 ISBN 978-80-87927-00-7, IISES

KONSTANTINS DIDENKO
Riga Technical University, Latvia

VITALIJS JURENOKS
Riga Technical University, Latvia

VLADIMIRS JANSONS
Riga Technical University, Latvia

VIKTORS NESPORS
Riga Technical University, Latvia

METHODOLOGY OF APPLICATION OF STATISTICAL MODELLING FOR RISK ASSESSMENT

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.

Keywords:

financial stability, risk statistical modelling, nonparametric methods

JEL Classification: C10, C14, C15

1. 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 modeling methods which due to the development of computer technologies are being more widely used in modeling insurance processes at all levels, starting with small insurance companies and ending with modeling 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 LVL[1] (448 million EUR). In 2010 insurance indemnities constituted 193.1 million LVL (274.8 million EUR). However, during the first two months of 2011 the volume of gross premiums totalled up to 31.46 million LVL (44.76 million EUR) 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 LVL (2.13 million EUR). Insurance companies gross premiums written by the 2011th the first three quarters year-on-year increase of 27.7% and were 176.8 million LVL, as well as the amount of gross claims paid increased by 12.1% and were 95.3 million LVL.
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.

2. Methodology of Investigation Using Statistical Modelling Method with Copula

Economic researches using statistical modelling methods have numerous challenges and opportunities in the waiting for the twenty-first century, calling for increasing numbers of non-traditional statistical approaches. Statistical modelling is one of the most widespread methods of research of economic systems. The selection of methods of modelling of the economic systems depends on a great number of conditions (modelling components) of the system being researched. The method of statistical modelling allows developing different scenarios of functioning of the investigated economic systems. Statistical modelling may be used for tackling a wide range of economic problems (design and analysis of industrial systems, stock management, balancing of production capacities, allocation of investment funds, optimization of investment funds, optimization of insurance system etc.). Modelling is frequently associated with the factor of uncertainty (or risk), who's description goes outside the confines of the traditional statistical modelling, which, in its turn, complicates the modelling process. However, most of the economic processes and systems are complex entities, consisting of a

[1] 1 LVL= 1.42 EUR

great number of interrelated subsystems (which in their turn also are complex objects and require a detailed study), changing their positions in space and time. For researching such economic systems it is impossible to create an absolutely accurate effective model by applying analytical methods. In such cases it is necessary to use the methods of modelling. In the process of modelling the most frequently method used to model multivariate distribution incidental values is the parametric method of modelling. In this case it is necessary to establish parameters of common distribution of incidental values characterizing the factors under consideration. Usually this is done by means of evaluation of parameters of multivariate distribution, i.e. by establishing the most suitable distribution (copula), deriving from the available empirical data. Copula like a tool for modelling different dependence structures more and more widely have been used in different fields of research: of economic, finance, insurance, risk theory. Copula functions are well studied object in the statistical literature. These functions have been introduced to model a joint distribution once the marginal distributions are known. When multivariate normal distribution is rejected by data, the copula may be used as an important alternative to represent the dependence in joint distributions. Copula theory makes it possible to approximate joint distribution of the significant factors of the economic system. On the basis of the obtained copula model it is possible to estimate the behaviour of the investigated economic system in relation with probabilities and therefore its expected values what is not possible to do with classical methods. When establishing the distribution of parameters describing the behaviour of investigated economic systems from empirical information most frequently is insufficient for a credible assessment of parameters offered by the function of distribution. In these cases it is necessary to use nonparametric modelling methods, given distribution of incidental values and then modelling parameters of distribution. The main objectives of the paper are:

- to describe the main ideas of using of statistical modelling methods for investigation of economic systems;
- to use statistical modelling method with copula for calculation of optimal insurance premium;
- to combine different marginal distributions using statistical modelling methods;
- to evaluate and model the risks of cereals production.

The technique of modelling using copula

In the real world there is often a non-linear dependence between different variables and correlation cannot be an appropriate measure of co-dependency. Therefore linear Spearmen's correlation coefficient is a limited measure of dependence[2]. It is not surprising that alternative methods (the copula method) for capturing co-dependency have been considered. The concept of copulas comes from Sklar[3] in 1959. Let us consider mathematical construction of copula.
Let F be a joint multivariate distribution with marginals F_1 and F_2. Then, for any x_1, x_2 there exists a copula C such (1):

$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2)) \tag{1}$$

Furthermore, if marginals F_1 and F_2 are continuous, the copula C is unique. Conversely, if F_1 and F_2 are marginal distributions and C is a copula, then the function F defined by

[2] Lindskog F., 2000, Modelling Dependence with Copulas, *Department of Mathematics, Royal Institute of Technology, Stockholm, Sweden*, 1-34.

[3] Sklar, A., 1973, Random variables, joint distributions, and copulas, *Kybernetica 9*, 449-460.

$C(F_1(x_1), F_2(x_2))$ is a joint distribution function with marginals F_1 and F_2. If we have a random vector X = (X_1, X_2) the copula of their joint distribution function may be extracted from equation (1):

$$C(u_1, u_2) = F(F_1^{-1}(u_1), F_2^{-1}(u_2)) \tag{2}$$

where the F_1^{-1}, F_2^{-1} are the quantile functions of the margins. Thus a copula is a function C:

$$C : [0,1]^2 \rightarrow [0,1] \tag{3}$$

Recently a lot of works have devoted to the modelling of the joint distribution and to see how these models could be better than the joint normal specification, for representing the tail behaviour. In rough terms, a n-dimensional copula is a function:

$$C : [0,1]^n \rightarrow [0,1] \tag{4}$$

with certain special properties. Alternatively we can say that it is a multivariate distribution function defined on the unit cube $[0,1]^n$. Copula functions are well studied object in the statistical literature. These functions have been introduced to model a joint distribution once the marginal distributions are known. When multivariate normal distribution is rejected by data, the copula may be used as an important alternative to represent the dependence in joint distributions. Copulas are used to combine different marginal distributions. They are unique, if marginal distributions are continuous, and like dependence measures use Kendall's tau and Spearman's rho, which are invariant under strictly increasing transformation. Therefore copulas have become a powerful tool for modelling dependence between random variables. Also copula methodology is effective for modelling joint distributions with fat tails. Fat tails in financial return data have been documented in numerous real cases.

Tails dependence meaningful measure of dependence. For example, real portfolio examples clearly show that we must use heavy tailed alternatives to the Gaussian law in order to obtain acceptable estimates of market losses. But can we substitute the Gaussian distribution with other distributions in Value at Risk (Expected Shortfall) calculations for whole portfolios of assets? Remember, that the definition of VaR utilizes the quartiles of the portfolio returns distribution and not the returns distribution of individual assets in the portfolio. If all asset return distributions are assumed to be Gaussian then the portfolio distribution is multivariate normal and we can apply well known statistical tool. However, when asset returns are distributed according to different laws then the multivariate distribution may not be multivariate normal. In particular, linear correlation may no longer be a meaningful measure of dependence.

In such cases multivariate statistics offers the concept of copulas. The technical definitions of copulas that can be found in the literature often look more complicated, but to a financial modelling, this definition is enough to build an intuition from. What is important for VaR calculations is that a copula enables us to construct a multivariate distribution function from the marginal (possibly different) distribution functions of n individual asset returns in a way that takes their dependence structure into account. This dependence structure may be no longer measured by correlation, but by other adequate functions like rank correlation and, especially, tail dependence. Moreover, it can be shown that for every multivariate distribution function there exists a copula which contains all information on dependence.

Elliptical and Archimedean copulas. Copula functions do not impose any restrictions on the financial model, so in order to reach a model that is to be useful in a given risk management problem, a particular specification of the copula must be chosen. From the wide variety of copulas that exist, probably the elliptical and Archimedean copulas are most often used in applications. Most widespread elliptical copulas are Gaussian and Student's copulas[4]. The Gaussian copula is given by:

$$C_\rho(u,v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} a \cdot \exp\left(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right) dxdy \tag{5}$$

where ρ is the parameter of the copula, and Φ^{-1} is the inverse of the standard univariate Gaussian distribution function and $\alpha = 1/(2\pi(1-\rho^2))^{1/2}$.

Student's t-copula. The Student's t-copula allows for joint fat tails and an increased probability of joint extreme events compared with the Gaussian copula. This copula can be written as:

$$C_{\rho,v}(u,v) = \int_{-\infty}^{t_v^{-1}(u)} \int_{-\infty}^{t_v^{-1}(v)} a \cdot \left(1 + \frac{s^2 - 2\rho st + t^2}{v(1-\rho^2)}\right)^{-\frac{v+2}{2}} dsdt \tag{6}$$

where ρ and v is the parameter of the copula, and tv-1 is the inverse of the standard univariate t-distribution function with v degrees of freedom, expectation 0 and variance v/(v-2). Rank correlation and tail dependence coefficients can be easily calculated for elliptical copulas. There are, however, drawbacks - elliptical copulas do not have closed form expressions, are restricted to have radial symmetry and have all marginal distributions of the same type. These restrictions may disqualify elliptical copulas from being used in some risk management problems. In particular, there is usually a stronger dependence between big losses (e.g. market crashes) than between big gains. Clearly, such asymmetries cannot be modelled with elliptical copulas. In contrast to elliptical copulas, all commonly encountered Archimedean copulas have closed form expressions. Their popularity also comes from the fact that they allow for a great variety of different dependence structures. Many interesting parametric families of copulas are Archimedean, including the well-known Clayton, Frank and Gumbel copulas.

After the marginal distributions are estimated and a particular copula type is selected, the copula parameters have to be estimated. The fit can be performed by least squares or maximum likelihood. In most cases we also have to use Monte Carlo simulations.

Archimedean copula. Let us consider a function φ: [0;1]→[0;1] (Archimedean copula's generator) which is continuous, strictly decreasing, convex and for which φ(0) = ∞ and φ(1)=0. We then define the pseudo inverse of φ[-1]: [0; ∞] → [0;1] such that:

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t) & 0 \le t \le \varphi(0) \\ 0 & \varphi(0) \le t \le \infty \end{cases} \tag{7}$$

[4] Glaz, J., 2001, 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.* Volume 1, 37-43.

As φ is convex, the function C: $[0; 1]^n \rightarrow [0; 1]$ defined as:

$$C(u_1, ..., u_n) = \varphi^{-1}[\varphi(u_1) +. . . + \varphi(u^n)], \qquad (8)$$

is an n-dimensional Archimedean copula if
and only if φ^{-1} is completely monotone on [0, ∞).

Copula-based dependence measures.

Since the copula of a multivariate distribution describes its dependence structure, it might be appropriate to use measures of dependence which are copula-based[5]. The bivariate concordance measures Kendall's tau and Spearman's rho, as well as the coefficient of tail dependence, can, as opposed to the linear correlation coefficient, be expressed in terms of the underlying copula alone. Since copulas are used to represent the dependence structure among the variables when margins are known or well estimated, it is useful to describe the basic dependence measures, which can be used to interpret the parameters appearing in parametric copula families.

Kendall’s tau – dependency measure.

Kendall’s tau of two variables X and Y is (in terms of copulas functions):

$$\tau(X,Y) = 4\int_{[0,1]^2} C(u,v)dC(u,v) - 1 \qquad (9)$$

where C(u, v) is the copula of the bivariate distribution function of X and Y. For the Gaussian and Student’s t-copulas and also all other elliptical copulas, the relationship between the linear correlation coefficient cor(X, Y) and Kendall’s tau is given by:

$$cor(X,Y) = \sin\left(\frac{\pi}{2}\tau\right) \qquad (10)$$

For an Archimedean copula τ(X, Y) can be evaluated directly from the generator of the copula:

$$\tau(X,Y) = 1 + 4\int_0^1 \frac{\varphi(t)}{\varphi'(t)} dt \qquad (11)$$

Spearman’s correlation.

The concept of linear correlation can also be applied to transformation of the basic variables. For instance, we can consider the correlation between the ranking associated with X and Y, that are U = F(X), V = G(Y). The correlation between the ranks, so-called Spearman’s rho is

$$S_\rho = Cor(U,V) = Cor(F(X), G(Y)) \qquad (12)$$

It depends on the copula only and is given by:

---

5 Embrechts, P., Lindskog, F., McNeil, A., 2003, Modelling dependence with copulas and applications to risk management. *In: Handbook of Heavy Tailed Distributions in Finance, Ed. E. Rachev, Elsevier*, 329-384.

$$S_{\rho}(X,Y) = 12\int_{[0,1]^2} C(u,v)dudv - 3 \quad (13)$$

where C(u, v) is the copula of the bivariate distribution function of X and Y. Let X and Y have distribution functions F and G, respectively. Then, we have the following relationship between Spearman's rho and the linear correlation coefficient

$$S_{\rho}(X,Y) = cor(F(X), F(Y)) \quad (14)$$

Simulating from copulas (common case). If in addition to equation (7) function ϕ equals the inverse of Laplase transform of a distribution function G on R^+ satisfying G(0) = 0, the following algorithm can be used for simulating from the copula:

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

II. Simulate n independent variates $V_1,\ldots, V_n$.

III. Return U = $(\phi^{-1}(-\log(V_1)/X),\ldots, (\phi^{-1}(-\log(V_n)/X))$.

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

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

II. Simulate n independent standart uniforms variates $V_1, \ldots, V_n$.

III. Return U = $((1-\log(V1)/X)^{-1/\theta},\ldots, (1-\log(V_n)/X)^{-1/\theta}))$.

For incidental factors under consideration characterising the logistics process in nonparametric modelling, discrete distribution is used represented in the form of histogram and correlation dependences. By means of histogram, marginal distributions are represented for constructing a copula, presenting a common distribution of factors.

3. Stochastic Modelling in Insurance

3.1. Example (1) – Modelling of Simple Insurance Fund in Agriculture

The research considers the possibilities of using modelling to evaluate the size of tariff and stability of the insurance portfolio. The case referred to illustrates the insurance of one of the agricultural risks – insurance of cereal sowings, using real data in Latvia. According to the classical theory of insurance of cereal sowings, the insurance object is insurance of revenues from cereal crops. One of problems of calculation of the insurance premium is definition of character of dependence between risks of a crop and price of agricultural product. It is difficult to calculate insurance premium for insurance of an expected crop with use of traditional methods of processing of the statistical information. Therefore in work the copula method for statistical modelling of multivariate dependences between risks of a crop and price of agricultural product is used.

The modelling approach ranges from non-parametric to parametric methods. Cobel K., Heifner R. and Zuniga M.[6] have investigated these correlations in revenue insurance and found that there is strong correlation between price and farm yield and between farm and national yield in certain crops and regions.
The international practice of agricultural risks insurance[7,8] shows that calculations of formation of insurance services use several options of obtaining statistical data, where data are split into three data groups, according to the phases of development of the service, as follows:
- statistical data for establishing the insurance coverage;
- statistical data for calculating the insurance premium;
- statistical data for calculating insurance indemnity.

The example shows the possibilities:
- to establish the insurance coverage for cereal sowings insurance process;
- to evaluate insurance tariffs and the insurance premium;
- to evaluate the dependence structure between the price and yield risks.

Let us consider the modelling scheme of the agricultural insurance fund, which later will allow us to model the process of developing the model and to establish the minimum amount of the insurance fund U (without a state subsidy). The minimum fund amount U guarantees that with certainty γ agricultural losses will be compensated. For modelling the insurance fund, we will use the simplest individual risk modelling scheme. Let us assume that the whole farm insurance fund is satisfactory, given the following conditions:
- the number of registered farms in the fund is constant;
- risks of individual farms are independent;
- payment of premiums is effected at the beginning of the period.

The loss distribution function is equal for all farms.
Let us designate that:
n – number of agreements in the fund;
j – ordinal number of the farm;
p – probability of setting in of the insurance event;
Y_j – possible losses of the farm j. Value Y_j has probability distribution function F(x);
X_j – satisfied loss of the farm j. $Xj = Ind_j * Y_j$;
Ind_j - binary index of the insurance event of the farm j.

By using variable Ind, we can calculate total number N of farms incurring losses:

$$N = \sum_{j=1}^{n} Ind_j \tag{15}$$

Total amount of losses is:

[6] Coble, K., R. Heifner, and M. Zuniga, 2000, Implications of Crop Yield and Revenue Insurance for Producer Hedging, *Journal of Agricultural and Resource Economics*, 432 - 452.

[7] Ray P.K., 1991, A Practical Guide to Multi-risk Crop Insurance for Developing Countries. *U.S.A.: Science Publishers, Inc.*, 10-162.

[8] Jansons V., Graudiņa A. 2006, Underlying Principles for Establishing Crop Insurances Services. *Int. Finance Symposium, May 25-26, 2006, Turkey*.

$$Z = X_1 + X_2 + ... + X_n \quad (16)$$

or by using indices of setting in of the events:

$$Z = Ind_1 \cdot Y_1 + Ind_2 \cdot Y_2 + ... + Ind_n Y_n = \sum_{j=1}^{n} Ind_j \cdot Y_j \quad (17)$$

We are to compensate losses to farms with a certainty γ and are to ensure the required operation of the fund with cash funds L. It means that the amount of the fund after compensations must be positive with a certainty $P(U - Z \geq 0) = \gamma$. The degree of risk of the insurance fund can be established by the variation coefficient:

$$K_{var}(Z) = \frac{\sigma(Z)}{E(Z)} = \frac{\sqrt{D(Z)}}{E(Z)} \quad (18)$$

where $\sigma(Z)$ - standard deviation from the amount Z (standard error);
E(Z) - mathematical expectation of value Z, which in practice is measured with average value of Z;
D(Z) - variation of value Z.

Figure 1. Illustration of process of total loss formation

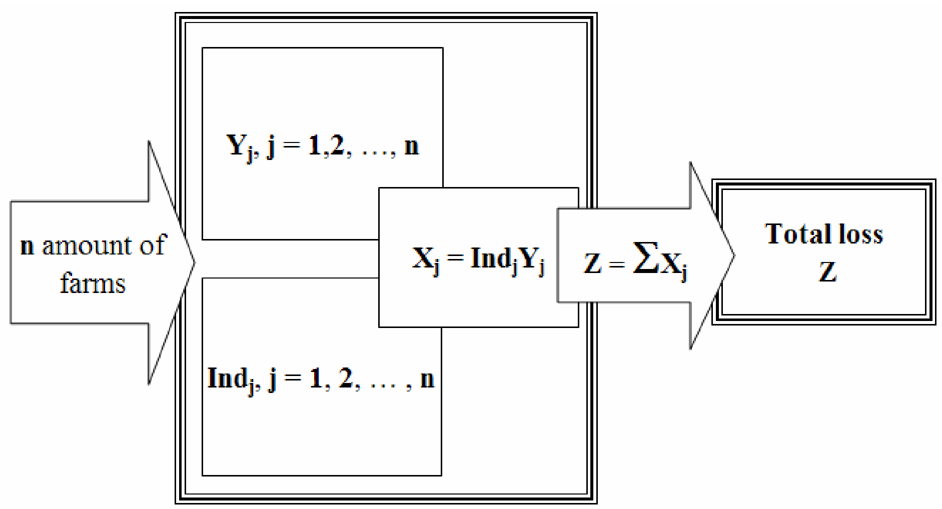


Figure 1 shows that total loss Z is formed in n farms during one time period. Yield risks X_1, X_2, …, X_n can be modelled by a family of Beta distributions (19), whereas price shocks can be modelled by log-normal distributions:

$$f(x) = \frac{x^{n1-1}(1-x)^{n2-1}}{\int_0^1 u^{n1-1}(1-u)^{n2-1} du} \quad (19)$$

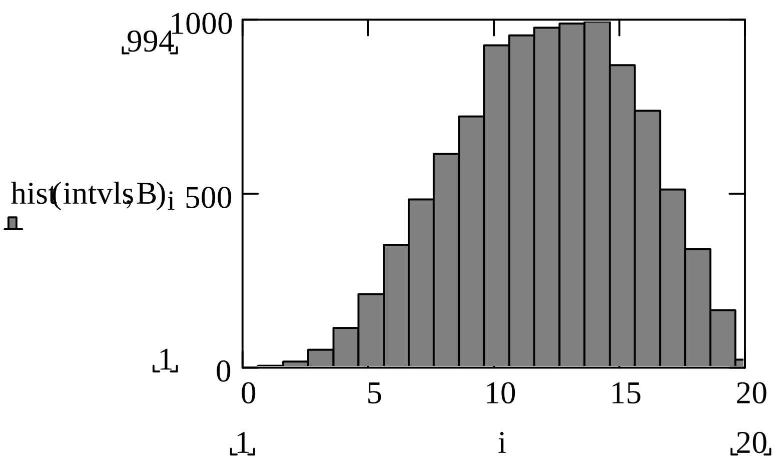
with parameters n_1 and n_2 in programme MathCad according of algorithm (20):

$$\mathrm{ORIGIN} := 1$$

$$
\begin{aligned}
B(n1,n2,N) := \;\; & u1 \leftarrow \mathrm{runif}(n1\cdot N,0,1) \\
& u2 \leftarrow \mathrm{runif}(n2\cdot N,0,1) \\
& \text{for } k \in 1..N \\
& \quad \text{for } i \in 1..n1 \\
& \qquad U1_i \leftarrow u1_{(k-1)\cdot n1+i} \\
& \quad \text{for } j \in 1..n2 \\
& \qquad U2_j \leftarrow u2_{(k-1)\cdot n2+j} \\
& \quad B_k \leftarrow \frac{-\log\left(\prod_{i=1}^{n1} U1_i\right)}{-\log\left(\prod_{i=1}^{n1} U1_i\right) - \log\left(\prod_{j=1}^{n2} U2_j\right)} \\
& B
\end{aligned}
\qquad (20)
$$

In most cases distributions of values X_i (i=1,2,…,n) have asymmetrical distributions (see Figure 2).

Figure 2. Illustration of distributions of values X_i



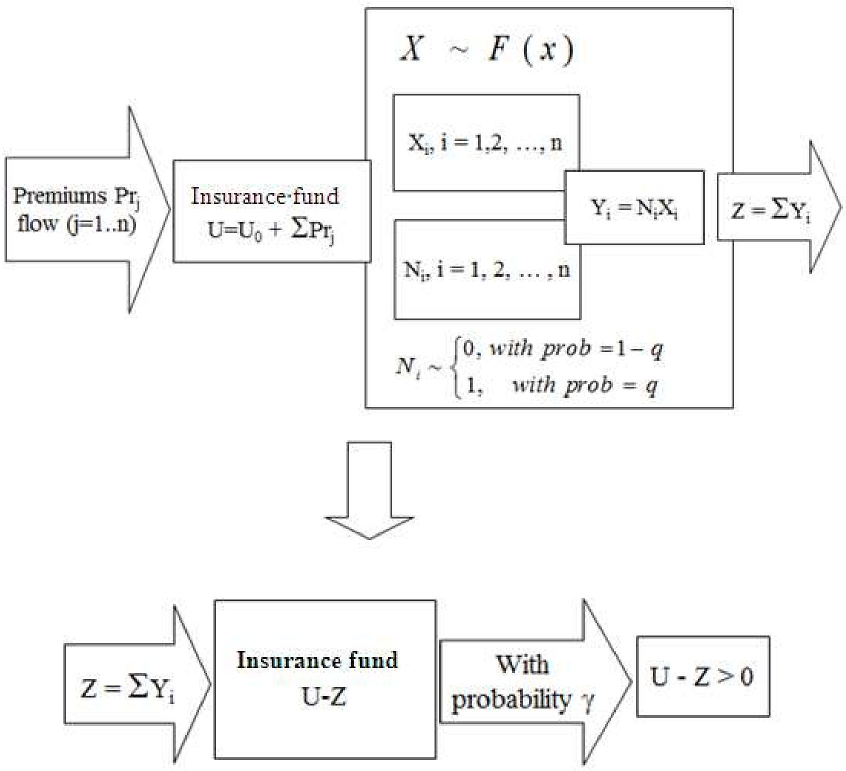
The amount of the insurance coverage in cereal sowings insurance depends on the average amount of crop received by years, in which no relevant losses took place. The results of modelling without price risk show that very often variation coefficient of insurance fund K_{var} fluctuates within the range from 20% to 50%, which testifies to the fact that insurance fund, is often not so stable and additional financing is required from the state. In the age of up-to-date information technologies the application of the Monte-Carlo statistical method using copula is sufficiently simple method and frequently allows avoiding from complicated theoretical calculations as well as allows obtaining sufficiently accurate practical results to take appropriate decisions on insurance parameters.

In order to develop a multivariate model that preserves a given set of marginals, a copula approach can be used to characterize the joint yield and price risk of cereals production, which are usually highly correlated. Results show that farm insurance, with use of new methodology of an estimation of risks, manages to farmers approximately on 20 % more cheaply, than it was earlier. It shows that accuracy and efficiency of modelling of a grain yield, and also modelling of price correlations and contracts of the insurance of a farm are improved using of multidimensional distributions – copula. Results of research are important for carrying out of the best risk-management and reduction of losses of manufacturers of agricultural production from various risks of manufacture of a crop.

3.2. Example (2) – Portfolio Risk Statistical 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 example focuses on the development of economic and mathematical models of non-life insurance and their statistical modeling 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 Figure 3).

Figure 3. Scheme of structure of individual risk model



where 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 modeling 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 Figure 4):

Figure 4. Insurance portfolio modeling 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:

$$U - Z = U_0 + \sum_{j=1}^{n} P_j - \sum_{j=1}^{n} N_j \cdot X_j > 0 \qquad (21)$$

with the given probability γ (usually γ = 0.1 or 0.05). Since U0 $\geq$ 0, the inequality (1) will follow from the inequality:

$$\sum_{j=1}^{n} P_j - \sum_{j=1}^{n} N_j \cdot X_j > 0 \qquad (22)$$

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 (21) as well as the inequality (22) hold. The value C shows the required level of aggregate premiums ensuring the stability of the insurance company with probability γ.

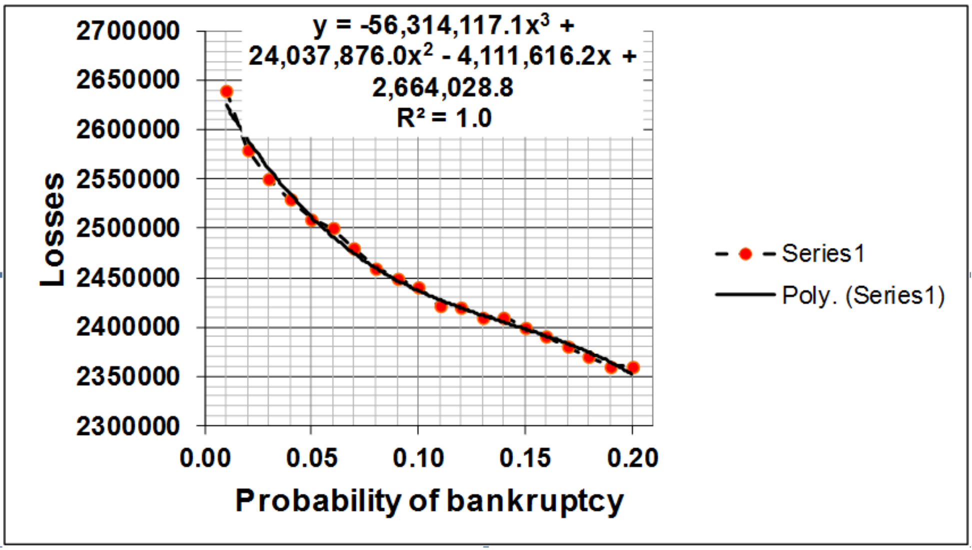
$$C = F^{-1}(1-\gamma) \qquad (23)$$

In many real life situations the analytical solution of the equation (23) 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=10000Ls, S_2=20000Ls, p=0.02 due to reinsurance when C=16000Ls, the company seeks to reduce the probability of ruin from 0.14 to 0.13;

b) having applied for modeling 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.

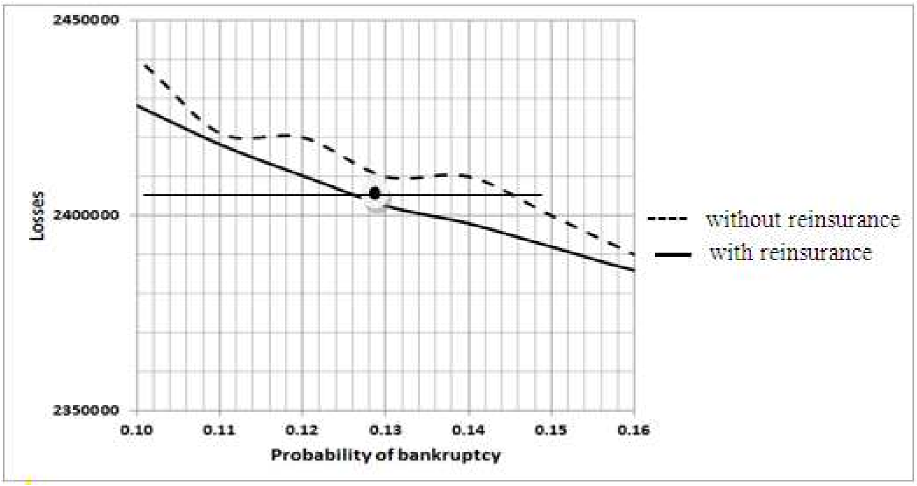
Figure 5. Ruin probability of the insurance portfolio without reinsurance – the dotted line and polynomial approximation – solid line



After the introduction of the reinsurance process the probability of bankruptcy decreased from

0.15 to 0.13 (see Figure 6). 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. Modeling ensures quicker and easier adaptation to various changes in the insurance situations which is practically very difficult to reach using analytical methods.

Figure 6. Ruin probabilities of the insurance portfolio without reinsurance – the dotted line and with reinsurance – the solid line



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):

$$F(x,t) = \sum_{0}^{\infty} q(n,t) G^{(n)}(x) \qquad (24)$$

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 $X_1, X_2,\ldots, X_n$ can be described by equation:

$$C(u_1,u_2,\ldots,u_n) = \Phi(F_1^{-1}(x_1), F_2^{-1}(x_2),\ldots, F_n^{-1}(x_n)), \qquad (25)$$

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 = $(\phi^{-1}(-\log(V1)/X),..., (\phi^{-1}(-\log(Vn)/X))$.

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 = $((1-\log(V1)/X)^{-1}/\theta,..., (1-\log(Vn)/X)^{-1}/\theta))$.

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('NorIzm'); zlabel('Laiks');
```

The illustration of the process of modelling of incidental value C= (C1, C2, C3) is presented in Figure 7.

Figure 7. Examples illustrating the process of modeling of incidental value X=(X1, X2, X3) with N=5000 and 50000 Monte-Carlo trials

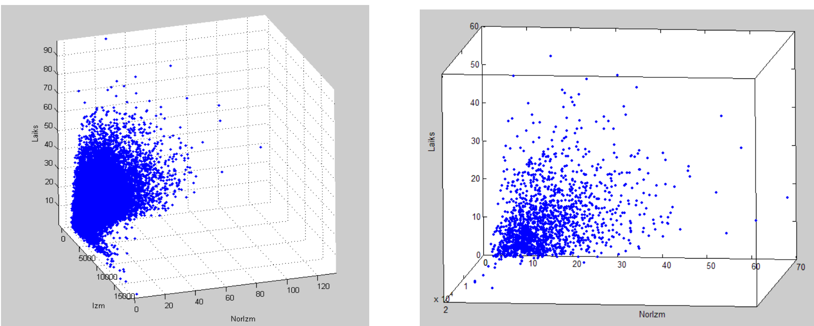


Figure 7 shows how different from the usual distribution the real joint distribution of three correlated random variables can be. In this case the nonparametric method of histograms is the most appropriate. By means of a histogram, marginal distributions are represented for constructing a copula, presenting 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.

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%).

3.3. Example (3) – Statistical Modelling and Estimation of Fire Insurance Parameters for Residential Buildings in Latvia

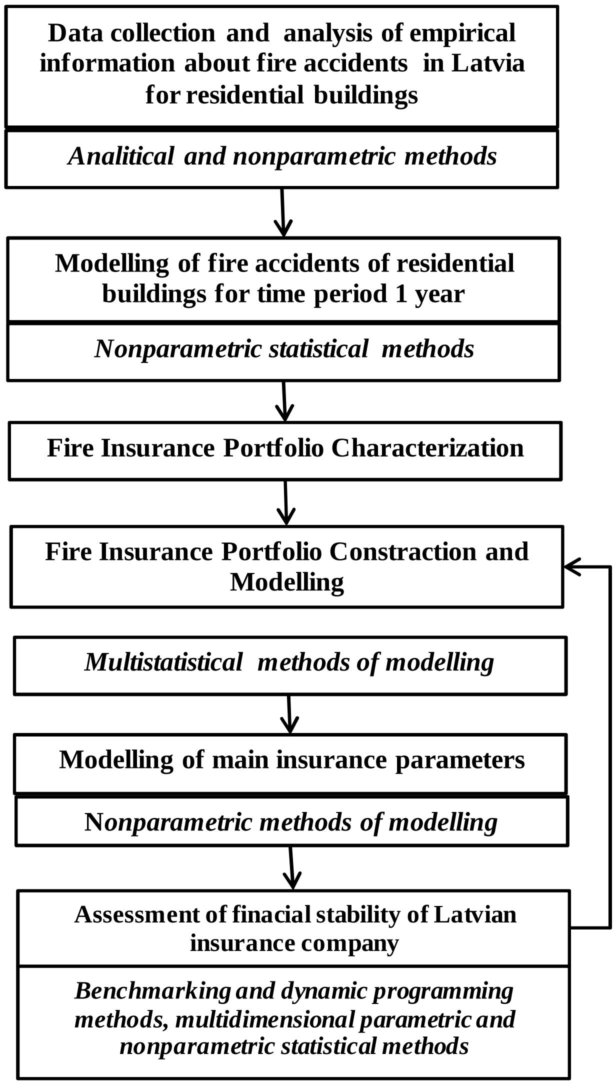
Approach to task modelling

To model the behavior of complex systems, wide use is made of non-parametric methods. These methods do not impose initial restrictions on the functional type of distribution functions. By means of the non-parametric approach it is possible to avoid specification (establishment of parameters of the system being investigated) of the model, which is typically encountered in the parametric approach. Thus the non-parametric method[9] allows a wide variety of types of non-linear behavior of the system, which are most frequently observed in the behavior of real insurance systems[10]. The common scheme of modelling process is presented in Figure 8.

[9] Goodwin, B. K., and A.P. Ker, 1998, Nonparametric Estimation of Crop Yield Distributions: Implications for Rating Group-Risk(GRP) Crop Insurance Contracts, *American Journal of Agricultural Economics,* 139- 153.

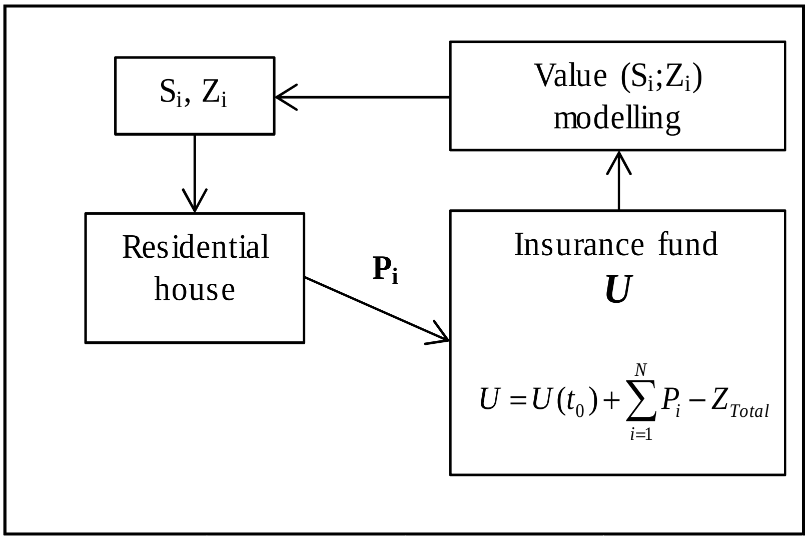
[10] Jansons, V., Didenko, K., Jurenoks, V., 2006, Insurance as a tool for steady development of agriculture. *VIII International scientific conference "Management and Sustainable Development", Bulgaria,* 40 -45.

Figure 8. Scheme of Modelling of Insurance Process



Insurance process of residential buildings in Latvia consists of financial flows of premiums and losses, just as suggested the estimation of the main parameters of the insurance. The common model of financial flows in insurance process is presented in Figure 9.

Figure 9. Model of Financial Flows in Insurance Process



where S_i – sum insured;
P_i – insurance premium;
Z_i – insurance indemnity;
U – insurance fund;
t_0 – initial time of insurance process;
Z_{Total} - general & administrative expenses;
N – number of insurance objects.

Very important for insurance process is possibility of performance of insurance obligations irrespective of intensity of a stream of insured events $\{t_j, Z_j\}$ and sizes of insurance payments $\{Z_j\}$ and also general & administrative expenses Z_{Total}. Insurance process should be financially stable during all time of functioning of insurance system. Steady financial position of insurance company is understood as the ability of the company to perform a complete set of functions and also maintain (or even increase) the services to be rendered for a long period of time in the conditions of uncertainty. Stability of insurance process we understand as performance of an inequality (1) during all time of functioning of insurance process with probability 1-α.

$$U - Z - Z_{Total} > 0 \qquad (26)$$

Parameter α is enough small - usually 5% or 1%. The application of modelling is connected with the fact that frequently it is not possible to provide a definite parametric description of the behavior of the economic system being investigated. Uncertainty in insurance process is understood as a situation when there is incomplete or no information at all about the possible conditions of the process itself and the environment in which the insurance company functions. Conditions of uncertainty are understood as various random fluctuations of factors of insurance process environment. In Latvia, taking into account short history of insurance (2008-2012), the amount of empirical information characterizing the process of formation of main insurance parameters, such as premium P is insufficient for establishing the dependence of insurance parameters on the factors characterizing insurance process. For modeling insurance process of residential buildings in Latvia authors have used empirical information, which is presented in Table 1.

Table 1. Data describing insurance process for various parameters characterizing residential buildings

Fire time	Sum insured, S	insurance indemnity, Z	Construction of house	Construction time	Area	Electrosystem	Heating system	Whether constantly live
2008	8900	631.8	C	1932	55	E	N	Yes
2008	18000	588	B	1975	120	F	O	No
2008	396000	3835.08	A	2004	792	D	L	Yes
2008	12500	179.17	B	1976	82	F	K	Yes
2008	325000	3468.93	A	2000	445	F	L	Yes
………………								
2012	12 044.00	6 480.00	C	1946	84	E	N	Yes
2012	12 044.00	810.00	C	1946	84	E	N	Yes
2012	21 000.00	50.00	A	1989	83	F	K	Yes
2012	27 010.00	680.00	A	1989	55	F	K	Yes

where A - stone houses; B - mixed construction of houses; C - wood houses; S - sum insured; Z - insurance indemnity; S_A - sum insured for stone houses; S_B - sum insured for houses with mixed construction (stone and wood); S_C - sum insured for wood houses; Z_A - insurance indemnity for stone houses; Z_B - insurance indemnity for houses with mixed construction (stone and wood); Z_C - insurance indemnity for wood houses; D – was repair of electricity within 5 years; E - fully new electricity installation; F – electricity had not been changed; K - central warming system; L – gas warming system; M - electrical warming system; N – using the furnace for warming; O – no warming system.

Modelling of Insurance Parameters

For modeling and researching of insurance process of residential buildings in Latvia it is impossible to create an appropriate effective model by applying analytical methods. In such cases it is necessary to use the methods of simulation modeling in combination with nonparametric methods. With Monte-Carlo method we can simulate any quantity of random variables, using initial empirical distributions of S_i, $Z_{i,}$ Z_{Total} and inverse transformation method. For generating values S_i and Z_i according to joint empirical distribution it's necessary to use two marginal empirical distributions for values S_i and Z_i. Here are two possible methods of modeling joint distribution for incidental vector (S, Z):
- using histogram (nonparametric) method, modeling values S and Z without parameterization
– copula method for modeling multivariate distributions;

- using marginal distributions functions for every value S_i and Z_i - parametric method (see Table 2).

Table 2. Initial empirical tables of marginal distributions of S_i and Z_i

S_C			Z_C	
Cumulative %	*Bin*		*Cumulative %*	*Bin*
0.00%	1700		0.00%	40
2.70%	13590		2.70%	9592.17
75.68%	25480		89.19%	19144.3
89.19%	37370		97.30%	28696.5
89.19%	49260		97.30%	38248.7
94.59%	61150		97.30%	47800.8
97.30%	62000		97.30%	48000

The illustration of the process of modelling of Values S, Z is presented in Figure 7. With Monte-Carlo method we can simulate any quantity of needed random variables, using initial empirical distributions of values S, Z. After that it is possible to calculate probability of a fire for each type of areas as:

$$P = \frac{S_{fire}}{S_{Total}} \tag{27}$$

In case of residential buildings this probability P is equal = 0.001471 (S_{fire}=S1+S2+S3, S_{Total}=122,2Mm^2).
Distribution of fires by various kinds of the property is presented in the Table 3.

Table 3. Distribution of Fires by Various Kinds of the Property

	Year			
	2008	**2009**	**2010**	**2011**
Total area	99965,75	100535,9	82467,11	80553,94
S1	76778,5	75173,43	60143,91	63896,77
S2	8923,14	8507,49	8326,06	5178,00
S3	853,55	345,36	924,82	71,3

where S_1 – total area of private property;
S_2 - total area of property of self-managements;
S_3 - total area of state ownership.
For calculation of quantity of standard units of area N_{2011} in 2011 year can be used the formula:

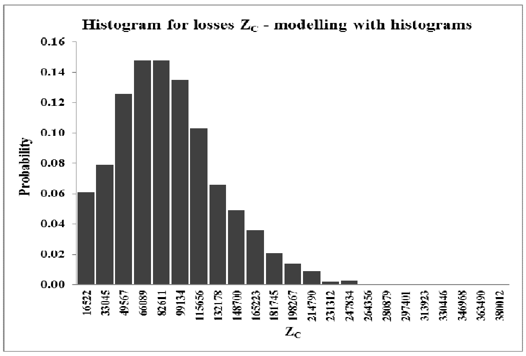
$$N_{2011} = \frac{S_{Total}}{S_0} \tag{28}$$

As a result we will receive (S_{Total} = 122.2M m^2, $S_0 = S_{Apartment\_average}$ = 82.04 m^2): $N_0 = N_{2011}$ = 1489517 units. It is approximately N_0 equal 1500000 units.
By means of a method of Monte Karlo $N_{M\text{-}C} = N_0/10$=15000 times we will simulate a vector (S, Z) and calculate the general financial losses as a result of fires using the formula:

$$Z_{j,Total} = \sum_{i=1}^{N_{M-C}} i_j(S_{j,i})S_{j,i} \quad (29)$$

where j – take values A, B, C; $i(S_j)$ coefficient characterizing connection between Sj and Z_j at various values of size S_j. For receiving distribution for value of total losses caused by fire Z_C it is necessary to model value Z_{Total} M times (in our case M = 3000) using Monte Carlo method. The histogram received for losses Z_C is shown in Figure 10.

Figure 10. Histogram for losses Z_C received by Monte-Carlo method using Histograms method



Using histograms methods for S_C and Z_C it has been received:
P_0 = 6.41 LVL;
P_1 = 10.97 LVL.
95% confidence intervals for premium P_0 and P_1 for wood buildings are:

$$\overline{P_0} - \Delta_0 < P_0 < \overline{P_0} + \Delta_0, \quad (30)$$
$$\overline{P_1} - \Delta_1 < P_1 < \overline{P_1} + \Delta_1$$

where Δ_0 = 0.21LVL, Δ_1 = 0.25 LVL.

Conclusion

The results of investigation describing in the paper shows that the copula approach has been spurred by the recent developments in the whole farm insurance, resulting in an increasing need for the modelling of multivariate risk factors and their interaction. Finally, the proposed copula approach is illustrated with simulated data to calculate the premium rate of the whole farm insurance. 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. After establishing the character of behavior of factors describing the characteristics of the system being researched using histograms, it is possible to undertake its imitation modelling. In insurance process a great number of participants are involved and linked in a unified insurance system and for investigating such type of complex system there is one effective method - simulating such system using Monte Carlo method.

The use of the imitation modelling allows:

1) to model important parameters of insurance process which allow to characterize the level and sensitivity of the financial stability of insurance systems;

2) to identify with probability $P = 1 - \alpha$ the amount of the insurance fund required for insurance system stability;

3) to model the behavior of financial stability in insurance system;

4) to enhance the performance efficiency of insurance system.

The modern economic analysis basing on the using of information technologies shows that in the real systems the parameters describing the economic objects, not always have the well-known Gauss distribution. The nonlinear dependence exists between various factors in our case between factors SC and ZC. In these cases it is impossible to use the linear correlation coefficient for evaluation of measure of dependences between factors. It requires using other methods for evaluation the measure of dependences between factors which had been used in the paper. The theoretical and practical results of this research can be applied for evaluation of premium values for different scenarios of insurance process in conditions of uncertainty.

References

Clemen, R. T., Reilly, T., 1999, Correlations and copulas for decision and risk analysis. *Management Science*, vol. 45, 208-224.

Coble, K., R. Heifner, and M. Zuniga, 2000, Implications of Crop Yield and Revenue Insurance for Producer Hedging, *Journal of Agricultural and Resource Economics*, 432 - 452.

Embrechts, P., Lindskog, F., McNeil, A., 2003, Modelling dependence with copulas and applications to risk management. *In: Handbook of Heavy Tailed Distributions in Finance, Ed. E. Rachev, Elsevier*, 329-384.

Glaz, J., 2001, 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.* Volume 1, 37-43.

Goodwin, B. K., and A.P. Ker, 1998, Nonparametric Estimation of Crop Yield Distributions: Implications for Rating Group-Risk(GRP) Crop Insurance Contracts, *American Journal of Agricultural Economics*, 139- 153.

Harrington S.E., Niehaus G.R., 2003, Risk Management and Insurance. *New York*, 672 p.

Jansons, V., Didenko, K., Jurenoks, V., 2006, Insurance as a tool for steady development of agriculture. *VIII International scientific conference "Management and Sustainable Development", Bulgaria*, 40 -45.

Jurenoks, V., Jansons, V., Didenko, K., 2008, Application of Benchmarking and Index Method in Research of Economic Systems. *X International Scientific Conference "Management and Sustainable Development", Yundola, Bulgaria*, 4-10.

Jurenoks, V., Jansons, V., Didenko, K., 2008, Modelling of Stability of Economic Systems Using Benchmarking and Dynamic Programming. *X International Conference on Computer Modelling and Simulation EUROSIM/UKSim, Cambridge, United Kingdom*, 295-300.

Jurenoks, V., Jansons, V., Didenko, K., 2009, Investigation of Economic Systems using Modelling Methods with Copula. *XI International Conference on Computer Modelling and Simulation UKSim, Cambridge, United Kingdom*, 311- 316.

Jansons, V., Jurenoks, V., Didenko, K., 2010, Investigation of Economic Systems using Modelling Methods with Copula. *The 13th International Conference on Harbour Maritime Multimodal Logistics Modelling & Simulation HMS, Fez, Morocco*, 1-6.

Lindskog F., 2000, Modelling Dependence with Copulas, *Department of Mathematics, Royal Institute of Technology, Stockholm, Sweden*, 1-34.

Pettere, G., Jansons, V., 2005, The best method for modelling a two-risk negative correlation portfolio, *Economics and Business, Scientific Proceedings of Riga Technical University 3 sērija, 10 sējums, RTU izdevn., Rīga, Latvia*, 90-100.

Jansons V., Graudiņa A. 2006, Underlying Principles for Establishing Crop Insurances Services. *Int. Finance Symposium, May 25-26, 2006, Turkey*.

Ray P.K., 1991, A Practical Guide to Multi-risk Crop Insurance for Developing Countries. *U.S.A.: Science Publishers, Inc.*, 10-162.

Sklar, A., 1973, Random variables, joint distributions, and copulas, *Kybernetica 9,* 449-460.