delta_i = T_i - T_{i-1}$ have exponential distribution with c.d.f. $F(x) = 1 - e^{-\alpha x}$ with parameter $\alpha = 1/\lambda$.

Exponential deviates X with parameter α can be obtained from uniform deviates in interval $[0,1]$ based on the so-called *inversion method*. Let $F(x)$ — be some c.d.f., then $F^{-1}(x)$ — is an inverse of the $F(x)$. Then

$$X \sim F^{-1}(U)$$

where $U \sim uniform(0,1)$. For the random deviate with *exponential* distribution we can write

$$Exp(\alpha) \sim -\frac{1}{\alpha}\ln(1-U) \sim -\frac{1}{\alpha}\ln(U). \qquad \textbf{(2.6)}$$

The same method can be easily used to produce *Pareto* random deviates ($F(x) = 1 - x^{-c}$):

$$Pa(c) \sim U^{-1/c}.$$

For the generation of a pair of *standard normal* random deviates, we can use the *Box-Muller method*:

$$\begin{aligned} N_1(0,1) &\sim \sqrt{-2\ln U_1}\sin(2\pi U_2), \\ N_2(0,1) &\sim \sqrt{-2\ln U_2}\cos(2\pi U_1), \end{aligned} \qquad \textbf{(2.7)}$$

where U_1, U_2 — is a pair of random deviates with a uniform distribution in the interval (0,1). To construct a deviate $N(\mu,\sigma^2)$ we can use the following formula

$$N(\mu,\sigma^2) \sim \sigma^2 N(0,1) + \mu. \qquad \textbf{(2.8)}$$

The random deviate with lognormal distribution can be obtained from the standard normal random deviates

$$Lognorm(\mu,\sigma) \sim \mu\exp(\sigma N(0,1)). \qquad \textbf{(2.9)}$$

Example 4. Estimation of the probability of ruin by simulation.

We construct N processes with parameters:

$$T_{max} = 2500\,,\ \ N = 1000\,,\quad p \sim Lognorm2(20.932,\, 0.8824) + 0{,}22\,,$$

$$Y \sim Lognorm2(155.1457,\, 1.1023) + 2{,}8385\ .$$

From **(2.1)** and **(2.2)** we can easily estimate the probability of ruin. For initial surplus varying from 0 to 40000 with step 1000 we have calculated following values of $\psi(u)$:

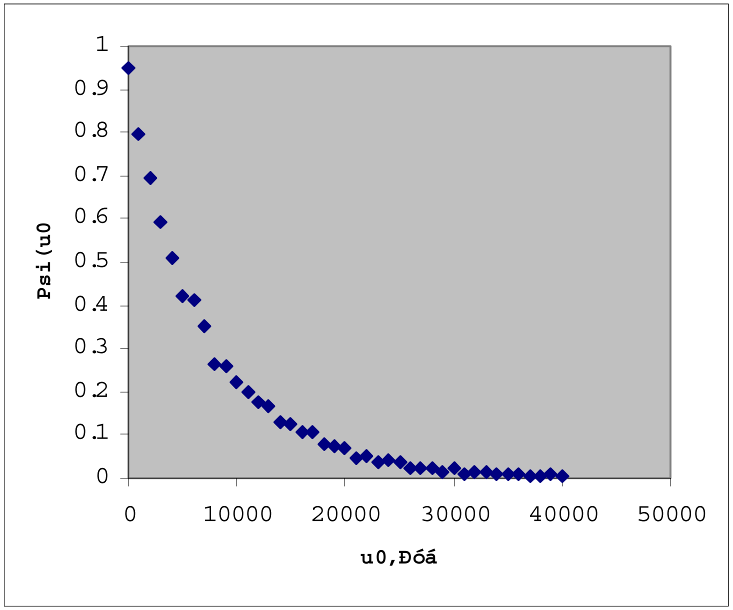

Fig. 4. The visualization of Table 2.

The obtained results can be fit by the exponential curve

$$\psi(u) = 0.91392e^{-0.000138152u}.$$

Example 5. Severity and time to ruin.

Let us show the results of simulation of 20 000 processes with parameters

$u = 5000$,

$T_{\max} = 3000$,

$p \sim Lognorm2(20.932, 0.8824) + 0.22$,

$Y \sim Lognorm2(155.1457, 1.1023) + 2.8385$.

The fitting of the Lomax (Pearson-2) distribution with c.d.f.

$$F(x) = 1 - \frac{b^q}{(x+b)^q}$$

to the data of severity gives the following parameters: $b = 581.5212$, $q = 1.8068$, shift = 0.0056 (see Fig. 5).

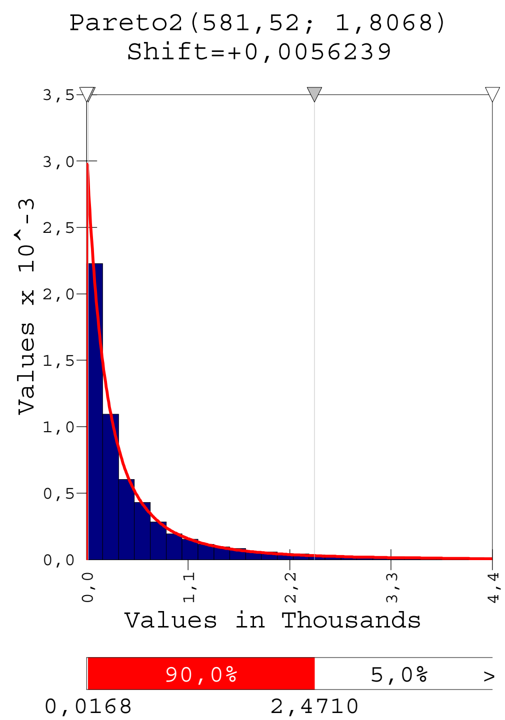

Fig. 5.

The fitting of the Inverse Gaussian distribution with density function

$$f(x)=\sqrt{\frac{\lambda}{2\pi x^3}}\exp\left(-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right)$$

to the data of the time to ruin (if the process is ruined) gives the following parameters

$\mu = 24.1549$, $\lambda = 9{,}0753$, shift = -1,336.

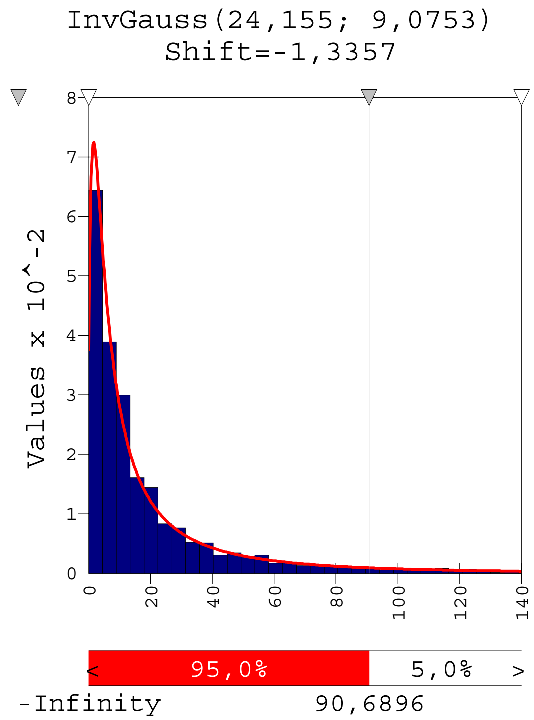

Fig. 6.

THE PRECISION OF THE RESULTS OBTAINED

Suppose the simulation gives us n values of ruin probability. Each of these values is obtained as a result of N runs. Then the estimate of ψ can be obtained as empirical mean:

$$\hat{\psi} = \frac{1}{n}(\psi_1 + \psi_2 + \ldots \psi_n), \qquad \textbf{(3.1)}$$

and variance $\sigma_\psi^2 = \mathrm{Var}\psi$ as empirical variance:

$$\hat{\sigma}_\psi^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(\psi_i - \hat{\psi}\right)^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n}\psi_i^2 - n\hat{\psi}^2\right). \qquad \textbf{(3.2)}$$

The question is how large n and N we should choose in order to obtain a given level of precision. For Monte-Carlo method we can only get some confidence interval for ruin probability. We will show what n we need to choose to obtain a result close enough to true answer with a given probability.

From central limit theorem it follows that

$$\sqrt{n}\left(\hat{\psi} - \psi\right) \xrightarrow{D} N\left(0, \sigma_\psi^2\right). \qquad \textbf{(3.3)}$$

Then 95% confidence interval is:

$$\bar{\psi} \pm \frac{1,96\bar{\sigma}_{\psi}}{\sqrt{n}} = \left[\bar{\psi} - \frac{1,96\bar{\sigma}_{\psi}}{\sqrt{n}}, \bar{\psi} + \frac{1,96\bar{\sigma}_{\psi}}{\sqrt{n}}\right]. \qquad \textbf{(3.4)}$$

For example, if we need to evaluate a number n of realizations so that the error is smaller than ε with probability 95% we can write the following formula

$$n = \frac{1,96^2 \sigma_{\psi}^2}{\varepsilon^2}. \qquad \textbf{(3.5)}$$

In practice the value σ_{ψ}^2 is unknown. But we can take some initial number n' of the values $\psi_1, \psi_2, ..., \psi_{n'}$, calculate the variance $\sigma_{\psi}'^2$ and then calculate $n = 1,96^2 \sigma_{\psi}'^2 / \varepsilon^2$.

Conclusion

The risk models described in this paper do not aim at an exhaustive coverage of real processes occurring in insurance practice. However, the impossibility of obtaining analytical results compels the researcher to use the methods proposed in this paper.

Regrettably, the Monte-Carlo method requires considerable time expenditure in order to reach an acceptable precision, and in direct simulation a slight gain in the precision of result is achieved at a considerable increase of calculation time. Besides, in this case we cannot obtain the ruin probability at infinity. But these disadvantages are offset by the simplicity of constructing models for a number of complex risk processes.

For example, the classical risk model assumes that the number of claims N_t by time t obeys the Poisson distribution with rate λ. More realistic is the method in which claims are put in in accordance with the inhomogeneous Poisson process whose rate changes with time. A change of the rate $\lambda(t)$ can describe various phenomena, for example, a seasonal increase in the frequency of fires in dry and hot weather, a rise in traumatism, a rise in road accidents on an ice-slick, and so on. The application the inhomogeneous Poisson process to premium income can help describe the various cycles and trends in concluding contracts with insurance clients. It is rather difficult to obtain specific results within the framework of classical theory, but simulation of the inhomogeneous Poisson process both for putting in claims and premium income can yield ruin probability values for similar processes.

It is assumed in the classical risk model that the number of claims by the moment t - N_t has a Poisson distribution with parameter λ. A more realistic model would be the one describing claims arising by time-inhomogeneous Poisson process with intensity varying over time. Changes in intensity $\lambda(t)$ can reflect different events, for example, seasonal factors: more fires in summer, more car accidents during snowy winter, etc.

The application of time-inhomogeneous Poisson process describing claims arising might be useful for understanding all sorts of cycles and seasonal trends in the work of an insurance company. |