APPLICATION OF STATISTICAL MODELING IN INSURANCE PROCESS

Vladimirs Jansons, Vitalijs Jurenoks, Konstantins Didenko
Riga Technical University, Riga, Latvia

Abstract

Risk assessment is one of the major challenges that must be addressed by each insurance company. To assess risk we need to know the value of losses as well as the probability of losses, since the risk cost is the basic component in evaluating the insurance indemnity. Statistical methods should be used for objective evaluation of insurance processes, but because of complexity in real life processes of insurance, statistical modelling techniques would be preferable. It is particularly important to develop and practically apply these methods in Latvia as in recent years (starting from 1992) the insurance market in Latvia has experienced steady growth. To improve the competitiveness of the insurance companies, especially small companies, it is simply impossible to do without methods allowing us to estimate the parameters of the insurance process. Taking this into consideration it becomes important to study information systems related to the processes of insurance and to use modern information technologies for processing the available empirical information and the dynamic scenario forecasting performance of the insurance process taking into account different assumptions about the factors that could affect the insurance process. The article deals with the various statistical models that assess the risks and losses of the insurance company allowing us to simplify the calculation of insurance premiums, insurance reserves and assess the financial stability of the insurance company with a sufficiently wide range of parameters of the real process of insurance. At the present time transition from local information systems to corporate information systems based on network technologies is being accomplished in the Baltic countries. Therefore, in the future it is important to include such statistical models into the integrated European information system of processing insurance information.

Key words: risk, statistical modeling, insurance company, financial stability, training process.

Introduction

Risk assessment is one the basic tasks to be tackled by any insurance company that wants to remain stable in the insurance market. To assess the insurance risk it is necessary to know the value and the probability of losses since the value of the insurance risk is the key component in assessing the insurance indemnity. For objective evaluation of insurance risks mathematical statistical methods and methods of actuarial mathematics should be used. The complexity of real life processes of insurance substantiates the necessity to apply the statistical modelling methods which due to the development of computer technologies are being more widely used in modelling insurance processes at all levels, starting with small insurance companies and ending with modelling of insurance processes at the level of big insurance companies. The development and application of statistical methods is of particular importance in Latvia since in recent years (starting from 1992) the insurance market in Latvia has experienced steady growth.

In 2010 the volume of non-life insurance grew up to 315 million lats (448 million Euros). In 2010 insurance indemnities constituted 193,1 million lats (274,8 million Euros). However, during the first two months of 2011 the volume of gross premiums totalled up to 31,46 million lats (44,76 million euros) which is 35% less than during the same period in 2010. To a certain extent it may be explained by falling prices of the insurance policies as well as by the overall financial instability in Europe. In 2011 three-quarters of the insurance company was working with profit of 1,5 million lats (2,13 million euros). Insurance companies gross premiums written by the 2011th the first three quarters year-on-year increase of 27,7% and was 176,8 million. lats, as well as the amount of gross claims paid increased by 12,1% and was 95,3 million lats.

Further improvement of the competitiveness of the insurance companies cannot be realized without applying methods which could estimate the parameters of the insurance process to guarantee sufficiently accurate and adequate decision-making process. This research deals with the various statistical models that assess the risks and losses of the insurance company and allow us to accurately enough assess insurance premiums, insurance reserves and the financial stability of the insurance company for a sufficiently wide range of parameters of the real process of insurance.

Approach to task modelling

The first works on the mathematical theory of insurance were published by F. Lundberg and X. Cramer who proposed and investigated the so-called classical model of the insurance process. The classical model allows us to calculate the probability of ruin and survival of the insurance company, the principles of choice of premium load and analyzes the survival time, the probability of insured accident, insurance rates and insurance claims.

This paper focuses on the development of economic and mathematical models of non-life insurance and their statistical modelling to estimate the overall losses of the insurance company. Let us consider the model of individual risk, which can be schematically represented as follows (see Fig. 1).

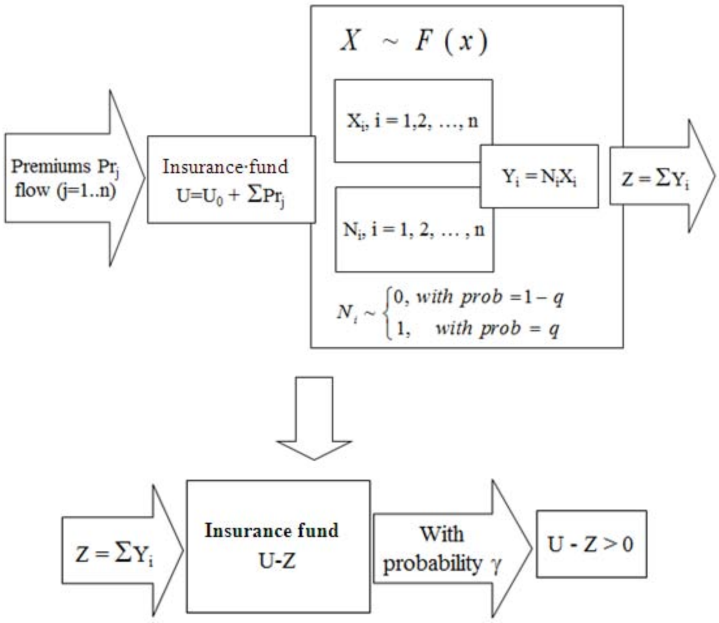

Fig. 1. Scheme of structure of individual risk model

Legend: n – number of contracts of insurance portfolio; j – index of the client; q – probability of the insured event; N_j – index of the insured event j, in the simplest case $N_j = \begin{cases} 1 & q \\ 0 & 1-q \end{cases}$ *; N – number of realized insurance events* $N = \sum_{j=1}^{n} N_j$ *; X_j – losses of the client with an index j, X_j, having the distribution function F(x); Y_j – insurance indemnity for the client with an index j losses,* $Y_i = N_i X_i$ *; Z – total compensation (of losses) of the insurance portfolio,* $Z = \sum_{j=1}^{n} Y_j = \sum_{j=1}^{n} N_j X_j$ *; γ - level of guarantees of the insurance company (usually in the range of 0,8 to 0,95); U0, U – value of the initial insurance fund and after a certain period of time.*

It is assumed that for the insurance portfolio the following conditions are met:

- number of contracts in the portfolio is constant;
- risks to customers are independent of each other;
- all payments are made without delay;
- function F(x) is equal for all clients.

Insurance portfolio modelling

Making use of the simplest Monte Carlo method when modelling the insurance portfolio with parameters: n = 1000, q = 0,1, F(x) - function of a uniform distribution in the interval (0; 1000) when assessing the average losses, dispersion of losses and coefficient of variation of the insurance portfolio, the relative errors compared with the exact results are as follows: (see Table 1).

Table 1. Insurance portfolio modelling results

	Model	Theory
M(Z)	49934,74	50000
D(Z)	31215305	30833000
Kvar(Z)	11,19%	11,11%

Relative errors	
e_M(Z) =	0,13%
e_D(Z) =	1,24%
e_Kvar(Z) =	0,75%

Insignificant relative errors (less than 2%) indicate the possibility of a sufficiently accurate analysis of the simplest insurance portfolio using the statistical Monte Carlo method. Further the possibility of using the statistical Monte Carlo method as an alternative to analytical methods for studying more complex insurance processes will be shown. To a large extent the insurance fund depends on how well the calculation of insurance premiums is done. To state the financial stability of the insurance company it is necessary to satisfy the following inequality:

$$U - Z = U_0 + \sum_{j=1}^{n} P_j - \sum_{j=1}^{n} N_j \cdot X_j > 0 \qquad (1)$$

with the given probability γ (usually $\gamma = 0{,}1$ or $0{,}05$). Since $U_0 > 0$, the inequality (1) will follow from the inequality:

$$\sum_{j=1}^{n} P_j - \sum_{j=1}^{n} N_j \cdot X_j > 0 \qquad (2)$$

Knowing the distribution F of the variable $\sum_{j=1}^{n} N_j \cdot X_j$, we may find such value C of the variable $\sum_{j=1}^{n} P_j$, where with probability γ, the inequality (2) as well as the inequality (1) hold. The value C shows the required level of aggregate premiums ensuring the stability of the insurance company with probability γ.

$$C = F^{-1}(1-\gamma) \qquad (3)$$

In many real life situations the analytical solution of the equation (3) turns out to be a complex mathematical problem not always having precise or sufficiently precise solution. In this case, a good alternative is the Monte Carlo method.

Assume that it is necessary to evaluate the possibility of reducing the ruin, using the process of reinsurance in the following case: the insurance portfolio contains N insurance contracts for 1 year from which the insurance sum of N_1 contracts is S_1 and the insurance sum of N_2 contracts is S_2. The probability of a claim is equal to q. We assume that the level of deductibles is C. Let us compare the solution of this problem by a) analytical method and b) using the Monte Carlo methods:

a) where $N = 8000$, $N_1 = 5000$, $N_2 = 3000$, $S_1 = 10000$Ls, $S_2 = 20000$Ls, $p = 0{,}02$ due to reinsurance when $C = 16000$Ls, the company seeks to reduce the probability of ruin from 0,14 to 0,13;

b) having applied for modelling the Monte Carlo method, we obtain a fairly accurate ($R^2 = 1{,}0$) regression dependence of the probability of bankruptcy depending on the value of losses (see Fig. 2).

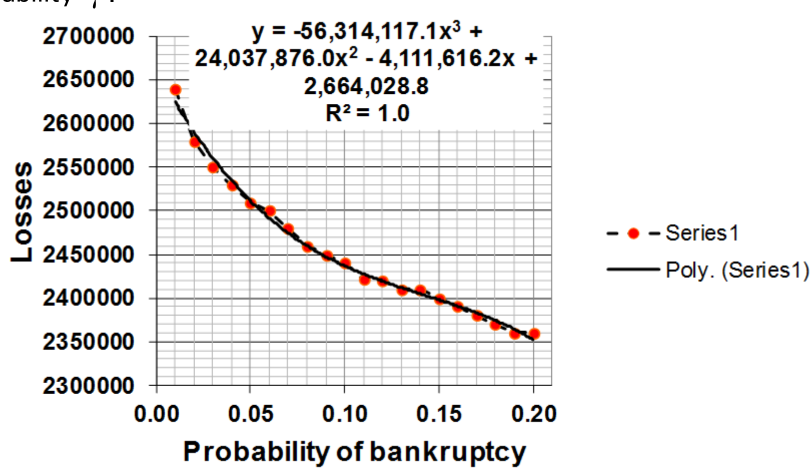

Fig. 2. Ruin probability of the insurance portfolio without reinsurance – the dotted line and polynomial approximation – solid line

After the introduction of the reinsurance process the probability of bankruptcy decreased from 0,15 to 0,13 (see Fig. 3). This agrees well with the result obtained by the analytical method, which in this case is rather time-consuming and requires a good knowledge of actuarial mathematics. Comparing the two methods for solving this problem conclusion may be made that the application of the Monte Carlo method is simpler than using the analytical method. Modelling ensures quicker and easier adaptation to various changes in the insurance situations which is practically very difficult to reach using analytical methods.

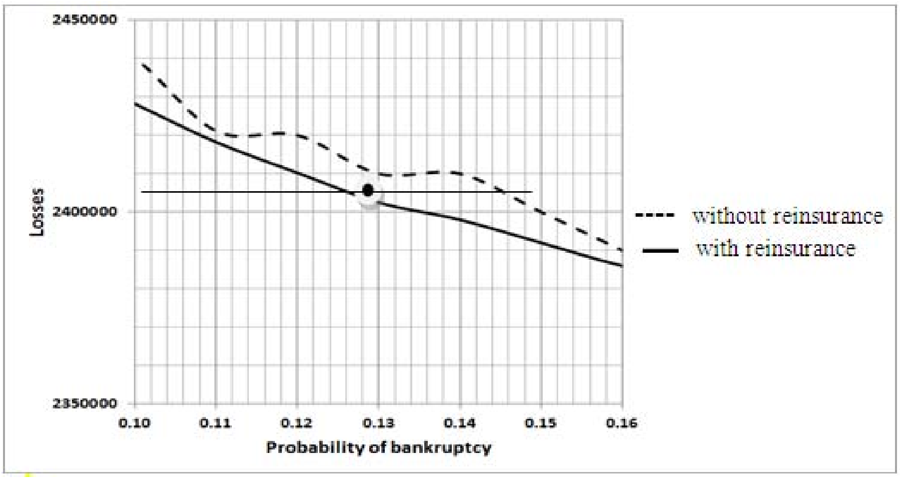

Fig. 3. Ruin probabilities of the insurance portfolio without reinsurance – the dotted line and with reinsurance – the solid line

Application of the modelling methods does not necessarily require knowledge of the analytical representation of the distribution functions, thus knowledge of the empirical distribution functions is quite sufficient (i.e., the existence of empirical information about the insurance situation in the country, about the values of insurance claims, etc.). If there is no empirical information about the losses of the insurance company which is typical at the initial stage, consider using benchmarking, finding and making comparison with a more or less similar insurance company in the given country or in the world. Experience shows that stable insurance companies in the same cluster of insurance (having approximately the same volume of services and providing the same kinds of insurance services, having the same insurance strategy) are similar and have very similar characteristic parameters.

The most complex subjects for the study of financial stability are the models of collective risk. The theory of collective risk was developed in 1909 by a small group of actuaries, mainly Scandinavian. In the theory of collective risk an insurance company is seen as a reservoir which produces a continuous stream of premiums and from which payments are made. The model consists of the following three elements:

1. the flow of premiums P (t) - the total amount of premiums received during the period (0, t);
2. q (N, t) - the probability that the n^{th} payment will be claimed during the period (0, t);
3. G (x) - probability with which the payments are made, and the amount paid does not exceed x.

From these three elements it may be derived that the probability for payment in the interval (0, t) does not exceed x what can be represented as a function F (x, t):

$$F(x,t)=\sum_{0}^{\infty} q(n,t)G^{(n)}(x)\,, \qquad (4)$$

where

$G^{(n)}(x)$ for N>0; n^{th} convolution of the function G(x) и $G^{(0)}=H(x)$ (Heaviside function).

For the statistical simulation of correlated random variables with distributions derived from empirical data, the authors used the method of copulas.

The copula for random values X_1, X_2,..., X_n can be described by equation:

$$C(u_1,u_2,...,u_n)=\Phi(F_1^{-1}(x_1),F_2^{-1}(x_2),...,F_n^{-1}(x_n)), \qquad (5)$$

where

F_i - marginal distribution for random value X_i, i=1,2,...,n.

The algorithm of simulation of random n-dimensional vector $X=(X_1, X_2,..., X_n)$ is:

I. Simulate a variable X with distribution function G such that the Laplace transform of G is the inverse of the generator.

II. Simulate n independent variates V1,... , Vn.

III. Return U=($^{-1}$(-log(V1)/X),..., ($^{-1}$(-log(Vn)/X)).

Frank, Clayton and the Gumbel copula can be simulated using this procedure. For example, for the Clayton copula simulation the algorithm is as follows:

I. Simulate a Gamma variate X~Gamma(1/θ,1)).

II. Simulate n independent standard uniform variates V1, ..., Vn.

III. Return U = ((1-log(V1)/X)$^{-1}$/θ,..., (1-log(Vn)/X)$^{-1}$/θ)).

The modeling of random vector $X=(X_1, X_2, X_3)$ has been realised by using of MatLab programme. The algorithm of simulation of random vector $X=(X_1, X_2, X_3)$ with known ρ(Rho) correlation matrix is:

MatLab code:

```
n=5000;
Rho=[1 -0,417 -0,522; -0,417 1 0,420; -0,522 0,42 1];
Z=mvnrnd([0 0 0], Rho, n);
U=normcdf(Z,0,1);
X=[logninv(U(:,1),4.75,1.32)
logninv(U(:,2),2.53,0.55)
wblinv(U(:,3),10.24,1.16)];
plot3(X(:,1),X(:,2),X(:,3),'.');
grid on; view([-50, 50]);
xlabel('Izm'); ylabel('Norlzm'); zlabel('Laiks');
```

The illustration of the process of modelling of incidental value C= (C1, C2, C3) is presented in Fig. 4.

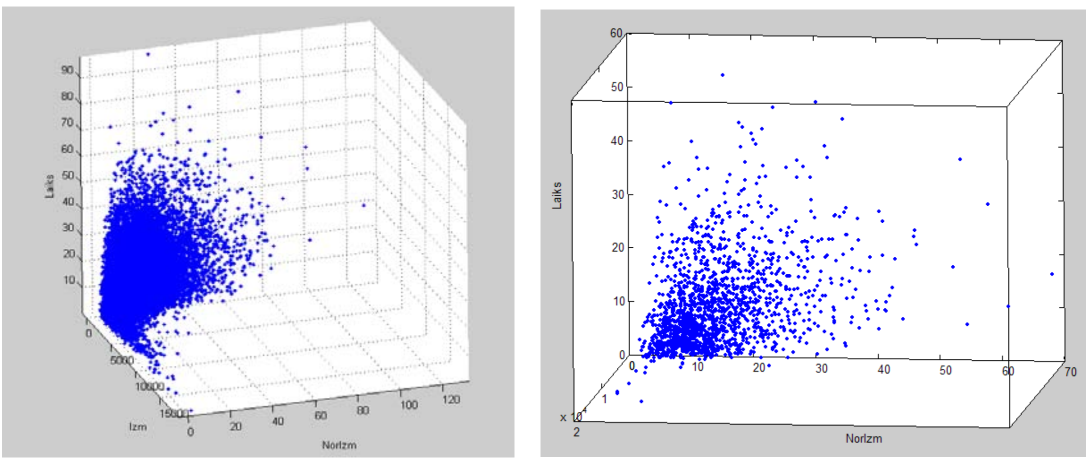

Fig. 4. Examples illustrating the process of modelling of incidental value X=(X1, X2, X3) with N=5000 and 50000 Monte-Carlo trials

Fig. 4 shows how different from the usual distribution the real joint distribution of three correlated random variables can be. In this case the nonparametric method of histograms is the most appropriate.

By means of a histogram, marginal distributions are represented for constructing a copula, presenting a common distribution of factors. In the simplest case, distribution of each incidental value may be represented by means of a nonparametric method – a block chart.

Conclusion

The research of the authors shows that unification of actuarial calculations after the creation of a software product that implements the application of Monte Carlo methods of statistical modeling to actuarial problems is possible. The application of Monte Carlo statistical methods is more natural and easier to deal with when solving urgent tasks of the insurance process. Setting objectives can be realized in a language close to the description of the real life insurance situation, which allows greater and more flexible practical application of methods of actuarial mathematics in real life. Methods for solving problems considered in this research can be applied in the training process for students of economic and engineering specialties.

It is obvious that analytical study of insurance processes described by functions of such kind may be performed only under some specific assumptions. It should be noted that in the real life insurance process the character of distributions of random variables is often not described by any known closed analytical distributions. In this case the Monte Carlo method can also be used to investigate the collective risk. Researches of the collective risk models conducted for educational purposes showed good agreement with those obtained by analytical methods (with the number of 50000-100000 Monte Carlo trials the relative error constitutes <5% -10%).

References

1. Clemen, R., Reilly, T. *Correlations and copulas for decision and risk analysis*. Management Science. 1999. vol. 45. pp. 208-224.
2. Embrechts, P., Lindskog, F., McNeil, A. *Modelling dependence with copulas and applications to risk management.* In: Handbook of Heavy Tailed Distributions in Finance. Ed. E. Rachev. 2003. Elsevier. pp. 329-384.
3. Glaz, J. *Approximations for the Multivariate Normal Distribution with Applications in Finance and Economics.* In: Applied Stochastic Models and Data Analysis. G. Govaert, J. Janssen and N. Limnios, eds. Universite de Technologie de Compiegne, Compiegne, France. 2001. Volume 1. 37-43.
4. Harrington, S., Niehaus, G. *Risk Management and Insurance.* New York. 2003. 672 p.
5. Jansons, V., Didenko, K., Jurenoks, V., *Insurance as a tool for steady development of agriculture.* VIII International scientific conference "Management and Sustainable Development". Bulgaria. 2006. p. 40 -45.
6. Jansons, V., Jurenoks, V., Didenko, K. *Investigation of Economic Systems using Modelling Methods with Copula.* The 13th International Conference on Harbor Maritime Multimodal Logistics Modelling & Simulation HMS 2010. October 13-15. Fez, Morocco. pp. 1-6.
7. Jurenoks, V., Jansons, V., Didenko, K., *Application of Benchmarking and Index Method in Research of Economic Systems.* X International Scientific Conference "Management and Sustainable Development". March 21-23, Yundola, Bulgaria. 2008. pp. 4-10.
8. Jurenoks, V., Jansons, V., Didenko, K., *Modelling of Stability of Economic Systems Using Benchmarking and Dynamic Programming*. X International Conference on Computer Modelling and Simulation EUROSIM/UKSim. 1-3 April, Cambridge, United Kingdom. 2008. pp. 295-300.
9. Jurenoks, V., Jansons, V., Didenko, K., *Investigation of Economic Systems using Modelling Methods with Copula*. XI International Conference on Computer Modelling and Simulation UKSim. March 25-27, Cambridge, United Kingdom. 2009. pp. 311- 316.
10. Lindskog, F. *Modelling Dependence with Copulas*, Master Thesis-MS-2000-06, Department of Mathematics, Royal Institute of Technology, Stockholm, Sweden. 2000.
11. *Par ieguldījumu pārvaldes sabiedr□bu un ieguldījumu fondu darb□bas rādītājiem 2011. gada 3. Ceturksnī*. [http://www.fktk.lv/lv/publikacijas/ pazinojumi_masu_informacijas_l/2012/2012-01-03_ par_ieguldijumu_parvalde/full_ver/].
12. Pettere, G., Jansons, V. *Stochastic analysis of insurance liabilities*. 9-th International Vilnius Conference on Probability Theory and Mathematical Statistics. 2006.
13. Ray, P. *A Practical Guide to Multi-risk Crop Insurance for Developing Countries.* U.S.A. Science Publishers Inc. 1991. 162 p.
14. Sklar, A., *Random variables, joint distributions, and copulas*. Kybernetica 9. 1973. pp. 449-460.
15. Корнилов, И. *Основы страховой математики.* М. Юнити-Дана. 2004.