UDC 004.942:519.216.3

MODELING RISK FACTORS INTERACTION AND RISK ESTIMATION WITH COPULAS

Kuznietsova N. V. – Dr. Sc., Associate Professor at the department of Mathematical Methods of System Analysis, Institute for Applied Systems Analysis at NTUU "Igor Sikorsky KPI", Kyiv, Ukraine.

Huskova V. H. – PhD, Assistant Professor at the Department of Mathematical Methods for System Analysis, Institute for Applied Systems Analysis at NTUU "Igor Sikorsky KPI", Kyiv, Ukraine.

Bidyuk P. I. – Dr. Sc., Professor at the Department of Mathematical Methods of System Analysis, Institute for Applied Systems Analysis at the NTUU "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine.

Matsuki Y. – Dr. Sc., Department of Natural Sciences at the National University «Petro Mohyla Academy», Kyiv, Ukraine.

Levenchuk L. B. – M. Sc., Postgraduate student at the Institute for Applied System Analysis, National Technical University of Ukraine "Igor Sikorsky KPI", Kyiv, Ukraine.

ABSTRACT

Context. Various risks are inherent to practically all types of human activities. Usually the risks are characterized by availability of multiple risk factors, uncertainties, incompleteness and low quality of data available. The problem of mathematical modeling of risks is very popular with taking into consideration possible uncertainties and interaction of risk factors. Such models are required for solving the problems of loss forecasting and making appropriate managerial decisions.

Objective. The purpose of the study is in development of multivariate risk modeling method using specialized copula functions. The models are developed in the form of multivariate distributions.

Method. The modeling methodology is based upon exploring the special features of various copula functions that are helpful to construct appropriate multivariate distributions for the risk factors selected. The study contains formal description of selected copulas, analysis of their specific features and possibilities for practical applications in the risk management area. Examples of practical applications of the copula based approach to constructing multivariate distributions using generated and actual statistical data are provided.

Results. The results achieved will be useful for further theoretical studies as well as for practical applications in the area of risk management. The distributions constructed with copula create a ground for solving the problems of forecasting possible loss and making appropriate decision regarding risk management.

Conclusions. Thus the problem of constructing multivariate distributions for multiple risk factors can be solved successfully using special copula functions.

KEYWORDS: multivariate stochastic processes, risk estimation, special copula functions, modeling multivariate distributions, combined marginal distributions.

ABBREVIATIONS

EVT is an extreme value theory;

MEVT is an multivariate extreme value theory.

NOMENCLATURE

C is -increasing function;

f_i is a function of density for marginal distributions;

ξ is the parameter that characterizes form of the distribution;

H is n-dimensional joint distribution function with marginal distributions;

β is additional scaling parameter;

N_u is a number of observations that exceed the threshold;

$F_i^{(-1)}(u_i)$ is inverse function to the function of marginal distribution;

ρ is symmetric positively defined matrix with the unity main diagonal;

φ is the function of standard scalar normal distribution;

φ_p is the function of multivariate standard normal distribution with correlation matrix ρ;

C is a number of coordinated couples;

D is a number of non-coordinated couples;

N is a number of observations.

INTRODUCTION

The studies related to risk analysis and management are very popular in the world today practically in every area of human endeavors due to widely spread necessity of various risk estimation, management, and minimization of possible loss. The risk management theory supposes mathematical modeling of risks themselves (including risk factors), and application of the models created to estimation and forecasting possible loss for a time horizon selected. One of the key elements of mathematical modeling is taking into consideration the interactions between risks and risk factors. Such interaction may serve as an amplifier for risk effects and it often results in increasing possible loss. Thus, it is important to take into consideration in the risk management procedures not only isolated risks but also their interaction and integrated estimates in the form of some portfolio risk. That is why the use of the results of former studies in the area, concentrated on the extreme value theory (EVT), in a scalar case imposes substantial restrictions on practical applicability of the

DOI 10.15588/1607-3274-2022-2-5

results because of focusing attention on separate risk factors and their influence [1–6]. In practice of risk management more often the problem arises in correct application of risk estimation procedures that would take into consideration complicated, very often asymmetric, non-stationary and nonlinear interaction between risk itself and risk factors. Correct mathematical and statistical description of the processes involved in the risk analysis procedures is the key point for solving the problems of making effective managerial decisions regarding risk management, for example regarding avoiding risk, its diversification, and minimization.

The natural approach to generalization and improvement of existing risk analysis methodology is the use of multivariate extreme value theory (MEVT) that considers modeling of tails for the multivariate distributions [7–9]. However, the use of extreme values only for corresponding vectors as it is done by scalar EVT, i.e. the vectors containing extreme values in each coordinate, provides the possibility for correct processing relatively small number of measurements. Taking into consideration that of basic interest are only about 3%–5% of observations the simultaneous threshold overcoming for each variable will be taking place very rarely. The more variables will be considered simultaneously the less number of coinciding extreme values will be met. Here the known problem of high dimensionality will not provide the possibility for reaching reliable data processing results when the number of observations is low.

For a long time as widely accepted measure of dependency between two random variables served well-known statistical correlation coefficient. According to the known hypothesis of normality for financial random variables the correlation coefficient was considered as necessary and sufficient measure in the case of multivariate normal distribution. However, in risk management problems very often the data does not correspond to the normality hypothesis, and the criticism arises regarding the correlation coefficient as inadequate measure for analysis of risks dependency [10–13]. The correlation coefficient does not provide appropriate formal description for the dependency structure between the risks, especially in the tails of distributions (Fig. 1). Thus, completely dependent random variables may exhibit the correlation coefficient distinctive from 1 or –1, and zero correlation coefficient does not support the hypothesis of risks independence. For example, this is true in the case of normally distributed risk, X, and completely dependent on its sequence of values X^2. The linear correlation is also not invariant to transformations of random variables.

The purpose of the study is in following:

– to perform analysis of a construction procedure for the class of special copula functions that are suitable for the formal description of multivariate distributions;

– to consider the special features of copula parameter estimation procedures using existing estimation techniques namely maximum likelihood method;

– to estimate the possibilities for practical applications of the copula families for performing statistical analysis of economic, financial, and other risk types represented by the extreme values of respective distributions.

1 PROBLEM STATEMENT

Let we have n independent random variables, $X_1,...,X_n$, for which the following relation is true: $P(X_1 \le x_1;..., X_n \le x_n) = P(X_1 \le x_1) \cdot ... \cdot P(X_n \le x_n)$. t means that the knowledge regarding one of the random variables does not provide the knowledge related to others. Studying the dependency between variables it is of interest to get information related to one random variable using the information about other variable and to compare the mutual dependences between the pairs of random variables. The dependency is inverse characteristic to independence of random variables but the ways of its identification can be different in different cases and is determined by specific problem statement.

This work is focusing on constructing of a class of special functions, copulas that are suitable for the formal description of multivariate distributions, and estimation of the possibilities for practical application of the copula family to statistical analysis of financial and other types of risk data represented by extreme values of corresponding distributions.

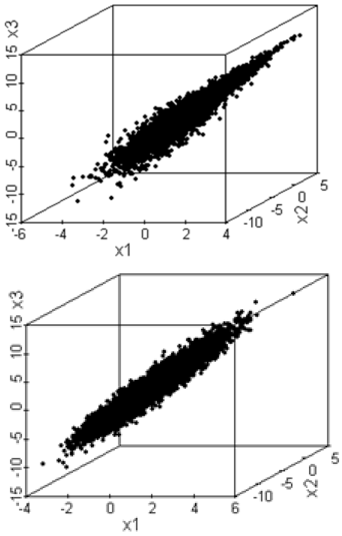

Figure 1 – Two multivariate observations with similar normal marginal distributions and coefficients of about, $\rho = 0.14$, but with different dependency structures

In the risk management procedures, especially regarding financial risks, withdrawing of marginal risk distributions for separate financial instruments from the dependency structure is a natural requirement. On one side each random variable has its scalar distribution but on the other side it is necessary to take into consideration existing dependences between random variables.

Such approach helps to improve the model adequacy and consequently enhance quality of final result – risk estimation. Here the copula notion is useful that allows

DOI 10.15588/1607-3274-2022-2-5

for clearly separation the information related to structure of the dependency and to produce its appropriate formal description. Now consider some theoretical formulations necessary for understanding basic material of the study [14, 15].

2 REVIEW OF THE LITERATURE

Before considering possible ways of describing dependency first return to the general definition of dependent random variables.

Definition 1: Distribution function is such a function F with the domain of definition, $[-\infty,\infty]$, that F is non-decreasing and having marginal the values of: $F(-\infty)=0$ and $F(\infty)=1$.

Definition 2: Distribution function of random variable X is such function, F, that for all, $x\in[-\infty,\infty]$, the following relation is true: $F(x)=P[X\le x]$.

Definition 3: The function H of n arguments is called n-increasing if for any n-dimensional interval, $B=[\vec{a},\vec{b}]$, such that $[\vec{a},\vec{b}]$ belong to the domain of definition of the function H with $\vec{a}\le\vec{b}$, the following is true: $V_H(B)=\Delta_{\vec{a}}^{\vec{b}}H(\vec{x})=\Delta_{a_n}^{b_n}...\Delta_{a_1}^{b_1}H(\vec{x})\ge 0$.

Definition 4: Joint distribution function is such a function H with the domain of definition, $[-\infty,\infty]^n$, that H is n-increasing, and $H(x_1,...,x_{i-1},-\infty,x_{i+1},x_n)=0\forall i$, $H(-\infty,\infty)=1$.

Definition 5: Joint distribution function for the random variables $X_1,...X_n$ is such a joint distribution function H with the domain of definition, $[-\infty,\infty]^n$, that $H(x_1,...,x_n)=P[X_1\le x_1,...,X_n\le x_n]$.

From the given definitions we have that distribution function for random variable X_i is marginal distribution function for joint distribution function of random variables,

$$X_1,...,X_n:$$
$$F(x_i)=P(X_i\le x_i)=\lim_{t_j\to\infty\forall j}H(t_1,...,t_{i-1},x_i)\cdot \quad (1)$$

The complete description of dependence or independence for random variables, $X_1,...X_n$, is $P(X_1\le x_1,...,X_n\le x_n)$, i.e. their joint distribution function, $H(x_1,...,x_n)$. But it also contains excessive information regarding marginal distribution for each random variable. In solving practical problems it is necessary to obtain information regarding the dependency structure separately.

Definition 6: The function, $C:[0,1]^n\to[0;1]$ is called n-copula if the following conditions are true.

$C(F_1,...,F_1)=0$, if there exists such j that $F_j=0$;

$$C(1,...1,F_i,1,...1)=F_i;$$

Theorem 1: Sklar theorem [15]: Let H is n-dimensional joint distribution function with marginal distributions, $F_1,...,F_n$. Then there exists such n-copula that for all, $\vec{x}\in R^n$, the following relation is true:

$$H(x_1,...,x_n)=C(F_1(x_1),...,F_n(x_n)). \quad (2)$$

If $F_1,...,F_n$ are continuous, then C is unique, otherwise the functions C are uniquely determined over the space, $Rng[F_1]\times...\times Rng[F_n]$. And vice versa: if the functions $F_1,...,F_n$ represent continuous distributions, and C is n-copula, then $H(x_1,...,x_n)$ is joint n-dimensional distribution function with marginal distributions, $F_1,...,F_n$.

Definition 7: The density distribution for copula C is the following function:

$$c(F_1,F_2,...,F_n)=\frac{\partial c(F_1,F_2,...,F_n)}{\partial F_1\partial F_2,...,\partial F_n}. \quad (3)$$

The joint density distribution function can be represented in the form:

$$h(x_1,x_2,...,x_n)=c(F_1(x_1),F_2(x_2),...,F_n(x_n)). \quad (4)$$

An important role for risk analysis is the copula feature of invariance to increasing transforms.

Theorem 2 [15]: Consider the transforms $Z_1,...,Z_n$ that are increasing in corresponding domains of definitions of random variables $X_1,...,X_n$ with continuous marginal distribution functions and n-copula, C. Then random variables $Z_1(X_1),...,Z_n(X_n)$ have the same n-copula, C.

3 MATERIALS AND METHODS

Constructing marginal distribution function. When copulas are used for modeling dependences between random variables it is necessary to construct separate model for marginal distributions for the variables. For the problems of risk management of particular importance are estimates of values that belong to the tails of distributions. For the formal description of right tail of loss distributions is recommended the use of the method of overriding based upon the generalized Pareto distribution. And the distribution for the other observations (that are closer to mathematical expectation of a sample) it is recommended to describe with normal distribution using the results of application of the central limit theorem and the experience of using the distribution in the problems of risk management. To separate the two parts of data sample it is necessary to estimate the empirical distribution quintile for a threshold selected. In the computational experiments car-

DOI 10.15588/1607-3274-2022-2-5

ried out by multiple researchers the threshold was selected at the level of 95 % according to the existing practice of risk management.

The sample mathematical expectation of normal distribution is estimated on all observations available. And the standard deviation should be estimated in a way so that the values of normal distribution function at the threshold point were equal to the value of empirical distribution function. Such approach to the computations allows for orienting the model to adequate formal description of the tail observations. Besides, this method of performing the computations provides for the continuous form of the combined distribution function the right part of which is constructed using the method of overriding. The distribution function constructed by the use of this method is equal to the empirical distribution function at the threshold point according to the constructing procedure selected. The continuity of the marginal distribution functions points out to the existence of unique copula in expression (1).

Definition 8: Let random variable, X_i, has distribution function, F_i, and right final point $x_{iF} \le \infty$. For a fixed value, $u < x_{iF}$, the distribution function of exceeding values for u is called the following one:

$$F_{iu}(x) = P(x_i - u_i \le x \mid x_i > u), x + u \le x_{iF} ; \quad (5)$$

and the function:

$$e(u) = E[X - u \mid X > u] = \int_u^{XiF} \frac{1 - F_i(t)}{1 - F_i(x)}, \quad (6)$$

is called the function of mean exceeding.

Denote, $z^+ = \max(z,0)$, and let $card\{A\}$ is a number of elements in the set A. An empirical estimate for the function of mean exceeding is the following one:

$$e(u) = \frac{1}{card\{j : x_{ij} > u, j = 1,...,N\}} \times \\ \times \sum_{j=1}^{N} (x_{ij} - u)^*, u \ge 0. \quad (7)$$

For the distribution function of the values exceeding selected threshold there exist an analogue of the Fisher-Tippet-Gnedenko theorem, i.e. Pikands-Balkema-de Haan theorem.

Theorem 3: Pikands-Balkema-de Haan theorem [3]:

$$\lim_{u \uparrow x_F} \frac{1 - F_i(u + x\beta(u))}{1 - F_i(u)} = \begin{cases} (1 + \xi x)^{-\frac{1}{\xi}}, \xi \ne 0, \\ \exp(-x), \xi = 0, \end{cases} \quad (8)$$

where, $1 + \xi x > 0$, for some positively defined function $\beta(u)$.

From (2) and definition of the distribution function for exceeding values we have a model for distributions of threshold exceeding values: the function of generalized Pareto distribution:

$$GPD_{\xi,\beta}(x) = \begin{cases} 1 - (1 + \xi x)^{-\frac{1}{\xi}}, \xi \ne 0, \\ 1 - \exp(-x), \xi = 0, \end{cases} \quad (9)$$

where, $\beta > 0$, and $x \ge 0$ with $\xi > 0$, and $0 \le x \le \frac{-\beta}{\xi}$ with $\xi < 0$.

The model for tail data distribution is constructed by the method of threshold exceeding that is based upon the marginal distribution law for the exceeding values (3) and includes the following steps.

1. For sample $\{x\}_i$ of power, N, the threshold u is selected. Then the observations $x_{i_1},...,x_{i_N}$ are determined exceeding the threshold, and respective exceeding values are computed: $y_j = x_{i_j} - u \ge 0$, where N_u is a number of observations that exceed the threshold.

2. Then the function $F_u(y)$ (distribution of exceeding values, $y_1,...,y_{N_u}$ in the form of (3), $GPD_{\xi,\beta}(x)$, is estimated, i.e. the parameters of form and scale are computed.

3. The distribution function for the tail region, X is estimated as follows:

$$F_{iu}(y) = \frac{F_i(y+u) - F_i(u)}{1 - F_i(u)}, \\ F_i(x) = (1 - F_i(u))F_{ij}(y) + F_i(u), \\ F_i(u) = \frac{N_u}{N}. \quad (10)$$

i.e. we have:

$$F_i(x) = (1 - \frac{N_u}{N})F_{iu}(y) + \frac{N_u}{N}. \quad (11)$$

Copula family constructing. The Sklar theorem guaranties copula existence and its uniqueness for definite conditions but it does not provide the method for its constructing. Consider some methods for copula constructing.

1. The method of inverse function

The idea of the method is in the following: from the Sklar theorem for joint distribution function, H, and continuous marginal distributions, $F_1,...,F_n$, copula, C, is defined as follows:

$$C(u_1,...,u_n) = H(F_1^{(-1)}(u_1),...,F_n^{(-1)}(u_n)). \quad (12)$$

The most widely spread in modeling random variables are elliptical distributions and more exactly multivariate

DOI 10.15588/1607-3274-2022-2-5

normal distribution. If we apply to the distribution inverse method then we will get multivariate normal copula or Gaussian copula.

Definition 9: Let ρ is symmetric positively defined matrix with the unity main diagonal. Then the following function is called Gaussian multivariate copula:

$$C(F_1, F_2 ..., F_n, \rho) = \varphi_\rho(\varphi_{-1}(F_1),...,\varphi_{-1}(F_n)) . \quad (13)$$

Fig. 2 (below) shows tree-dimensional distribution on the basis of normal copula.

The density of multivariate Gaussian copula is defined as follows [14]:

$$C(F_1, F_2 ..., F_n, \rho) = \frac{1}{\sqrt{\rho}} \exp(-\frac{1}{2}(\varphi^{-1}(F_1), \\ \varphi^{-1}(F_n)) \times (\rho^{-1} - 1) \times \begin{pmatrix} \varphi^{-1}(F_1) \\ \varphi^{-1}(F_n) \end{pmatrix}). \quad (14)$$

Thus, the Gaussian normal copula is completely defined by the correlation matrix, ρ, and its parameters can be easily computed.

According to its constructing procedure the Gaussian copula can be naturally used for modeling multivariate normal distributions, and it can be hired in the risk management procedures for constructing meta-normal distributions. This multivariate distribution is created in the way of modeling the dependences between random variables using normal copulas, and marginal distributions are built with some other distributions appropriate for each variable under consideration.

Another type of copula that is constructed with the method of inverse function from elliptical distribution is Student's copula that corresponds to the multivariate Student's t-distribution. The form of this copula looks to some extent like the normal one in its central part, and approaches even more to this form in the tail part where the number of degrees of freedom for Student's t-distribution is growing. Fig. 3 shows three-dimensional distribution of this type for 5 degrees of freedom. However, in the risk management problems the Student's copula can find its extended application due to substantial difference in modeling of dependences in the tail of distributions far from the central part.

The t-copula that is derived from t-distribution with ν degrees of freedom and positively defined matrix $\sum$, has the following density function:

$$c(x)\Gamma((v+d)/2)/\Gamma(v/2(\pi v)^2 |\sum \ |)1) . \quad (15)$$

As it can be seen, the elliptical copulas also have the advantage that they can be easily computed for large, n.

2. The Archimedean copulas

Definition 10: Let φ is continuous strictly increasing function defined on region 1 of $[-\infty;0]$ such that, $\varphi(1) = 0$. The pseudo-inverse for φ is such function $\varphi^{[-1]}$ with $Dom\varphi = [0,\infty]$:

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t), 0 \le t \le \phi(0), \\ 0, \varphi(0) \le t \le \infty. \end{cases} \quad (16)$$

Note, that $\varphi^{[-1]}$ is continuous and non-decreasing function over the domain of, $[0;\infty]$, and strictly decreasing over $[0;\varphi(0)]$. Moreover, $\varphi^{[-1]}(\varphi(u)) = u$ over, 1, and

$$\varphi(\varphi^{[-1]}(t)) = \begin{cases} t, 0 \le t \le \varphi(0) \\ \varphi(0), \varphi(0) \le t \le \infty \end{cases} . \quad (17)$$

If $\varphi(0) = \infty$, then φ^{-1}.

Lemma 1: Let φ is continuous strictly decreasing function in the domain 1 over the interval, $[0;\infty]$, and such that, $\varphi(1) = 0$, and let $\varphi^{[-1]}$ is pseudo-inverse for, φ, and function, C, with, I^2, defined over I for the function, φ, is defined as follows:

$$C(u, v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)) . \quad (18)$$

Then, C, satisfies the existing restriction conditions for a copula.

Proof: Here the function, $C(u,0) = \varphi^{[-1]}(\varphi(u) + \varphi(0)) = 0$, and, $C(u,1) = \varphi^{[-1]}(\varphi(u) + \varphi(1)) = \varphi^{[-1]}(\varphi(u)) = u$.

If $\varphi(0) = \infty$, then φ, is considered as a strict generating function. In this case, $\varphi^{[-1]} = \varphi^{-1}$, and

$$C(u, v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)) , \quad (19)$$

is considered as a strict Archimedean copula.

All the copulas that can be represented in the form (18) are called Archimedean. In this case

The function φ is called copula generator. This copula class is one of the most often used in practice due to the fact that it includes a substantial number of parametric copulas reflecting a variety of structural mutual dependences. Besides, the constructing procedure for the copulas is relatively simple.

For example, consider, $\varphi(t) = (-\ln t)^\theta$, where $\theta \ge 1$. It is clear that $\varphi(t)$ is continuous with $\varphi(1) = 0$; and $\varphi(t) = -\theta \frac{1}{t(-\ln t)^{\theta-1}}$; that is, φ is strictly decreasing

DOI 10.15588/1607-3274-2022-2-5

function from I defined over domain $[0;\infty]$. Further on, $\varphi(t)^t = \frac{\theta(\theta-1)1}{t^2(-\ln t)^{\theta-2}} + \theta\frac{1}{t^2(-\ln t)^{\theta-1}} \geq 0$ from, I, and thus φ is convex. Moreover, $\varphi(0)=\infty$, and thus φ can be considered as a strict generating function. Now, from (6) we have:

$$\exp = ((-\ln u) + (-\ln v)) . \tag{20}$$

This copula family is called Gumbel copula in the case of describing two variables. The Archimedean copulas also include Frank copula family in the case of describing two variables:

$$C(F_1 1, F_1 2) = -1/\beta \times \\ \times \frac{\ln(1+((e(-\beta F_1 1)-1))(e(-\beta F_1 2)-1)))}{(e(-\beta)-1)}. \tag{21}$$

and Clayton copulas also in the case of two variables:

$$C(F_1 1, F_1 2) = \max((F1(-\beta) + F2(-\beta) - 1)\times \\ \times(-1/\beta), 0)\beta \in [-1,\infty]\{0\}). \tag{22}$$

The Fig. 4–6 show three-dimensional joint distributions on the bases of Frank, Gumbel, and Clayton copulas respectively.

One of natural approaches to constructing multidimensional copulas supposes constructing first the families of two-dimensional copulas exhibiting necessary features. After this step the next procedure touches upon constructing multidimensional generalization. The family of multidimensional copulas belongs to generalization of two-dimensional family if the following conditions are true:

– all two-dimensional components of the multidimensional copula belong to the family of the two-dimensional ones;

– all the two-dimensional components of order 3 and up to $n-1$ have the same multidimensional form.

The problem of constructing the multidimensional generalizations of two-dimensional copulas is not trivial. Very often two-dimensional link functions do not have multidimensional generalizations or they result in such specific dependency structure that has very restricted application.

The function C^n is a result of sequential application of Archimedean two-dimensional copulas based on the ϕ generator. Thus, $C^2(u_1,u_2) = C(u_1,u_2) = \varphi^{[-1]}(\varphi(u_1)+\varphi(u_2))$, and for $n \geq 3$, $C^n(u_1,u_2,...,u_n) = C(C^{n-1}(u_1,u_2),u_n)$. Let's stress that such approach to constructing copulas usually does result in a success. But Archimedean copulas are symmetric and associative what often results in copula C^n with, $n > 3$.

Parameter estimation procedures for copulas. The simplest copula parameter estimation method is in application of known data sample characteristics. Using the correlation matrix estimates it is possible to find an estimate for the form parameters of elliptical copula, ρ. Another useful in this case is such concordance measure as Kendall statistic, τ, that is computed as follows:

$$\bar{\tau} = 2\frac{C-D}{N(N-1)} . \tag{23}$$

The Kendal τ can be used, for example, for estimating Clayton copula parameters, $\bar{\theta} = \frac{2\bar{\tau}}{1-\bar{\tau}}$, and elliptical copula parameters, $\rho = \sin\frac{\pi}{2}\bar{\tau}$. This method is rather simple but it restricts substantially possibilities for selection of copula families.

Estimation of copula parameters for any family can be performed with maximum likelihood method. Consider joint distribution function, $H(\vec{x},\vec{\theta})$, where parameter vector includes marginal distribution and copula parameters $\vec{\theta} = (\beta_1,...,\beta_n,\vec{\alpha})$. The joint distribution function can be represented in the following form:

$$H(\vec{x},\vec{\theta}) = C(F_1(x_1,\beta_1), F_2(x_2,\beta_2),..., \\ F_n(x_n,\beta_n)). \tag{24}$$

For a sample $S_x = \{\vec{x}\}$ of power N the likelihood function is computed as follows:

$$\ln L(S_x,\beta_1,...,\beta_n,\vec{\alpha}) = \sum_{i=1}^{N} C(F_1(x_1(i),\beta_1), \\ F_2(x_2(i),\beta_2),...,F_n(x_n(i),\beta_n),\vec{\alpha}) + \\ + \sum_{i=1}^{N}\ln f_1(x_1(i),\beta_1) + ... + \sum_{i=1}^{N}\ln f_n(x_n(i),\beta_n). \tag{25}$$

This expression allows for application of two-stage algorithm for parameters estimation:

$$(\beta_1,...,\beta_n,\vec{\alpha}) = \underset{\beta_1,...,\beta_n,\vec{\alpha}}{\arg\max} \\ \ln L(S_x,\beta_1,...,\beta_n,\vec{\alpha}). \tag{26}$$

At the first step we can compute the parameter estimates for the marginal distributions as follows:

$$\beta_j = \underset{\beta_j}{\arg\max}\sum_{i=1}^{N}\ln f_i(x_j(i),\beta_j) , \tag{27}$$

and, at the second step it will be possible to compute the estimates of the copula parameters:

$$\vec{\alpha} = \underset{\beta_j}{\arg\max}\sum_{i=1}^{N}\ln f_i(x_j(i),\beta_j) . \tag{28}$$

DOI 10.15588/1607-3274-2022-2-5

It is shown in [16] that the vector parameter estimate $\vec{\theta} = (\beta_1, ..., \beta_n, \vec{\alpha})$ is consistent and asymptotically normal. An advantage of the two-step estimation procedure is in dimension reduction what makes the numerical optimization procedure easier, and allows for hiring extra information necessary for parameter estimation of separate risk distributions. It is also helpful for more complete usage of measuring information when data samples available have different power when new financial instruments or risk factors come to being.

4 EXPERIMENTS

Here we consider the problem of constructing joint distribution of risk factors for generated three-dimensional distributions for Cauchy, t-Student, normal data as well as three-dimensional distribution for the three currency exchange rates of EUR, CHF, GBP against USD. For each dataset the Archimedean copulas were estimated from the Gumbel, Clayton and Frank families as well as the elliptical copulas of the Student family and Normal copula [18].

5 RESULTS

Together with estimates of marginal distributions the computational experiment provided the possibility for modeling the joint distribution function for the processes under consideration. The Fig. 2–6 show graphical illustrations of the joint distributions for the selected currencies exchange rates using different dependency structures.

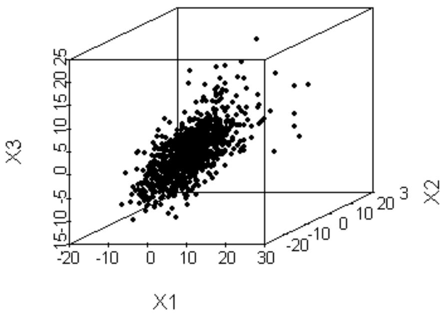

Figure 2 – Graphical representation of joint distribution function on the basis of normal copula

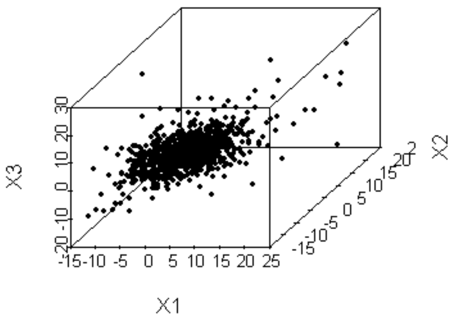

Figure 3 – Graphical representation of joint distribution function on the basis of Student t-copula

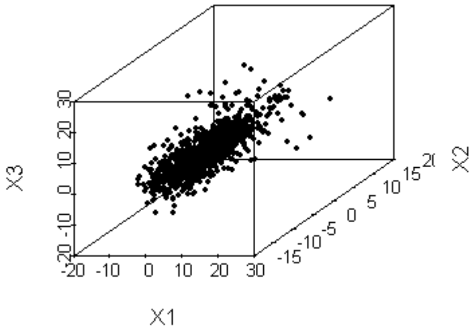

Figure 4 – Graphical representation of joint distribution function on the basis of Frank copula

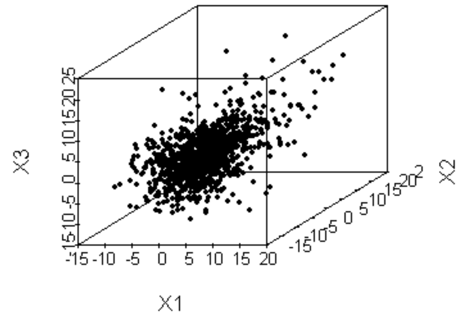

Figure 5 – Graphical representation of joint distribution function on the basis of Gumbel copula

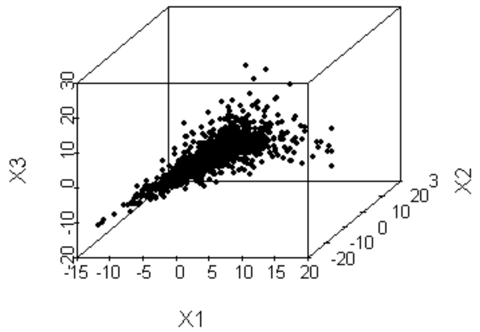

Figure 6 – Graphical representation of joint distribution function on the basis of Clayton copula

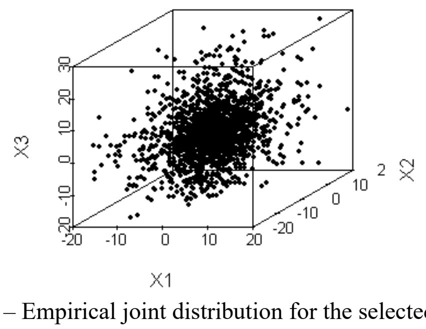

Figure 7 – Empirical joint distribution for the selected currency exchange rates

The results of the computations performed illustrate practical usefulness of the study performed.

6 DISCUSSION

The main goal of the analysis was in constructing the joint distributions for risk factors and further use of the

DOI 10.15588/1607-3274-2022-2-5

models for solving the risk management problems, for example to forecast possible loss in investment problems.

The results of the study show that together with estimates of marginal distributions the computational experiment provided the possibility for modeling the joint distribution function for the processes under consideration. The Fig. 2–6 show graphical illustrations of the joint distributions. The results will also be helpful for constructing decision support systems for financial risk forecasting and management using appropriate statistical data.

The copula constructed according to the method of inverse function from elliptical distribution is the Student copula that corresponds to the multidimensional Student-distribution. The form of the copula in its central part resembles very much a normal one and becomes even more alike in the tail part with growing the number of degrees of freedom for the Student-distribution.

The proposed approaches will be used by the authors of the article in further studies, as well as by other scientists, which will allow obtaining more accurate results using minimal costs for experiments.

CONCLUSIONS

The analysis has been performed regarding the possibility of the special class of copula functions application to formal description of multidimensional distributions in the problems of financial risk modeling.

The scientific novelty of obtained results is that the method of constructing combined marginal distributions was proposed that allows for taking into consideration heavy tails of one-dimensional risk distributions. The marginal distributions are combined into joint distributions of risks via the dependency structure that is characterized by copulas. On the basis of analysis the methods of constructing the copula families it was proposed to use for modeling risk several families of copulas with useful for risk management features.

The practical significance of obtained results is that the computational experiment has been carried out with two generated, theoretically known tree-dimensional distributions, and one empirical three-dimensional distribution created for a formal description of selected currencies exchange rate. The experiment demonstrated the possibility of the method proposed for modeling multidimensional distribution using appropriately combined marginal distributions and the dependency structure between them.

Prospects for further research is that In the future studies it is planned to expand application of the copula families and develop the methods for estimating basic risk measures for a portfolio of financial instruments on the basis of the multidimensional distribution model proposed. It is also planned to apply the method proposed for developing the systemic methodology for risks management including market and credit risks. The problems mentioned will be solved in the frames of a specialized intellectual decision support system designed on the purpose of automatizing risk modeling and management procedures and generating appropriate decision alternatives.

ACKNOWLEDGMENTS

The work was performed at the Institute for Applied System Analysis of the National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute” and Department of Natural Sciences of the National University “Petro Mohyla Academy” within the framework of scientific research conducted by the departments.

REFERENCES

1. Nystrom K., Skoglund J. Univariate Extreme Value Theory, GARCH and Measures of Risk, *Stockholm: Swedbank, Group of Financial Risk Control,* 2002, Vol. 11, pp. 1–27.
2. Bhatti M. I., Hung Quang Do Recent Developments in Copula and Its Applications to the energy, Forestry and Environmental Sciences, *International Journal of Hydrogen Energy*, 2019. Vol. 44, Number 36, pp. 19453–19473.
3. Kluppelberg C. Risk Management with Extreme Value Theory. New York, Chapman & Hall, 2003, 405 p.
4. Bassi F., Embrechts P., Kafetzaki M. Risk management and quantile estimation, *A Practical Guide to Heavy Tails; Ed. R.J. Adler*. Boston, Birkhauser, 1998, Vol. 4, pp. 111–130.
5. Embrechts P. Eds.: B. Finkenstadt, H. Rootzen. Extremes in Economics and the Economics of Extremes, 2001, Vol. 12. pp. 169–183.
6. Embrechts P., Haan L., Huang X. Modelling multivariate extremes, *RISK Books,* 2000, Vol. 12, pp. 59–67.
7. Bouye E. Multivariate extremes at work for portfolio risk measurement, *Finance: Presses universitaires de France,* 2002, Vol. 23, Nubmer 2, pp. 125–144.
8. Rootzen H., Tajvidi N. The multivariate generalized Pareto distribution, *Bernoulli,* 2006, Vol. 12, pp. 917–930.
9. Bidyuk P. I., Kroptya A. V. Analysis and methods of solving the problem of extreme values estimation, *Scientific Bulletin of NTUU KPI*, 2005, No. 4, pp. 34–47.
10. Blum P., Dias A., Embrechts P. The ART of dependence modelling: the latest advances in correlation analysis, *Alternative Risk Strategies: RISK Books*, 2002, Vol. 7, pp. 339–356.
11. Embrechts P., McNeil A. J., Straumann D. Correlation: Pitfalls and alternatives, *Risk*, 1999, Vol. 12, pp. 69–71.
12. Embrechts P., McNeil A. J., Straumann D. Correlation and dependence in risk management. Cambridge, Cambridge University Press, 2002, pp. 176–223.
13. Edward W., Valdez E. Understanding relationships using copulas, *North American Actuarial Journal,* 1998, Vol. 2. Number 1, pp. 1–25.
14. Nelsen R. B. An Introduction to Copulas. New York, Springer, 2006, 270 p.
15. Newey W. K., McFadden D. Large Sample Estimation and Hypothesis Testing, *Handbook of Econometrics*, 1994, Vol. 4, pp. 2111–2245.
16. Bashkov E., Dmitrieva O., Huskova N. et al. Parallel multiple blocked methods of Bickart type, *Proceedings of SPIE – The International Society for Optical Engineering*, 2019, No. 11176, P. 9.
17. Miroshkin O., Kovalov S., Dmitrieva O. et al. ParSimTech Research and Training Center, *IEEE European Technology and Engineering Management Summit (E-TEMS)*, 2020, pp. 1–4.
18. Siddhartha M. Gold Price Prediction Dataset [E’lektronnyj resurs]. Rezhim dostupa, https://www.kaggle.com/sid321axn/gold-price-prediction-dataset.

Received 29.11.2021.
Accepted 25.01.2022.

DOI 10.15588/1607-3274-2022-2-5

УДК 004.942:519.216.3

МОДЕЛЮВАННЯ ВЗАЄМОДІЇ ФАКТОРІВ РИЗИКУ І ОЦІНЮВАННЯ РИЗИКУ З ВИКОРИСТАННЯМ КОПУЛ

Кузнецова Н. В. – д-р техн. наук, доцент кафедри математичних методів системного аналізу, Інститут прикладного системного аналізу, НТУУ «Київський політехнічний інститут ім. Ігоря Сікорського », Київ, Україна.

Гуськова В. Г. – д-р філософії, старший викладач кафедри математичних методів системного аналізу, Інститут прикладного системного аналізу, НТУУ «Київський політехнічний інститут ім. Ігоря Сікорського », Київ, Україна.

Бідюк П. І. – д-р техн. наук, професор кафедри математичних методів системного аналізу, Інститут прикладного системного аналізу, НТУУ «Київський політехнічний інститут ім. Ігоря Сікорського », Київ, Україна.

Мацукі Й. – д-р техн. наук, факультет природничих наук, національний університет «Києво-Могилянська Академія», Київ, Україна.

Левенчук Л. Б. – магістр, аспірантка Інституту прикладногосистемного аналізу НТУУ «КПІ ім. І. Сікорського», Київ, Україна.

АНОТАЦІЯ

Актуальність. Різні типи ризиків притаманні практично всім видам людської діяльності. Зазвичай ризики характеризуються наявністю множини факторів ризику, невизначеностями, неповнотою і низькою якістю наявних даних. Задача математичного моделювання ризиків є досить популярною, беручи до уваги можливі невизначеності і взаємодію факторів ризику. Такі моделі необхідні для розв'язання задач прогнозування втрат і прийняття належних управлінських рішень.

Мета роботи. Метою цього дослідження є розробка методу моделювання багатовимірного ризику з використанням спеціальних функцій копул. Моделі пропонуються у формі багатовимірних розподілів.

Метод. Технологія моделювання ґрунтується на використанні спеціальних властивостей копул, які дають можливість побудувати коректні багатовимірні розподіли для вибраних факторів ризику. У статті подано формальний опис вибраних копул, аналіз їх властивостей і можливостей практичного застосування у системах менеджменту ризиків. Подані приклади практичного застосування копул до побудови багатовимірних розподілів з використанням згенерованих і фактичних статистичних даних.

Результати. Отримані результати будуть корисними для подальших теоретичних досліджень, а також для практичного використання у системах менеджменту ризиків. Розподіли, побудовані за допомогою копул, створюють основу для розв'язання задач прогнозування можливих втрат і прийняття належних рішень стосовно менеджменту ризиків.

Висновки. Таким чином, задача побудови багатовимірних розподілів для множини факторів ризику може бути успішно розв'язана завдяки використанню спеціальних функцій копул.

КЛЮЧОВІ СЛОВА: багатовимірні стохастичні процеси, оцінювання ризику, спеціальні функції копули, моделювання багатовимірних розподілів, комбіновані маргінальні розподіли.

УДК 004.942:519.216.3

МОДЕЛИРОВАНИЕ ВЗАИМОДЕЙСТВИЯ ФАКТОРОВ РИСКА И ОЦЕНКА РИСКА С ИСПОЛЬЗОВАНИЕМ КОПУЛ

Кузнецова Н. В. – д-р техн. наук, доцент кафедры математических методов системного анализа, Институт прикладного системного анализа НТУУ «Киевский политехнический институт им. Игоря Сикорского», Киев, Украина.

Гуськова В. Г. – доктор философии, старший преподаватель кафедры математических методов системного анализа, Институт прикладного системного анализа, НТУУ «Киевский политехнический институт им. Игоря Сикорского», Киев, Украина.

Бидюк П. И. – д-р техн. наук, профессор кафедры математических методов системного анализа, Институт прикладного системного анализа, НТУУ «Киевский политехнический институт им. Игоря Сикорского», Киев, Украина.

Мацуки Й. – д-р техн. наук, факультет естественных наук, национальный университет «Киево-Могилянская Академия», Киев, Украина.

Левенчук Л. Б. – магистр, аспирантка Института прикладного системного анализа НТУУ «КПИ им. И. Сикорского», Киев, Украина.

АННОТАЦИЯ

Актуальность. Различные типы рисков присущи практически всем видам человеческой деятельности. Обычно риски характеризуются наличием множества факторов риска, неопределенностями, неполнотой и низким качеством имеющихся данных. Задача математического моделирования рисков достаточно популярна, учитывая возможные неопределенности и взаимодействие факторов риска. Такие модели необходимы для решения задач прогнозирования потерь и принятия надлежащих управленческих решений.

Цель работы. Целью данного исследования является разработка метода моделирования многомерного риска с использованием специальных функций копул. Модели предлагаются в форме многомерных распределений.

Метод. Технология моделирования основывается на использовании специальных свойств копул, позволяющих построить корректные многомерные распределения для выбранных факторов риска. В статье представлено формальное описание избранных копул, анализ их свойств и возможностей практического применения в системах менеджмента рисков. Представлены примеры практического применения копул к построению многомерных распределений с использованием сгенерированных и фактических статистических данных.

DOI 10.15588/1607-3274-2022-2-5

Результаты. Полученные результаты будут полезны для дальнейших теоретических исследований, а также для практического использования в системах менеджмента рисков. Распределения, построенные с помощью копул, создают основу для решения задач прогнозирования возможных потерь и принятия надлежащих решений по менеджменту рисков.

Выводы. Таким образом, задача построения многомерных распределений для множества факторов риска может быть успешно решена благодаря использованию специальных функций копул.

КЛЮЧЕВЫЕ СЛОВА: многомерные стохастические процессы; оценка риска; специальные функции капусты; моделирование многомерных распределений; комбинированные маргинальные распределения.

ЛІТЕРАТУРА / ЛИТЕРАТУРА

1. Nystrom K. Univariate Extreme Value Theory, GARCH and Measures of Risk / K. Nystrom, J. Skoglund // Stockholm: Swedbank, Group of Financial Risk Control. – 2002. – Vol. 11. – P. 1–27.
2. Bhatti M. I. Recent Developments in Copula and Its Applications to the energy, Forestry and Environmental Sciences / M. I. Bhatti, Hung Quang Do // International Journal of Hydrogen Energy. – 2019. – Vol. 44, Number 36. – P. 19453–19473.
3. Kluppelberg C. Risk Management with Extreme Value Theory / C. Kluppelberg. – New York : Chapman & Hall, 2003. – 405 p.
4. Bassi F. Risk management and quantile estimation / F. Bassi, P. Embrechts, M. Kafetzaki // A Practical Guide to Heavy Tails; Ed. R. J. Adler. – Boston : Birkhauser. – 1998. Vol. 4. – P. 111–130.
5. Embrechts P. Extremes in Economics and the Economics of Extremes / P. Embrechts // Eds.: B. Finkenstadt, H. Rootzen. – 2001. – Vol. 12. – P. 169–183.
6. Embrechts P. Modelling multivariate extremes / P. Embrechts, L. Haan, X. Huang // RISK Books. – 2000. – Vol. 12. – P. 59–67.
7. Bouye E. Multivariate extremes at work for portfolio risk measurement / E. Bouye // Finance: Presses universitaires de France. – 2002. – Vol. 23, Number 2. – P. 125–144.
8. Rootzen H. The multivariate generalized Pareto distribution / H. Rootzen, N. Tajvidi // Bernoulli. – 2006. – Vol. 12. – P. 917–930.
9. Bidyuk P. I. Analysis and methods of solving the problem of extreme values estimation / P. I. Bidyuk, A. V. Kroptya // Scientific Bulletin of NTUU KPI. – 2005. – No. 4. – P. 34–47.
10. Blum P. The ART of dependence modelling: the latest advances in correlation analysis/ P. Blum, A. Dias, P. Embrechts // Alternative Risk Strategies: RISK Books. – 2002. – Vol. 7. – P. 339–356.
11. Embrechts P. Correlation: Pitfalls and alternatives / P. Embrechts, A. J. McNeil, D. Straumann // Risk. – 1999. – Vol. 12. – P. 69–71.
12. Embrechts P. Correlation and dependence in risk management / P. Embrechts, A. J. McNeil, D. Straumann // Cambridge: Cambridge University Press. – 2002. – P. 176–223.
13. Edward W. Understanding relationships using copulas / W. Edward, E. Valdez // North American Actuarial Journal. – 1998. – Vol. 2. Number 1. – P. 1– 25.
14. Nelsen R. B. An Introduction to Copulas / R. B. Nelsen. – New York : Springer, 2006. – 270 p.
15. Newey W.K. Large Sample Estimation and Hypothesis Testing / W. K. Newey, D. McFadden //Handbook of Econometrics. – 1994. Vol. 4. – P. 2111–2245.
16. Parallel multiple blocked methods of Bickart type / [E. Bashkov, O. Dmitrieva , N. Huskova et al.] // Proceedings of SPIE – The International Society for Optical Engineering. – 2019. – No. 11176. – P. 9.
17. ParSimTech Research and Training Center / [O. Miroshkin, S. Kovalov, O. Dmitrieva et al.] // IEEE European Technology and Engineering Management Summit (E-TEMS). – 2020. – P. 1–4.
18. Siddhartha M. Gold Price Prediction Dataset [Электронный ресурс] / M. Siddhartha. – Режим доступа: https://www.kaggle.com/sid321axn/gold-price-prediction-dataset.

DOI 10.15588/1607-3274-2022-2-5