Viera Pacáková
University of Economics
in Bratislava
Faculty of Economic Informatics
Department of Statistics
Dolnozemská cesta 1
852 35 Bratislava
viera.pacakova@euba.sk

SIMULATION OF COLLECTIVE RISK MODEL

Summary

The article focuses on providing brief theoretical definitions of the basic terms and methods of modelling and simulations of insurance risks in non-life insurance by means of mathematical and statistical methods and operational research, and on practical examples of applications of these methods using statistical software.

While the risk assessment of insurance company in connection with its solvency is a rather complex and comprehensible problem, its solution starts with statistical modelling of number and amount of individual claims. Successful solution of these fundamental problems enables solving of curtail problems of insurance such as modelling and simulation of collective risk, premium and reinsurance premium calculation, estimation of probability of ruin etc.

The article also presents some essential ideas underlying Monte Carlo methods and their application to modelling of insurance risks. Solving problem is to find the probability distribution of the collective risk in non-life insurance portfolio. Simulation of the compound distribution function of the aggregate claim amount can be carried out, if the distribution functions of the claim number process and the claim size are assumed given. The Monte Carlo simulation is suitable method to confirm the results of other methods and for treatments of catastrophic claims, when small collectives are studied.

Analysis of insurance risks using risk theory is important part of the project Solvency II. Risk theory is analysis of stochastic features of non-life insurance processes. The field of application of risk theory has grown rapidly. There is a need to develop the theory into form suitable for practical purposes and to demonstrate their application. Modern computer simulation techniques open up a wide field of practical applications for risk theory concepts, without requiring the restrictive assumptions and sophisticated mathematics. This article presents some comparisons of the traditional actuarial methods and of simulation methods of the collective risk model.

Keywords

insurance risks, loss distributions, collective risk model, goodness of fit tests, simulation, Monte Carlo method

ACM classification

J.4 SOCIAL AND BEHAVIORAL SCIENCES, *Economics,* I.6 SIMULATION AND MODELING, *Monte Carlo*

JEL classification

C1 - Econometric and Statistical Methods: General, C15 - Simulation Methods

1. MODELLING OF INSURANCE RISKS

We shall consider a short term insurance contract covering a risk. By a risk we mean either a single policy or a specified group of policies. The random variable S denotes the aggregate claims paid by the insurer in the year in respect of this risk. We are going to construct the models for the random variable S – so called the collective risk models.

A first step in the construction of a collective risk model is to write S in terms of the number of claims arising in the year, denoted by the random variable N, and the amount of each individual claim. Let the random variable X_i denote the amount of the i-th claim. Then

$$S = X_1 + X_2 + \dots + X_N \qquad (1)$$

where $X_1, X_2, ..., X_N$ are independent and identically distributed variables, $N, X_1, X_2, ..., X_N$ are mutually independent, and if $N = 0$, than $S = 0$ too.

The problems we will be solving are the derivation of the moments and distribution of S in terms of the moments and distributions of N and the X_i's.

We will assume that the moments and the distributions of N and X_i's are known with certainty. In practice these would probably be estimated from some relevant data using methods of parameters' estimation and goodness of fit tests.

We shall denote by $G(s)$ distribution function of S and $F(x)$ the distribution function of X_i, so that $G(s) = P(S \leq s) = F_S(s)$ and $F(x) = P(X_i \leq x)$. The k-th moment of X_i about zero, $k = 1, 2, 3, \ldots$, will be denoted $m_k = E\left(X_i^k\right)$.

We will not use exact methods for evaluating $G(s)$, but we will use approximate methods. For the approximate methods we need to know the moments of S. Basic expressions, known from actuarial literature there are $E(S) = E(N)m_1$

$$D(S) = E(N)\left(m_2 - m_1^2\right) + D(N)\, m_1^2 \quad (2)$$

$$M_S(z) = M_N(\ln M_X(z))$$

The distribution of S is an example of a compound distribution. We consider the most important case when N has a Poisson distribution. We say that S has a compound Poisson distribution with parameters λ and $F(x)$. In this case results (2) we can express in simple forms:

$$E(S) = \lambda m_1 \quad (3)$$

$$D(S) = \lambda\left(m_2 - m_1^2\right) + \lambda m_1^2 = \lambda m_2 \quad (4)$$

$$M_S(z) = e^{\lambda[M_x(z)-1]} \quad (5)$$

$$\mu_3(S) = \lambda m_3 \quad (6)$$

$$\gamma = \frac{\mu_3(S)}{[D(S)]^{\frac{3}{2}}} = \frac{\lambda m_3}{(\lambda m_2)^{\frac{3}{2}}} > 0 \quad (7)$$

The coefficient of skewness γ shows that the distribution of S is positively skewed and for large values of λ the distribution of S is almost symmetric.

We suppose that all we know or can confidently estimate about S are its mean and variance. Bearing in mind the Central Limit Theorem, this suggests assuming S is approximately normally distributed. An important drawback of this approximation may be that normal density is symmetric, i.e. has zero skewness, and has a right hand tail which goes to zero very quickly. For many types of insurance the distribution of S is positively skewed with a fairly heavy right hand tail and so normal approximation will tend to underestimate $P(S > x)$ for large values of x.

Suppose we know or can estimate with reasonable confidence the first three moments of S. One way of avoiding or at least reducing the problem of underestimating tail probabilities is to approximate the distribution of S by a translated gamma distribution. Let μ, σ^2 and γ denote the mean, variance and coefficient of skewness of S. We assume S has approximately the same distribution as the random variable $k+Y$, where k is a constant and Y has a gamma distribution $G(\alpha, \beta)$. The parameters k, α and β are chosen so that $k + Y$ has the same first three moments as S.

Equating the coefficients of skewness, variance and means of S and $k + Y$ gives the following three formulae

$$\gamma = \frac{2}{\sqrt{\alpha}}, \quad \sigma^2 = \frac{\alpha}{\beta^2}, \quad \mu = k + \frac{\alpha}{\beta} \quad (8)$$

from which α, β and k can be calculated from the known values of μ, σ^2 and γ.

2. MONTE CARLO SIMULATION OF COLLECTIVE RISK

The simulation of the values of S consists of the following steps:

1. Generate the number of claims n_1 from the known distribution of variable N (Poisson, negative binomial ...) using the random number generator.
2. Generate from the known distribution of the individual claim amount X just n_1 values of the individual losses $x_1, x_2, ..., x_{n_1}$.
3. The sum $s_1 = x_1 + x_2 + ... + x_{n_1}$ gives the first random number s_1 of the aggregate claim amount (collective risk) S.
4. The steps 1 to 3 repeat n-times to get generated random numbers $s_1, s_2, ..., s_n$ from unknown distribution of S.

Simulated values $s_1, s_2, ..., s_n$ enable us to solve two important tasks:

1. To verify suitability of probability models of *S*, those found by exactness actuarial methods.
2. To find the probability model of S by application of Goodness of Fit Tests using sampling data generated by Monte Carlo simulation.

3. EXAMPLE OF APPLICATION

Suppose that the number of claims *N* incurred in time period of one year follows a Poisson distribution with parameter $\lambda = 10\,000$.

We know the values of 91 individual claims made on an insurance portfolio. We will assume that these individual claim amounts are drawn from a particular distribution, called a loss distribution. Using Maximum Likelihood estimation and Goodness of fit tests in statistical analytical system SAS 9.1 we have verified that lognormal distribution with parameters $\mu = 9{,}741$ and $\sigma^2 = 2{,}165$ give a very good fit to the empirical data of individual claims amounts.

The first three moment of lognormal distribution can be calculated from the formula

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

Then

$m_1 = 50\ 171{,}2$

$m_2 = 21936129213$

$m_3 = 8{,}35827 \cdot 10^{16}$

Using formulaes (3) to (7) we calculated the descriptive measures of the collective risk S in Table 1.

Table 1 **Descriptive Measures of the Collective Risk *S***

E(S)	D(S)	$\mu_3(S)$	$\gamma(S)$
501712000	2,19361E+14	8,35827E+20	0,257262525

Source: Own Calculations

Except of the normal approximation of S with parameters $\mu = 501\,712\,000$ and $\sigma^2 = 2{,}1936 * 10^{14}$ we can approximate the distribution of S by a translated gamma distribution with parameters in table 2.

Table 2 **The parameters of Translated Gamma Distribution of S**

α	β	k
60,43756191	5,24896E-07	386570080,3

Source: Own Calculations

Using the own computer programme of Monte Carlo simulation in SAS system we have generated 10000 values of S (See: http://www.fhi.sk/files/katedry/ks/pacakova-inauguracna/Priloha-6.xls).

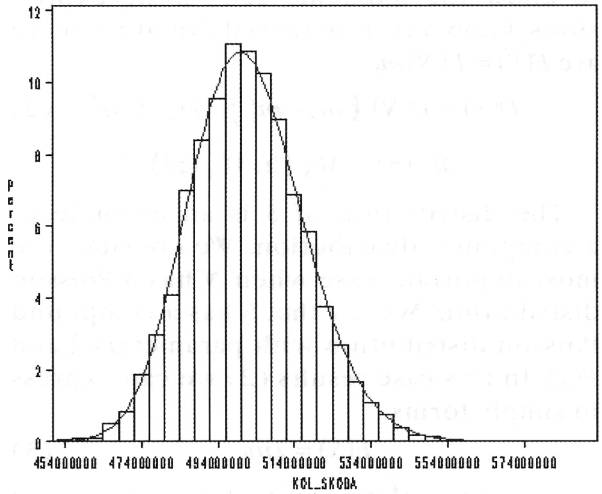
We use these simulated values of S to verify the suitability of translated gamma distribution with parameters in table 2 by goodness of fit tests in SAS Enterprise Guide 3.0 (table 3). Because of p-value > 0, 05, all three tests in table 3 confirm the translated gamma distribution gives a good fit to collective risk model of *S*.

Table 3 **Goodness of Fit Tests for Translated Gamma Distribution**

Goodness-of-Fit Tests for gamma distribution				
Test	**Statistic**		**p-value**	
Kolmogorov-Smirnov	D	0.00766886	Pr > D	>0.250
Cramer-von Mises	W-Kv	0.10683175	Pr > W-Kv	>0.250
Anderson-Darling	A-Kv	0.89623670	Pr > A-Kv	>0.250

Source: Own Calculations in SAS Enterprise Guide 3.0

Figure 1 **Histogram and fitted translated gamma model**



Using sampling data generated by Monte Carlo simulation we have find also transacted lognormal distribution in SAS system (table 4) with parameters estimated by maximum likelihood method (table 5) as a model that gives a good fit to simulated data of *S*.

Table 4 **Goodness of Fit Tests for Translated Lognormal Distribution**

Goodness-of-Fit Tests for lognormal distribution				
Test	**Statistic**		**p-Value**	
Kolmogorov-Smirnov	D	0.00649898	Pr > D	>0.250
Cramer-von Mises	W-Kv	0.05637045	Pr > W-Kv	>0.250
Anderson-Darling	A-Kv	0.42641107	Pr > A-Kv	0.189

Source: Own Calculations in SAS Enterprise Guide 3.0

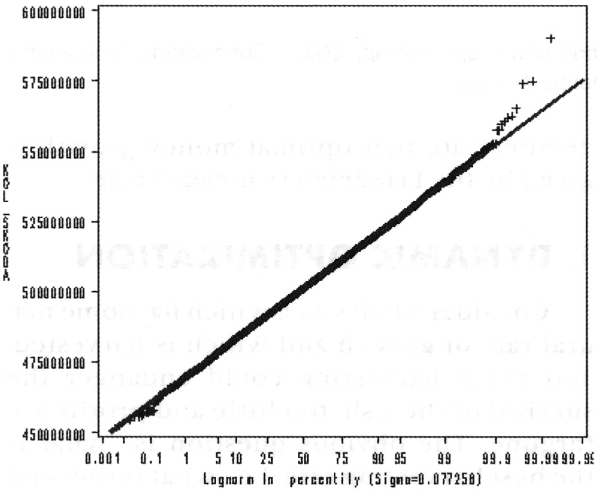
Table 5 **Parameters Estimation for Lognormal Distribution**

Parameters' Estimation for Lognormal Distribution		
Parameter	**Symbol**	**Est. parameters**
Threshold	Théta	3.1093E8
Scale	Zeta	19.06412
Shape	Sigma	0.077258
Mean		5.018E8
Std. deviation		14768374

Source: Own Calculations in SAS Enterprise Guide 3.0

The Q-Q plot (Figure 2) shows the quantiles of sampling data simulated by the Monte Carlo method plotted versus the equivalent percentiles of the fitted lognormal distribution. The fact that the points lie close to the diagonal line confirms the fact that the transacted lognormal distribution provides a good model for the simulated data of *S*.

Figure 2 **Q-*Q* plot of the empirical and expected quantiles**



5. CONCLUSION

The article presents the application of Monte Carlo methods in non-life insurance. Solving problem is to find the probability distribution of the collective risk in non-life insurance portfolio. Simulation of the compound distribution function of the aggregate claim amount can be carried out, if the distribution functions of the claim number process and the claim size are assumed given. The Monte Carlo simulation is suitable method to confirm the results of other methods and for treatments of catastrophic claims, when small collectives are studied.

REFERENCES

[1] BEARD, R. E.: *Risk Theory, the Stochastic Basis of Insurance*, Chapman and Hall, London 1994.

[2] BEARD, R. E. - PENTIKÄINEN, T. - PESONEN, E.: *Risk Theory*, Chapman and Hall, Cambridge 1990.

[3] BOHDALOVÁ, M.: Meranie finančných a poistných rizík metódou Value at Risk. In: *Aktuárska veda v teórii a v praxi* - zborník zo 6. vedeckého seminára, organoizovaného Ksatedrou matematiky FHI EU v Bratislave, Bratislava, 7. december 2007.

[4] BOHDALOVÁ, M, - STANKOVIČOVÁ, I.: Using the PCA in the Analyse of the risk Factors of the investment Portfolio. In: *Forum Statisticum Slovacum* 3/2006, s. 41-52. Bratislava: SŠDS, 2006.

[5] CURRIE, I. D.: *Loss Distributions*, Institute of Actuaries and Faculty of Actuaries, London and Edinburgh 1993.

[6] DAYKIN, C. D. - PENTIKÄINEN, T. - PESONEN, M.: *Practical Risk Theory for Actuaries*, Chapman and Hall, London 1996.

[7] DICKSON, D. C. M. - WATERS, H. R.: *Risk models*, Institute of Actuaries and Faculty of Actuaries, London and Edinburgh 1993.

[8] FECENKO, J.: *Neživotné poistenie*. Bratislava: Vydavateľstvo EKONÓM, 2006.

[9] HOGG, R. V. - KLUGMAN, S. A.: *Loss Distributions*, John Wiley & Sons, New York 1984.

[10] HORÁKOVÁ, G. - MUCHA, V.: *Teória rizika v poistení.* Bratislava: Vydavateľstvo EKONÓM, 2006.

[11] KAAS, R. - GOOVAERTS, M. - DHAENE, J. - DENUIT, M.: *Modern Actuarial Risk Theory*. Boston: Kluwer Academic Publishers, 2001.

[12] LINDA, B.: *Stochastické modely operačního výzkumu*. Bratislava: STATIS, 2004.

[13] PACÁKOVÁ, V.: *Aplikovaná poistná štatistika* - 3. prepracované a doplnené vydanie. Bratislava: IURA EDITION, 2004.

[14] PACÁKOVÁ, V.: *Modelovanie a simulácie poistných rizík*. Edícia: Habilitačné a inauguračné prednášky. Bratislava: Vydavateľstvo Ekonóm, 2006.

[15] RUBINSTEIN, REUVEN Y.: *Simulation and the Monte Carlo Method*. New York: John Wiley &Sons, Inc., 1981.

Biography:

prof. RNDr. Viera Pacáková, PhD. is a member of Department of Statistics at Faculty of Economic Informatics, University of Economics in Bratislava. Her focus of professional activity is application of quantitative methods, especially methods of statistics, theory of probability and actuarial science in economics practice.