ACTA UNIVERSITATIS APULENSIS No 11/2006

Proceedings of the *International Conference on Theory and Application of Mathematics and Informatics ICTAMI 2005* - Alba Iulia, Romania

THE USE OF COPULAS IN THE STUDY OF CERTAIN TRANSFORMS OF RANDOM VARIABLES WITH APPLICATIONS IN FINANCE

Edith Kovacs

Abstract: In this paper we estimate the probability $P(Z \in (A, B))$, $A, B \in R$ for a transform $Z = \varphi(\mathbf{X})$ of the n-dimensional random variable $\mathbf{X} = (X_1, \cdots, X_n)$, without the knowledge of the joint distribution function of $\mathbf{X}$. For this purpose, we use a Copula, estimated from a sample of the n-dimensional random variable $\mathbf{X}$.

Keywords: Joint distribution function, Marginals, Copula, Value at risk.

2000 *Mathematics Subject Classification*: 62H15

1. First section

The financial life often requires the calculation of some sort of expectations: expected utility, expected payoffs, expected losses, which are expressed as certain transforms of some random variables. We shall provide a method, based on the copula concept, to calculate the probabilities that these transforms belong to a certain neighbourhood of their expected values.

First, we recall briefly the concept of copula and some useful results.

We denote $I = [0, 1]$.

Definition 1.1. *The bivariate copula is a function $C : I^2 \to I$, with the following properties:*

a) $\forall\ u, v \in I, \quad C(u, 0) = C(0, v) = 0$

b) $\forall\ u, v \in I, \quad C(u, 1) = u, \quad C(1, v) = v$

c) $\forall\ u_1,\ u_2,\ v_1,\ v_2 \in I, \quad u_1 \leq u_2, \quad v_1 \leq v_2$ *the inequality*

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0.$$

Let X and Y be two continuous random variable, with the joint distribution function $H(x, y) : \mathbf{R}^2 \to [0, 1]$ and marginals $F(x) : \mathbf{R} \to [0, 1]$ respectively $G(y) : \mathbf{R} \to [0, 1]$.

Definition 1.2. *The function* $C : I^2 \to I$,

$$\forall u, v \in I \; C(u, v) = H\left(F^{-1}(u), G^{-1}(v)\right) \tag{1}$$

is called the copula associated to the pair of continuous random variables X *and* Y.

The above defined function C satisfies the requirements a), b), c) of the Definition 1.1.

Theorem 1.1.(Sklar [1]). *If* X *and* Y *are two continuous random variables with joint distribution* $H(x, y)$ *and marginals* $F(x)$ *and* $G(y)$, *then a unique bivariate copula* $C : I^2 \to I$ *exists, so that the relation (1) holds. Conversely, if* $C : I^2 \to I$ *is a bivariate copula, then only one joint distribution* $H(x, y)$, *with marginals* $F(x)$ *and* $G(y)$, *exists, so that*

$$\forall\, x, y \in \mathbf{R} \quad H(x, y) = C\left(F(x), G(y)\right). \tag{2}$$

Definition 1.3. *The n-variate copula is a function* $C : I^n \to I$, *with the following properties:*

a) If $i \in \overline{1, n}$ *exists so that* $u_i = 0$, *then* $C\left(u_1, \cdots, u_{i-1}, 0, u_{i+1}, \cdots, u_n\right) = 0$;

b) $C\left(1, \cdots, 1, u_i, 1, \cdots, 1\right) = u_i, \quad i = \overline{1, n}$.

c) If $[a_i, b_i] \subseteq I, \quad i \in \overline{1, n}$ *and* $B = [a_1, b_1] \times \cdots \times [a_n, b_n]$, *an n-box, then*

$$V_C(B) = \Delta_{a_n}^{b_n} \; \Delta_{a_{n-1}}^{b_{n-1}} \; \cdots \Delta_{a_2}^{b_2} \; \Delta_{a_1}^{b_1} \; C(t) \geq 0, where$$

$$\Delta_{a_k}^{b_k} \; C(t) = C\left(t_1, \cdots, t_{k-1}, b_k, t_{k+1}, \cdots, t_n\right).$$

Theorem 1.2.(Sklar [1]). *If* X_i *are continuous random variables with joint distribution* $H(\mathbf{x})$, $\mathbf{x} \in \mathbf{R}^n$, $\mathbf{x} = (x_1, \cdots, x_n)$ *and marginals* $F_i(x_i)$,*then a unique n-variate copula* $C : I^n \to I$ *exists, so that*

$$\forall \mathbf{u} \in I^n \quad \mathbf{u} = (u_1, \cdots, u_n) \quad C\left(\mathbf{u}\right) = H\left(F_1^{-1}(u_1), \cdots, F_n^{-1}(u_n)\right). \tag{3}$$

Conversely, if $C : I^n \to I$ *is a n-variate copula, then only one joint distribution* $H(\mathbf{x})$, $\mathbf{x} \in \mathbf{R}^n$, *with continuous marginals, exists, so that*

$$\forall \mathbf{x} \in \mathbf{R}^n \quad H(\mathbf{x}) = C\left(F_1(x), \cdots, F_n(x_n)\right). \tag{4}$$

2. The sceond section

In this section, we consider a bidimensional random variable (X, Y), having the joint distribution function $H : \mathbf{R}^2 \to [0,1], \quad H(x,y) = P(X < x,\ Y < y)$ and the marginal distribution functions $F : \mathbf{R} \to [0,1],\ F(x) = P(X < x)$ and $G : \mathbf{R} \to [0,1],\ G(y) = P(Y < y)$. Generally, when the independence of X and Y is not guaranteed, only the knowledge of marginal distribution functions F and G is not sufficient for the determination of the joint distribution function H, nor of the distribution function of a random variable $Z = \varphi(X, Y)$, which is the transform of (X, Y) by the function $\varphi : \mathbf{R}^2 \to \mathbf{R}$.

Many practical problems need to estimate the probability $P\left(Z \in (A, B)\right)$, $A, B \in \mathbf{R}, A < B$, when we know the marginal distribution functions F and G and a sample $\left\{(x_i, y_i),\ i \in \overline{1,m}\right\}$ of the bidimensional random variable (X, Y).

In the following, we suppose:

a)F and G are *strictly* monotonic and admit continuous derivative of first order;

b) the copula $C : [0,1]^2 \to [0,1]$, which corresponds to H, admits continuous second partial derivatives.

We denote

$D = \left\{(x,y) \in \mathbf{R}^2,\ \ \varphi(x,y) \in (A,B)\right\},$
$D^* = \left\{(u,v) \in [0,1]^2,\ \ \varphi\left(F^{-1}(u), G^{-1}(v)\right) \in (A,B)\right\}$

Theorem 2.1.*If the functions $x \mapsto F(x)$, $\quad y \mapsto G(y)$, $(u,v) \mapsto C(u,v)$ satisfies the conditions a) and b), then the following equality holds*

$$P\left(Z \in (A,B)\right) = \int\int_{D^*} \frac{\partial^2 C(u,v)}{\partial u \partial v}\, du\ dv. \tag{5}$$

Proof. Since the conditions for the validity of (2) fulfills, from hypothesis it results that H also admits the partial derivatives

$$\frac{\partial H(x,y)}{\partial x}, \quad \frac{\partial^2 H(x,y)}{\partial x \partial y}.$$

In the relation (2) we perform successively a partial derivation with respect to x, then with respect to y and obtain

$$\frac{\partial H(x,y)}{\partial x} = \frac{\partial C\left(F(x), G(y)\right)}{\partial u}\ F'(x)$$

$$\frac{\partial^2 H(x,y)}{\partial x \partial y} = \frac{\partial^2 C\left(F(x), G(y)\right)}{\partial u \partial v}\; F'(x)\; G'(y). \tag{6}$$

Using (6), the following equalities take place

$$P\left(Z \in (A,B)\right) = P\left(\varphi(X,Y) \in (A,B)\right) = P\left((X,Y) \in D\right) =$$

$$= \int\int_D \frac{\partial^2 H(x,y)}{\partial x \partial y}\; dx\; dy = \int\int_D \frac{\partial^2 C\left(F(x), G(y)\right)}{\partial u \partial v}\; F'(x)\; G'(y)\; dx\; dy. \tag{7}$$

The application

$$t:\, D \,\rightarrow\, D^*, \quad t(x,y) = \left(F(x), G(y)\right)$$

defines a pointwise transformation between D and D^* and its Jacobian is equal to $F'(x)\; G'(y)$. Using this transformation for a change of variables in the last integral of (7), we obtain

$$\int \cdots \int_D \frac{\partial^2 C\left(F(x), G(y)\right)}{\partial u \partial v}\; F'(x)\; G'(y)\; dx\; dy = \int \cdots \int_{D^*} \frac{\partial^2 C(u,v)}{\partial u \partial v}\; du\; dv. \tag{8}$$

From (7) and (8), the conclusion (5) results.

Remark 2.1. We fit a copula to the known sample $\left\{(x_i, y_i),\; i \in \overline{1,m}\right\}$ of the bidimensional random variable (X,Y). Some methods for fitting Archimedean copulas, based on Genest's and MacKay's results, are presented in [5]. The copula found in this way has the required differential properties imposed to C, and will replace C in the relation (5).

Remark 2.2.For a large class of functions $\varphi : \mathbf{R}^2 \rightarrow \mathbf{R}$, the form of the domain D^*,allows us to express the double integral of (5) as an iterated one.

Theorem 2.2.*If the functions* $x \mapsto F(x)$**,** $\quad y \mapsto G(y),\; (u,v) \mapsto C(u,v)$ *satisfies the conditions a) and b) and*

1) the function φ *is continuous on* $\mathbf{R}^2$ *and admits a partial derivative*

$$\frac{\partial \varphi(x,y)}{\partial y} > 0 \quad for\;\, (x,y) \in \mathbf{R}^2,$$

2) the functions $x \mapsto a(x)$ *and* $x \mapsto b(x)$ *defined on* $\mathbf{R}$ *exist, so that*

$$\forall x \in \mathbf{R} \quad \varphi(x,\ a(x)) = A, \quad \varphi(x,\ b(x)) = B,$$

then

$$P(Z \in (A,B)) = \int_0^1 du \int_{G(a(F^{-1}(u)))}^{G\left(b\left(F^{-1}(u)\right)\right)} \frac{\partial^2 C(u,v)}{\partial u \partial v}\, dv. \tag{9}$$

Note that the limits of the inner integral can be effectively determined, when the functions $x \mapsto a(x)$, $x \mapsto b(x)$ *and* $u \mapsto F^{-1}(u)$ *can be written as explicit functions (in a closed form).*

Proof. According to the hypotheses, the condition $A < B$ implies for each $x \in \mathbf{R}$, $a(x) < b(x)$. Since the inverse of an increasing function is also increasing, the following conditions are equivalent

$$(u,v) \in D^* \Leftrightarrow (x,y) \in D \Leftrightarrow x \in \mathbf{R}\, and\, a(x) < y < b(x) \Leftrightarrow$$

$$\Leftrightarrow F^{-1}(u) \in \mathbf{R}\, and\, a\left(F^{-1}(u)\right) < G^{-1}(v) < b\left(F^{-1}(u)\right) \Leftrightarrow$$

$$\Leftrightarrow 0 < u < 1 \text{ and } G\left(a\left(F^{-1}(u)\right)\right) < v < G\left(b\left(F^{-1}(u)\right)\right).$$

The double integral of (1) can be expressed as an iterated one

$$\int\int_{D^*} \frac{\partial^2 C(u,v)}{\partial u \partial v}\, du\, dv = \int_0^1 du \int_{G(a(F^{-1}(u)))}^{G\left(b\left(F^{-1}(u)\right)\right)} \frac{\partial^2 C(u,v)}{\partial u \partial v}\, dv$$

whence (2) results.

Remark 2.3.If in the hypothesis of the Theorem 2.2. the condition 1) is replaced by the condition

1') the function φ is continuous on $\mathbf{R}^2$ and admits a partial derivative

$$\frac{\partial \varphi(x,y)}{\partial y} < 0 \quad for\ (x,y) \in \mathbf{R}^2,$$

then in the relation (9) the limits of inner integral will be interchanged.

Remark 2.4.In the above theorem, the role of variable x and y, respectively of u and v can be interchanged.

Remark 2.5.In the particular case, when the function $\varphi : \mathbf{R}^2 \to \mathbf{R}$ is given by the relation $\varphi(X,Y) = X + Y$, we obtain

$$a(x) = A - x, \quad b(x) = B - x, \quad a\left(F^{-1}(u)\right) = A - F^{-1}(u), \quad b\left(F^{-1}(u)\right) = B - F^{-1}(u).$$

In this case, the formula (2) became

$$P\left(Z \in (A,B)\right) = \int_0^1 du \int_{G(B-F^{-1}(u))}^{G\left(A-F^{-1}(u)\right)} \frac{\partial^2 C(u,v)}{\partial u \partial v}\, dv.$$

3. The third section

Now, we will generalize the Theorem 2.1, for the case of a random variable $Z = \varphi(\mathbf{X})$, which is the transform of the n-dimensional random variable $\mathbf{X} = (X_1, \cdots, X_n)$ by the function $\varphi : \mathbf{R}^n \to \mathbf{R}$. We denote by H the joint distribution of $\mathbf{X}$

$$H(\mathbf{x}) = H(x_1, \cdots, x_n) = P(X_1 < x_1, \cdots, X_n < x_n)$$

which is unknown and by $F_i,\ i = \overline{1,n}$, its marginal distribution functions

$$F_i(x_i) = P(X_i < x_i)$$

which are known.

We also know a sample

$$\left\{\mathbf{x_j} = (x_{j1}, \cdots, x_{jn}), \quad j = \overline{1,m}\right\}$$

of the n-dimensional random variable $\mathbf{X}$.

We suppose that

$a')$ F_i are *strictly* monotonic and admit continuous derivatives of first order;

$b')$ the copula $C : [0,1]^n \to [0,1]$, which correspond to H, admits continuous n-th order partial derivatives.

We denote $\mathbf{x} = (x_1 \cdots, x_n) \in \mathbf{R}^n, \quad \mathbf{u} = (u_1, \cdots, u_n) \in [0,1]^n$,

$D = \{\mathbf{x} \in \mathbf{R}^n, \quad \varphi(x_1, \cdots, x_n) \in (A,B)\}$,

$D^* = \left\{\mathbf{u} \in [0,1]^n, \quad \varphi\left(F_1^{-1}(u_1), \cdots, F_n^{-1}(u_n)\right) \in (A,B)\right\}$.

Theorem 3.1. *If the functions* $x_i \mapsto F(x_i), \quad i = \overline{1,n}$ *and* $\mathbf{u} \mapsto C(u_1, \cdots, u_n)$ *satisfy the conditions* $a')$ *and* $b')$, *then the following equality holds*

$$P\left(Z \in (A,B)\right) = \int \cdots \int_{D^*} \frac{\partial^n C\left(u_1, \cdots u_n\right)}{\partial u_1 \cdots \partial u_n}\, du_1 \cdots du_n. \tag{10}$$

The proof of this theorem is similar with that of the Theorem 2.1 and is based on the relation (4) of the Sklar's Theorem 1.2, concerning the n-variate copulas.

Example.Let $X_i,\ i = \overline{1,n}$ be continuous random variables with the known strictly monotonic distribution functions F_i. We suppose that F_i have continuous derivatives of first order and their inverses F_i^{-1} can be written in a closed form.

We want to express, by means of copulas, the probability that the value of the sum

$$Z = \sum_{i=1}^{n} \alpha^i X_i; \quad \alpha^i > 0$$

belongs in a neighbourhood with radius $\delta,\ \delta > 0$ of the expected value of Z

$$\mu = M(Z) = \sum_{i=1}^{n} \alpha^i M(X_i).$$

Let C be the n-variate copula corresponding to the n-dimensional random variable $(X_1, \cdots, X_n)$,

$$D = \left\{ (x_1, \cdots, x_n) \in \mathbf{R}^n, \quad \left| \sum_{i=1}^{n} \alpha^i x_i - \mu \right| < \delta \right\}$$

and D^* the image of D, by the pointwise transformation defined by the relations $u_i = F_i(x_i), \quad i = \overline{1,n}$. According to (10), the equality

$$P\left(|Z - \mu| < \delta\right) = \int \cdots \int_{D^*} \frac{\partial^n C\left(u_1, \cdots u_n\right)}{\partial u_1 \cdots \partial u_n}\, du_1 \cdots du_n \tag{11}$$

holds.

The following equivalences takes place:

$$\begin{aligned} (x_1, \cdots, x_n) \ &\in \ D \Leftrightarrow x_i \in \mathbf{R},\ i = \overline{1, n-1}, \\ \frac{1}{\alpha^n}\left(\mu - \sum_{i=1}^{n} \alpha^i x_i - \delta\right) \ &< \ x_n < \frac{1}{\alpha^n}\left(\mu - \sum_{i=1}^{n} \alpha^i x_i + \delta\right) \end{aligned}$$

$$
\begin{aligned}
& \Leftrightarrow \quad F_i^{-1}(u_i) \in \mathbf{R},\ i=\overline{1,n-1}, \\
\frac{1}{\alpha^n}\left(\mu-\sum_{i=1}^{n}\alpha^i F_i^{-1}(u_i)-\delta\right) & < \quad F_n^{-1}(u_n) < \frac{1}{\alpha^n}\left(\mu-\sum_{i=1}^{n}\alpha^i F_i^{-1}(u_i)+\delta\right) \Leftrightarrow \\
& \Leftrightarrow \quad 0<u_i<1,\ i=\overline{1,n-1}, \\
a(\{u_i\}_{\overline{i=1;n-1}};\left\{\alpha^k\right\}_{k=\overline{1;n}}) & < \quad u_n < b(\{u_i\}_{i=\overline{1;n-1}};\left\{\alpha^k\right\}_{k=\overline{1;n}}) where
\end{aligned}
$$

$$
\begin{aligned}
a(\{u_i\}_{i=\overline{1;n-1}};\left\{\alpha^k\right\}_{k=\overline{1;n}}) &= F_n\left[\frac{1}{\alpha^n}\left(\mu-\sum_{i=1}^{n-1}\alpha^i F_i^{-1}(u_i)-\delta\right)\right], \\
b(\{u_i\}_{i=\overline{1;n-1}};\left\{\alpha^k\right\}_{k=\overline{1;n}}) &= F_n\left[\frac{1}{\alpha^n}\left(\mu-\sum_{i=1}^{n-1}\alpha^i F_i^{-1}(u_i)+\delta\right)\right].
\end{aligned}
$$

Writing the second member of (11) as an iterated integral, we obtain

$$
P(|Z-\mu|<\delta)=\int_0^1 du_1\cdots\int_0^1 du_{n-1}\int_{b(\{u_i\}_{i=\overline{1;n-1}};\left\{\alpha^k\right\}_{k=\overline{1;n}})}^{a(\{u_i\}_{i=\overline{1;n-1}};\left\{\alpha^k\right\}_{k=\overline{1;n}})}\frac{\partial^n C\left(u_1,\cdots u_n\right)}{\partial u_1\cdots\partial u_n}du_n.
$$

Remark 3.1.The procedure indicated in Remark 2.1 can be applied also in this general case.

4. The fourth section

In this section we present certain financial applications of our results and some conclusions.

One of the main issues of risk management is the aggregation of individual risks. The risk can split into two parts: the individual risks and the dependence structure between them. For this purpose the copula function is a powerful concept for describing the joint distribution of the aggregate risk.

In many cases is very useful to know how probable is that the aggregate risk belongs to a certain neighbourhood of its expected value (as in the previous example).

Another field of risk management, where we can use successfully the idea presented in this paper, is the determination of the value of risk, denoted "VaR ". One of the most frequent questions concerning the risk management in finance involves the estimation of extreme quantiles. This correspond to the determination of the value of a given random variable Z (obtained as a transform of other random variables X_i), which is exceeded with a given low probability q :

$P(Z > VaR) = q$, where $q = 0,05$ or $q = 0,01$ or other low probability.

This is equivalent to $P(Z < VaR) = 1 - q = p$.

The value at risk, called "risk capital" is generally defined as the capital which is sufficient, with a great probability (as 0,95 or 0,99), to cover losses from a portfolio.

Because $P(Z < a)$ is increasing with respect to a, we can determine successively the values of a, till we obtain $P(Z < a) = p$. This is a procedure to determinate the value of VaR.

Now we consider that the risk is represented by the random variable Z, from the example of the previous point. The value of VaR can be determined by the above procedure. But for that, we must estimate the probability $P(Z < a)$,as a function of a.

We take in the Theorem 3.1., $A = -\infty$, $B = a$ and obtain, as in previous example:

$$P(Z < a) = \int_0^1 du_1 \cdots \int_0^1 du_{n-1} \int_0^{b\left(\{u_i\}_{i=\overline{1;n-1}};\left\{\alpha^k\right\}_{k=\overline{1;n}}\right)} \frac{\partial^n C\left(u_1, \cdots u_n\right)}{\partial u_1 \cdots \partial u_n}\, du_n,$$

where $b\left(\{u_i\}_{i=\overline{1;n-1}};\left\{\alpha^k\right\}_{k=\overline{1;n}}\right) = F_n\left[\frac{1}{\alpha^n}\left(a - \sum_1^{n-1} \alpha^i F_i^{-1}(u_i)\right)\right].$

5.Conclusions

Because in practice the joint distribution is often not known, the copulas provide a very important tool to model this, by modeling a lot of types of dependences between the underlying random variables. For this purpose our results are useful since we can determine with them the probabilities, using only copulas and the marginal distribution functions. These are useful when the quantiles can be written in an explicit form.

Acknowledgment. I would like to express thanks to the head of our department, economist Dr. György Lipėcz, for the helpful discussions.

References

[1] Sklar, A.: *Random Variables, Distribution Functions and Copulas - a personal look backward and forward.* In Distribution with fixed Marginals and

related topics. Ed. by L. Rüschendorff, B. Schweizer and M. Taylor. pp.1-14. Institute of Mathematical Statistics, Hayward, C.A.

[2] Genest, C. and MacKay: *The Joy of Copulas: Bivariate Distribution with Uniform Marginals.* The American Statistician, 40, 1986, pp. 280-283 .

[3] Embrechts, P. Lintskog, F. and Mc Neil A.: *Modelling Dependence with Copulas and Applications to Risk Management.* Department of Mathematics EthZ September 10, 2001, pp.1-48.

[4] Nelsen, R.: *An Introduction to Copulas.* 1998. Springer, New York.

[5] Kovacs, E.: *Algorithms for the Estimation of Joint Distributions by Means of Copulas* (Hungarian). Proceedings of XXIX Conference of Math. Phys. and Comp. Sciences, Szeged 2005 (to be printed).

[6] Kovacs, E.:*Testing the Goodness-of-fit of a Bivariate Copula.* Annals of the Tiberiu Popoviciu Seminar, Cluj-Napoca (to be printed).

Edith Kovacs
Department of Mathematics
College of Management
Villanyi ut, 11
e-mail: kovacs.edith@villanyi.avf.hu