Scientific Research

Transfer of Global Measures of Dependence into Cumulative Local

Boyan Dimitrov[1], Sahib Esa[2], Nikolai Kolev[3], Georgios Pitselis[4]

[1]Kettering University, Flint, USA
[2]Salahaddin University, Erbil, Iraq
[3]University of Sao Paulo, Sao Paulo, Brazil
[4]University of Piraeus, Piraeus, Greece
Email: bdimitro@kettering.edu, sahib_esa@yahoo.com, nkolev@ime.usp.br, pitselis@unipi.gr

Abstract

We explore an idea of transferring some classic measures of global dependence between random variables $X_1, X_2, \cdots, X_n$ into cumulative measures of dependence relative at any point $(x_1, x_2, \cdots, x_n)$ in the sample space. It allows studying the behavior of these measures throughout the sample space, and better understanding and use of dependence. Some examples on popular copula distributions are also provided.

Keywords

Analysis of Variance; Copula; Correlation; Covariance; Multivariate Analysis; Measures of Dependence; Probability Modeling

1. Introduction

The ideas of transferring integral measure of dependence into local measures are probably not new. We got the clue from the works of [1]-[3]. Cooke raised the idea of the use of the indicator random variables (r.v.'s) in regression models to consider the regression coefficients as measures of local dependence. Genest [4] gave some more motivations on the need of transfer of overall (integrated) measures to the cumulative local functions. Dimitrov [2] revived some old measures of dependence between random events proposed by Obreshkov [5], and noticed their convenience in the use of the studies of local dependence between r.v.s. [6] is an attempt looking at the cumulative dependence from the point of view of the so called Sibuya function, which is also used in studies of dependence.

We propose here a review of the classic statistical analysis approach, where the procedures are more or less routine to be used in the analysis of local dependence. The manuscript is built in the following way: Section 2 introduces the idea of the transfer in detail. Section 3 considers the covariance and correlation matrices between indicator functions for a finite family of random events. At this point it is worth noticing that the expressions for the entries $\Sigma_{(A_i, A_j)}$ of the covariance matrix for identificators of random events $A_1, \cdots, A_n$ coincide with the quantity $\delta(A_i, A_j) = P(A_i \cap A_j) - P(A_i)P(A_j)$, introduced as *connection between the events* A_i *and* A_j. Similarly, the entries R_{A_i, A_j} of the correlation matrix coincide with the expression called in [5] as *correlation coefficient between the events* A_i *and* A_j. Obreshkov approach is using different categories. Following these ideas, [2] exposes a relatively complete analysis of the two-dimensional dependence between random events, and establishes some more detailed properties of the Obreshkov measures of dependence. Similar expressions for the correlation between random events, are used in [4]. They call it *correlation coefficient associated with pairs of dichotomous random variables*. Long before [7], p.311 uses *correlation coefficient between indicator functions of two events*, but does not elaborate any details.

Let $\mathcal{F}_1$ and $\mathcal{F}_2$ be two σ-sub-algebras of $\mathcal{F}$. The survey of [3] proposes the measure

$$\alpha(\mathcal{F}_1, \mathcal{F}_2) = \sup |P(A \cap B) - P(A)P(B)|, \quad A \in \mathcal{F}_1, B \in \mathcal{F}_2, \tag{1.1}$$

to be used as magnitude of dependence between $\mathcal{F}_1$ and $\mathcal{F}_2$. This measure is used for studying properties of a sequence $\{X_n\}_0^\infty$ of r.v.'s, when choosing $\mathcal{F}_1 = \mathcal{F}_{X_m}$ and $\mathcal{F}_2 = \mathcal{F}_{X_{m+n}}$ generated by two remote members— X_m and X_{m+n} of the sequence. The $\alpha(\mathcal{F}_1, \mathcal{F}_2) > 0$ is called *mixing*. The similarity with the Obreshkov's connection measure between events is obvious. However, this measure of dependence $\alpha(\mathcal{F}_1, \mathcal{F}_2)$ is a global measure limited within the interval $\left[-\frac{4}{4}, \frac{1}{4}\right]$, and does not have clear meaning of magnitude. Some of its properties are summarized in the statements of Section 3. Section 4 extends these results as integrated local measures of dependence between finite sequence of r.v.'s. Explicit presentations in terms of joint and marginal cumulative distribution functions (c.d.f.'s) are obtained. Some invariant properties are established in Section 5; Examples (explicit analytic and graphic) with popular copula distributions, and a bivariate Poisson distribution with dependent components illustrate these properties. 3-d surfaces with level curves consisting of points with constant cumulative magnitude of local dependence (covariance or correlation magnitude) in the sample space are drawn (Sections 5, 6, and 7).

2. The Idea for Measure Transfer

Let $X_1, X_2, \cdots, X_n$ be random variables (r.v.'s) defined on a probability space $(\Omega, \mathcal{F}, P)$ with joint cumulative distribution function (c.d.f.)

$$F(x_1, x_2, \cdots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \cdots, X_n \leq x_n)$$

and marginals

$$F_i(x_i) = P(X_i \leq x_i); \quad i = 1, 2, \cdots, n, \quad x_i \in \mathfrak{R} = (-\infty, \infty).$$

Most of the conventional measures of dependence between r.v.'s used in the practice are *integral* (*global*). Usually they have the form of expected values of some specifically selected function $g(x_1, \cdots, x_n)$. Sometimes $g(., \cdots, .)$ can be a real number, a vector, or a matrix. The dependence measure "of type g" is expressed by the equation

$$d_{1,2,\cdots,n} = E\left[g(X_1, X_2, \cdots, X_n)\right] = \int_{\mathfrak{R}^n} g(x_1, x_2, \cdots, x_n) dF(X_1, X_2, \cdots, X_n).$$

Let $A_1, A_2, \cdots, A_n$ be a finite set of random events in the sample space Ω of an experiment, where the r.v.'s X_i are defined. Introduce the indicator r.v.'s

$$I_{A_i}(\omega) = \begin{cases} 1, & \text{if } \omega \in A_i; \\ 0, & \text{elsewhere}, \end{cases} \quad i = 1, 2, \cdots, n.$$

Then the domain of the r.v.'s $I_{A_1}, I_{A_2}, \cdots, I_{A_n}$ is a subset of the sample space Ω of the r.v.'s $X_1, X_2, \cdots, X_n$.

Definition 1. The quantity

$$d_{1,2,\cdots,n}\left(A_1, A_2, \cdots, A_n\right) = E\left[g\left(I_{A_1}, I_{A_2}, \cdots, I_{A_n}\right)\right] = \int_{\Omega} g\left(I_{A_1}(\omega), I_{A_2}(\omega), \cdots, I_{A_n}(\omega)\right) \mathrm{d}P(\omega)$$

is called *measure of dependence of the type* $g(.)$ *between the random events* $A_1, A_2, \cdots, A_n$.

Assume, the measure of dependence of type $g(.)$ is introduced, and its properties are well established and understood. For any choice of the point $(x_1, x_2, \cdots, x_n) \in \Re^n$ define the random events

$$A_i(x_i) = \{\omega : X_i(\omega) \le x_i\}, \ i = 1, 2, \cdots, n. \tag{1.2}$$

Obviously, $P\{A_i(x_i)\}$ define the marginal distribution functions of the r.v.'s $X_i, i = 1, \cdots, n$. Put these events in Definition 1, and get naturally the next:

Definition 2. The quantities

$$d_{1,2,\cdots,n}\left(x_1, x_2, \cdots, x_n\right) = d_{1,2,\cdots,n}\left(A_1(x_1), A_2(x_2), \cdots, A_n(x_n)\right)$$

are called *cumulative local measure of dependence of type* $g(.)$ *between the* r.v.'s $X_1, X_2, \cdots, X_n$ *at the point* $x_1, x_2, \cdots, x_n$.

As an illustration of the above idea we consider the frequently used and well known integral measures within a finite sequence of r.v.'s, the covariance and correlation matrices.

3. Covariance and Correlation between Random Events

Introduce vector-matrices $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)$ and $\boldsymbol{\mu} = (\mu_1, \mu_2, \cdots, \mu_n) = E(\boldsymbol{X})$, and their transposed $\boldsymbol{X}^{\mathrm{T}}$ and $\boldsymbol{\mu}^{\mathrm{T}}$. The covariance matrix between r.v.'s $X_1, X_2, \cdots, X_n$ is

$$\boldsymbol{\Sigma} = E\left[(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})^{\mathrm{T}}\right]. \tag{1.3}$$

By implementing the idea of the previous section, we introduce *Covariance matrix between random events* $A_1, A_2, \cdots, A_n$.

First, observe that $E\left(I_{A_i}(\omega)\right) = P(A_i)$, and

$$Var\left(I_{A_i}\right) = E\left[I_{A_i}^2\right] - \left[P(A_i)\right]^2 = P(A_i)\left(1 - P(A_i)\right) = P(A_i)P\left(\overline{A_i}\right).$$

Introduce vectors

$$\boldsymbol{I}_A(\omega) = \left(I_{A_1}(\omega), I_{A_2}(\omega), \cdots, I_{A_n}(\omega)\right),$$

and

$$\boldsymbol{P}(A) = \left(P(A_1), P(A_2), \cdots, P(A_n)\right).$$

Formally, the covariance matrix $\boldsymbol{\Sigma}_{(A_i, A_2, \cdots, A_n)}$ between the random events $A_1, A_2, \cdots, A_n$ is defined by

$$\boldsymbol{\Sigma}_{(A_i, A_2, \cdots, A_n)} = E\left[(\boldsymbol{I}_A - \boldsymbol{P}(A))(\boldsymbol{I}_A - \boldsymbol{P}(A))^{\mathrm{T}}\right].$$

Its entries are

$$\begin{aligned}
\Sigma_{(A_1, A_j)} &= E\left[\left(I_{A_i} - P(A_i)\right)\left(I_{A_j} - P(A_j)\right)\right] \\
&= E\left(I_{A_i} I_{A_j}\right) - E\left(I_{A_i}\right)P(A_j) - E\left(I_{A_j}\right)P(A_i) + P(A_i)P(A_j) \\
&= E\left(I_{A_i \cap A_j}\right) - E\left(I_{A_i}\right)P(A_j) - E\left(I_{A_j}\right)P(A_i) + P(A_i)P(A_j) \\
&= P(A_i \cap A_j) - P(A_i)P(A_j).
\end{aligned} \tag{1.4}$$

On the main diagonal of the covariance matrix $\boldsymbol{\Sigma}_{(A_i, A_2, \cdots, A_n)}$ we recognize the entries

$$\sigma_{ii}^2 = \Sigma_{(A_i, A_i)} = P(A_i)P\left(\overline{A_i}\right) = Var\left(I_{A_i}\right), \tag{1.5}$$

where $\overline{A_i}$ is the complement event to the event A_i. By making use of the standard approach we introduce the correlation matrix between random events

$$\boldsymbol{R}_{A_1,A_2,\cdots,A_n} = \left\{ R_{A_i,A_j} \right\};\; i,j = 1,2,\cdots,n,$$

whose entries are the correlation coefficients between the indicators of respective events. Based on (1.4) and (1.5) for $i,j = 1,2,\cdots,n$, they have the form

$$R_{A_i,A_j} = \frac{Cov\left(I_{A_i}, I_{A_j}\right)}{\sigma_{ii}\sigma_{jj}} = \frac{P\left(A_i \cap A_j\right) - P\left(A_i\right)P\left(A_j\right)}{\sqrt{P\left(A_i\right)P\left(\overline{A_i}\right)P\left(A_j\right)P\left(\overline{A_j}\right)}}. \tag{1.6}$$

We call these entries *correlation coefficients between random events* A_i, *and* A_j. As one should expect, the entries on main diagonal are equal to 1. Each entry of the correlation matrix for $i,j = 1,2,\cdots,n$, satisfies the inequalities

$$-1 \leq \frac{P\left(A_i \cap A_j\right) - P\left(A_i\right)P\left(A_j\right)}{\sqrt{P\left(A_i\right)P\left(\overline{A_i}\right)}\sqrt{P\left(A_j\right)P\left(\overline{A_j}\right)}} \leq 1. \tag{1.7}$$

Let us focus on some basic properties of the entries of covariance and correlation matrices. They reveal also the information contained in terms of these matrices. Details can be seen in [2].

Theorem 1. a) The covariance $\Sigma_{\left(A_i,A_j\right)}$ and the correlation coefficient R_{A_i,A_j} between events A_i, and A_j equal 0 if and only if A_i, and A_j are independent. The case includes the situation when one of the events A_i, or A_j is the sure, or the impossible event (then the two events also are independent);

b) $R_{A_i,A_j} = 1$ if and only if events A_i, and A_j coincide;

c) $R_{A_i,A_j} = -1$ if and only if events A_i, and $\overline{A_j}$ coincide (or equivalently, if and only if the events A_j, and $\overline{A_i}$ coincide);

d) The covariances between events A_i, and A_j satisfy the inequalities

$$-\frac{1}{4} \leq -\sqrt{P\left(A_i\right)P\left(\overline{A_i}\right)P\left(A_j\right)P\left(\overline{A_j}\right)} \leq \Sigma_{\left(A_i,A_j\right)} \leq \sqrt{P\left(A_i\right)P\left(\overline{A_i}\right)P\left(A_j\right)P\left(\overline{A_j}\right)} \leq \frac{1}{4}. \tag{1.8}$$

The equalities in the middle of (0.8) hold only if the two events A_i, and A_j obey the hypotheses stated in parts a), b) and c);

e) The covariance and the correlation coefficients between the events A_i and A_j, and their complements $\overline{A_i}$ and $\overline{A_j}$ are related by the equalities

$$\Sigma_{A_i,\overline{A_j}} = -\Sigma_{A_i,A_j};\; R_{A_i,A_j} = -R_{A_i,\overline{A_j}},$$

and

$$\Sigma_{\left(\overline{A_i},\overline{A_j}\right)} = \Sigma_{\left(A_i,A_j\right)};\; R_{\overline{A_i},\overline{A_j}} = R_{A_i,A_j}.$$

The advantage of the considered measures is that they are completely determined by the individual (marginal) probabilities $P\left(A_i\right)$ of the random events, and by their joint probabilities $P\left(A_i \cap A_j\right)$. An interesting fact is that we do not need the joint probability to calculate conditional probabilities between the events. If either of the measures of dependence $\Sigma_{\left(A_i,A_j\right)}$ or R_{A_i,A_j} is known then the the conditional probability $P\left(A_i|A_j\right)$ can be found when marginal probabilities $P\left(A_i\right),\; P\left(A_i\right)$ are known. It holds

Theorem 2. Given the measures $\Sigma_{\left(A_i,A_j\right)}$ or R_{A_i,A_j}, and marginal probabilities $P\left(A_i\right),\; P\left(A_i\right)$, the conditional probabilities of one event, given the other, can be calculated by the equations

$$P\left(A_j|A_i\right) = \begin{cases} P\left(A_j\right) + \dfrac{\Sigma_{\left(A_i,A_j\right)}}{P\left(A_j\right)}, & \text{or} \\[3ex] P\left(A_j\right) + R_{A_i,A_j}\sqrt{\dfrac{P\left(\overline{A_i}\right)P\left(A_j\right)P\left(\overline{A_j}\right)}{P\left(A_i\right)}}. \end{cases} \tag{1.9}$$

Remark 1. If the covariance, or the correlation between two events A_i and A_j is positive, the conditional

probability of one event, given the other one, increases, compare to its unconditional probability, *i.e.* $P(A_j|A_i) > P(A_j)$ and the exact quantity of increase equals to

$$P(A_j|A_i) - P(A_j) = \begin{cases} \dfrac{\Sigma_{(A_i, A_j)}}{P(A_j)}, \text{ or} \\ R_{A_i, A_j} \sqrt{\dfrac{P(\overline{A_i})P(A_j)P(\overline{A_j})}{P(A_i)}}. \end{cases}$$ (1.10)

Conversely, if the correlation coefficient is negative, the information that one of the events has occurred, the chances that the other one occurs decrease. The net amount of decrease in probability is also given by (1.10).

4. Covariance and Correlation Functions Describe Cumulative Local Dependence for Random Variables

Here we continue developing the idea of Section 2 to illustrate how the cumulative local measures of dependence between r.v.'s are obtained and expressed in terms of the conventional probability distribution functions. Let $X_1, X_2, \cdots, X_n$ be arbitrary finite sequence of r.v.'s. For any choice of the real numbers $x_1, x_2, \cdots, x_n$, let $A_i(x_i)$ be the random events as defined in (1) $i = 1, \cdots, n$. The covariance and correlation matrices for random events $A_i(x_i)$ immediately turn into respective integrated measures of local dependence between the r.v.'s $X_1, X_2, \cdots, X_n$ at the selected point $(x_1, x_2, \cdots, x_n)$. We discuss these matrices here, and focus on their specific forms in terms of the respective distribution functions. The covariance matrix function

$$\mathbf{\Sigma}(x_1, x_2, \cdots, x_n) = \left\{ \Sigma_{ij}(x_i, x_j) \right\}_{i, j=1}^n,$$

whose entries are defined for any pair (i, j), $i, j = 1, 2, \cdots, n$ by the equations

$$\Sigma_{ij}(x_i, x_j) = P\{X_i \le x_i, X_j \le x_j\} - P\{X_i \le x_i\} P\{X_j \le x_j\} = F_{ij}(x_i, x_j) - F_i(x_i) F_j(x_j), \quad (1.11)$$

represents the cumulative local covariance structure between r.v.'s $X_1, X_2, \cdots, X_n$ at the point $(x_1, x_2, \cdots, x_n) \in \mathfrak{R}^n$. Analogously, we obtain the correlation matrix function

$$\mathbf{R}(x_1, x_2, \cdots, x_n) = \left\{ R_{ij}(x_i, x_j) \right\}_{i, j=1}^n,$$

whose entries according to Equation (1.6), are

$$R_{ij}(x_i, x_j) = \frac{F_{ij}(x_i, x_j) - F_i(x_i) F_j(x_j)}{\sqrt{F_i(x_i)[1 - F_i(x_i)]}\sqrt{F_j(x_j)[1 - F_j(x_j)]}}.$$ (1.12)

Obviously, these matrices are functions of the joint c.d.f.'s $F_{ij}(x_i, x_j)$ and of the marginals $F_i(x_i)$, $F_j(x_j)$ of the participating r.v.'s. The properties of these entries are described by Theorem 1, and Theorem 2. We notice here the following particular and important facts.

Theorem 3. a) The equality

$$\Sigma_{ij}(x_i, x_j) = R_{ij}(x_i, x_j) = 0$$

holds if and only if the two events $\{X_i \le x_i\}$, and $\{X_j \le x_j\}$ are independent at the point (x_i, x_j);

b) The equality

$$R_{ij}(x_i, x_j) = 1$$

holds if and only if the two events $\{X_i \le x_i\}$, and $\{X_j \le x_j\}$ coincide at the point (x_i, x_j);

c) The equality

$$R_{ij}(x_i, x_j) = -1$$

holds if and only if the two events $\{X_i \le x_i\}$, and the complement $\overline{\{X_j \le x_j\}} = \{X_j > x_j\}$ coincide at the

point (x_i, x_j). Then also the events $\{X_i > x_i\}$, and $\{X_j \leq x_j\}$ coincide.

d) Given the dependence function measures $\{\Sigma_{ij}(x_i, x_j)\}$ or $\{R_{ij}(x_i, x_j)\}$ and marginal c.d.f.'s $F_i(x_i)$, $F_j(x_j)$, the conditional probabilities $P\{X_j \leq x_j | X_i \leq x_i\}$, and $P\{X_j \leq x_j | X_i > x_i\}$ can be evaluated by the expressions

$$P\{X_j \leq x_j | X_i \leq x_i\} = \begin{cases} F_j(x_j) + \dfrac{\Sigma_{ij}(x_i, x_j)}{F_i(x_i)}, \\ \text{or} \\ F_j(x_j) + R_{ij}(x_i, x_j) \times \sqrt{\dfrac{\overline{F}_i(x_i) F_j(x_j) \overline{F}_j(x_j)}{F_i(x_i)}}, \end{cases}$$

and

$$P\{X_j \leq x_j | X_i > x_i\} = \begin{cases} F_j(x_j) - \dfrac{\Sigma_{ij}(x_i, x_j)}{1 - F_i(x_i)}, \\ \text{or} \\ F_j(x_j) - R_{ij}(x_i, x_j) \times \sqrt{\dfrac{\overline{F}_i(x_i) F_j(x_j) \overline{F}_j(x_j)}{1 - F_i(x_i)}}. \end{cases}$$

Proof. The statement is just an illustration of the transfer of the results of Theorems 1 and 2 from random events to the case of events depending on parameters. We omit details.

Theorem 3 d) provides the opportunity to make predictions for either of the events $\{X_i \leq x_i\}$, or $\{X_j \leq x_j\}$, given that the other event occurs, or does not occur. The absolute values of the increase (or decrease) of the chances of the r.v. X_j to get values less than a number x_j under conditions of known values of the variable X_i compare to these chances without conditions, can be evaluated as shown in Remark 1. The evaluation is expressed in terms of the local measures of dependence and the marginal c.d.f.'d of participating r.v.'s.

Remark 4. [8] introduces *positive quadrant dependence* between two r.v.'s by the requirement: for all $(x_i, x_j) \in \Re^2$ to be fulfilled $F_{ij}(x_i, x_j) \geq F_i(x_i) F_j(x_j)$. This is equivalent to the validity of $\Sigma_{ij}(x_i, x_j) \geq 0$ everywhere. The reverse inequality, when it remains true for all $(x_i, x_j) \in \Re^2$ Nelsen calls as *negative quadrant dependence*. Also there ([8] p. 189) the author gives some discussion about *Local Positive (and Local Negative) Quadrant Dependencies* between r.v.'s X and Y at a point $(x, y) \in \Re^2$, and does not go any further in its possible use. We are confident, that having local measure of dependence at a point, and especially, the measure of its magnitude given by the correlations functions $R_{ij}(x_i, x_j)$, offers the opportunity to use more. Namely, it is possible to classify the points in the plane according to their magnitude of dependence. For instance, there might be points of marginal independence, where the covariance equals zero. There can be areas of positive and negative dependence, depending on the sign of the covariance. There can be curves of constant cumulative correlations, e.g. curves given by the equations, like $R_{ij}(x_i, x_j) = 0.5$, or $R_{ij}(x_i, x_j) = -0.75$, etc. Our examples below provide effective proofs of such facts.

5. Monotone Transformations of the Random Variables and Related Dependence Functions. Copula

Here we study how the dependence functions behave under increasing or decreasing transformations. Before we move on, we introduce the following notations in order to make exposition clearer. We will use the notations $\Sigma_g(y_1, \cdots, y_n)$, and $R_g(y_1, \cdots, y_n)$ for the covariance and correlation matrices of the random variables $Y_1 = g_1(X_1), \cdots, Y_n = g_n(X_n)$, where $g_1(x_1), \cdots, g_n(x_n)$ are arbitrary functions defined on the range of the arguments $X_1, \cdots, X_n$. Also we introduce notations $\Sigma_{ij}(y_i, y_j)$ and $R_{ij}(y_i, y_j)$ to mark the variables for which the respective dependence functions are pertained.

Now we present the results with the following statement:

Theorem 4. Given arbitrary continuous functions $g_1(x_1), \cdots, g_n(x_n)$ on the supports of the continuous r.v.'s

$X_1, \cdots,$ and X_n respectively, such that $g_1^{-1}(y_1), \cdots,$ and $g_n^{-1}(y_n)$ do exist, we have:

If $g_1(x_1), \cdots,$ and $g_n(x_n)$ are all increasing functions, or if $g_1(x_1), \cdots, g_n(x_n)$ are all decreasing functions, then the values of the respective covariance and correlation functions at the respective points are related by the equations

$$\mathbf{\Sigma}_g(y_1, \cdots, y_n) = \mathbf{\Sigma}\left(g_1^{-1}(y_1), \cdots, g_n^{-1}(y_n)\right),$$

and

$$\mathbf{R}_g(y_1, \cdots, y_n) = \mathbf{R}\left(g_1^{-1}(y_1), \cdots, g_n^{-1}(y_n)\right).$$

Also the relations

$$\mathbf{\Sigma}(x_1, \cdots, x_n) = \mathbf{\Sigma}_g\left(g_1(x_1), \cdots, g_n(x_n)\right),$$

and

$$\mathbf{R}(x_1, \cdots, x_n) = \mathbf{R}_g\left(g_1(x_1), \cdots, g_n(x_n)\right)$$

are valid.

Proof. The statement follows from the fact that the following equivalence between the random events at any point $(x_1, \cdots, x_n)$ are valid:

If $g_i(x_i)$ and $g_j(x_j)$ are both increasing functions, then the following random events are equivalent

$$\left\{g_i(X_i) = Y_i \leq y_i\right\} = \left\{X_i \leq g_i^{-1}(y_i)\right\};$$
$$\left\{X_j \leq x_j\right\} = \left\{g_j(X_j) = Y_j \leq g_j(x_j)\right\},$$

and equivalent are also the events

$$\left\{Y_i \leq y_i, Y_j \leq y_j\right\} = \left\{g_i^{-1}(Y_i) = X_i \leq g_i^{-1}(y_i), g_j^{-1}(Y_j) = X_j \leq g_j^{-1}(y_j)\right\}$$

for all $i, j = 1, 2, \cdots, n$.

If $g_i(x_i)$ and $g_j(x_j)$ are both decreasing functions, then the following random events are equivalent

$$\left\{g_i(X_i) = Y_i \leq y_i\right\} = \left\{X_i \geq g_i^{-1}(y_i)\right\};$$
$$\left\{X_j \leq x_j\right\} = \left\{g_j(X_j) = Y_j \geq g_j(x_j)\right\},$$

and also equivalent are

$$\left\{Y_i \leq y_i, Y_j \leq y_j\right\} = \left\{g_i^{-1}(Y_i) = X_i \geq g_i^{-1}(y_i), g_j^{-1}(Y_j) = X_j \geq g_j^{-1}(y_j)\right\}$$
$$= \left\{\overline{g_i^{-1}(Y_i) = X_i \leq g_i^{-1}(y_i)}, \overline{g_j^{-1}(Y_j) = X_j \leq g_j^{-1}(y_j)}\right\}.$$

for all $i, j = 1, 2, \cdots, n$.

Next, we use the facts that the covariance and correlation matrix functions introduced in (1.11) and (1.12) for pairs of random events $\{Y_i \leq y_i\}$ and $\{Y_j \leq y_j\}$ are the same for any pairs of equivalent events $\left\{X_i \leq g_i^{-1}(y_i), X_j \leq g_j^{-1}(y_j)\right\}$, as well as for the pair of their complements $\left\{X_i \geq g_i^{-1}(y_i), X_j \geq g_j^{-1}(y_j)\right\}$, according to Lemma 2 part b) are equal. The reason is that the respective probabilities used in the definition of the matrix entries of equivalent events also are equal.

Assume now, that $X_1, \cdots, X_n$ are continuous r.v.'s, and $F(x_1, \cdots, x_n)$ is their continuous distribution function, as well as its all marginals are. According to [8], there exists unique *n*-dimensional copula $C(u_1, \cdots, u_n)$ (a *n*-dimensional distribution on the *n*-dimensional hyper-cube $[0,1] \times \cdots \times [0,1]$ with uniform 1-dimensional marginals and any 2- and higher-dimensional marginals also being copulas) associated with $F(x_1, \cdots, x_n)$, such that it is true

$$C(u_1, \cdots, u_n) = F\left(F_1^{-1}(u_1), \cdots, F_n^{-1}(u_n)\right), \; u_i \in [0,1], \; i = 1, \cdots, n. \tag{1.13}$$

Then it is true also

$$F(x_1,\cdots,x_n) = C(F_1(x_1),\cdots,F_n(x_n)),\ x_i \in \text{Range } X_i,\ i = 1,\cdots,n. \qquad (1.14)$$

Denote by $\boldsymbol{\Sigma}_C(u_1,\cdots,u_n)$ and by $\boldsymbol{R}_C(u_1,\cdots,u_n)$ the covariance and the correlation matrix functions corresponding to the uniformly distributed on $[0,1]$ r.v.'s $U_1,\cdots,U_n$ whose joint distribution function is the copula $C(u_1,\cdots,u_n)$. We introduce the notations $\boldsymbol{\Sigma}_F(x_1,\cdots,x_n)$ and $\boldsymbol{R}_F(x_1,\cdots,x_n)$ for the covariance and correlation matrix functions corresponding to the original r.v.'s $X_1,\cdots,X_n$.

The following statement holds:

Theorem 5. If $F(x_1,\cdots,x_n)$ is a continuous n-dimensional distribution function and $C(u_1,\cdots,u_n)$ is its associated copula, then the following relationships are true

$$\boldsymbol{\Sigma}_F(x_1,\cdots,x_n) = \boldsymbol{\Sigma}_C(F_1(x_1),\cdots,F_n(x_n));$$

$$\boldsymbol{R}_F(x_1,\cdots,x_n) = \boldsymbol{R}_C(F_1(x_1),\cdots,F_n(x_n)),$$

and also

$$\boldsymbol{\Sigma}_C(u_1,\cdots,u_n) = \boldsymbol{\Sigma}_F\left(F_1^{-1}(u_1),\cdots,F_n^{-1}(u_n)\right);$$

$$\boldsymbol{R}_C(u_1,\cdots,u_n) = \boldsymbol{R}_F\left(F_1^{-1}(u_1),\cdots,F_n^{-1}(u_n)\right).$$

Proof. The proof of these relationships is a simple consequence from Theorem 4, when we notice that the copula is just the distribution of a set of random variables, obtained after a monotonically increasing transformation $U_i = F_i(X_i)$ or $X_i = F_i^{-1}(U_i)$ is applied to each component of the original random vector $X = (X_1,\cdots,X_n)$, or to the vector $U = (U_1,\cdots,U_n)$.

Both theorems explain how after transformation the local cumulative dependence structure for the new variables is transfered to the points related with the same transformations at each coordinate of the other set of random variables. The fact of curiosity is that the relationships in Theorem 5 completely repeats the relationships (1.13) and (1.14) which define the copula.

The next examples illustrate the covariance and correlation functions for two popular multivariate copula, which we borrow from [8].

Examples

1) Farlie-Gumbel-Morgenstern (FGM) n-copula

The FGM *n*-copula is given by the equation

$$C(u_1,\cdots,u_n) = u_1\cdots u_n\left(1 + \sum_{k=2}^{n}\sum_{1 \le j_1 < \cdots < j_k \le n}\theta_{j_1\cdots j_k}\bar{u}_{j_1}\cdots\bar{u}_{j_k}\right)$$

where the coefficients $\theta_{j_1\cdots j_k} \in [0,1]$, and $\bar{u}_i = 1 - u_i;\ u_1,\cdots,u_n \in [0,1]$.

The entries of the covariance and the correlation matrices are respectively

$$\Sigma_{i,j}(u_i,u_j) = \theta_{ij}u_iu_j(1-u_i)(1-u_j),$$

and

$$R_{i,j}(u_i,u_j) = \theta_{ij}\sqrt{u_i(1-u_i)u_j(1-u_j)}.$$

We observe here, that the covariance and correlation functions keep constantly the sign of the parameter θ_{ij}.

Moreover, the correlation functions $R_{i,j}(u_i,u_j)$ do not exceed the numbers $\dfrac{\theta_{ij}}{4}$ for all $(u_i,u_j) \in \Re^2$.

2) Clayton Multivariate Archimedean copula

The copula is given for any $n \ge 2$, by the equation

$$C(u_1,\cdots,u_n) = \left(u_1^{-\theta} + \cdots + u_n^{-\theta} - n + 1\right)^{-\frac{1}{\theta}},$$

where $u_1,\cdots,u_n \in [0,1]$, and the parameter $\theta \ge 0$. For negative values of $\theta \in [-1,0)$ the maximum of the expression in the brackets and 0 must be taken, and $0 > \theta > -\dfrac{1}{n-1}$ must be fulfilled. For instance, if $\theta = -\dfrac{1}{3}$

we should have $n < 4$.

The entries of the covariance and the correlation matrices in case of positive θ, are respectively

$$\Sigma_{i,j}\left(u_i,u_j\right) = \left(\frac{1}{\left(u_i^\theta + u_j^\theta - u_i^\theta u_j^\theta\right)^{\frac{1}{\theta}}} - 1\right) \cdot u_i u_j,$$

and

$$R_{i,j}\left(u_i,u_j\right) = \left(\frac{1}{\left(u_i^\theta + u_j^\theta - u_i^\theta u_j^\theta\right)^{\frac{1}{\theta}}} - 1\right) \times \sqrt{\frac{u_i}{1-u_i} \cdot \frac{u_j}{1-u_j}}.$$

6. Maps for Dependence Structures on the Sample Space

A thorough analysis of the results in previous sections leads to an interesting opportunity to divide the sample space for any pair of r.v.'s $\left(X_i, X_j\right)$ into areas of positive, or negative dependence between the events $\{X_i \le x_i\}$ and $\{X_j \le x_j\}$. For this purpose, one can use the following curves:

Curve of Independence, C_0, the set of all points $\left(x_i, x_j\right)$ in $\Re^2$, where the correlation function $R_{ij}\left(x_i, x_j\right) = 0$. In other words,

$$C_0 := \left\{\left(x_i, x_j\right); F_{ij}\left(x_i, x_j\right) - F_i\left(x_i\right)F_j\left(x_j\right) = 0, \left(x_i, x_j\right) \in \Re^2\right\}.$$

Here the events $\{X_i \le x_i\}$ and $\{X_j \le x_j\}$ are independent for all points $\left(x_i, x_j\right) \in C_0$.

Curve of Positive Complete Dependence, C_1, the set of all points $\left(x_i, x_j\right) \in \Re^2$, where the correlation function $R_{ij}\left(x_i, x_j\right) = 1$. In other words,

$$C_1 := \left\{\left(x_i, x_j\right); F_{ij}\left(x_i, x_j\right) - F_i\left(x_i\right)F_j\left(x_j\right)\right\} = \sqrt{F_i\left(x_i\right)\overline{F}_i\left(x_i\right)F_j\left(x_j\right)\overline{F}_j\left(x_j\right)}.$$

Here the events $\{X_i \le x_i\}$ and $\{X_j \le x_j\}$ coincide for all points $\left(x_i, x_j\right) \in C_1$.

Curve of Negative Complete Dependence, C_{-1}, the set of all points $\left(x_i, x_j\right) \in \Re^2$, where the correlation function $R_{ij}\left(x_i, x_j\right) = -1$. In other words,

$$C_{-1} := \left\{\left(x_i, x_j\right); F_{ij}\left(x_i, x_j\right) - F_i\left(x_i\right)F_j\left(x_j\right)\right\} = -\sqrt{F_i\left(x_i\right)\overline{F}_i\left(x_i\right)F_j\left(x_j\right)\overline{F}_j\left(x_j\right)}.$$

Here either of the events $\{X_i \le x_i\}$ and $\{X_j \le x_j\}$ is equivalent to the complement of the other one for all points $\left(x_i, x_j\right) \in C_{-1}$.

Additional mapping can be drown by making use of the *Correlation Level Curves at level* α:

$$C_\alpha := \left\{\left(x_i, x_j\right); F_{ij}\left(x_i, x_j\right) - F_i\left(x_i\right)F_j\left(x_j\right)\right\} = \alpha\sqrt{F_i\left(x_i\right)\overline{F}_i\left(x_i\right)F_j\left(x_j\right)\overline{F}_j\left(x_j\right)}.$$

At each point of such a curve $\left(x_i, x_j\right) \in C_\alpha$, the correlation function $R_{ij}\left(x_i, x_j\right)$ between the events $\{X_i \le x_i\}$ and $\{X_j \le x_j\}$ equals to α. and $\alpha \in [-1, 1]$.

When α is positive, the random variables $\left(X_i, X_j\right)$ are positively associated at the points $\left(x_i, x_j\right) \in C_\alpha$. When α is negative, the random variables $\left(X_i, X_j\right)$ are negatively associated ate the points $\left(x_i, x_j\right) \in C_\alpha$.

In particular applications such curves may have important impact on understanding the dependence structure between the random variables $\left(X_i, X_j\right)$. By the way, *some of these curves may happen to be empty sets*.

One can ilustrate these curves in the next examples, using 3-d plots of the surfaces $z = \Sigma_{X,Y}\left(x, y\right)$ and $z = R_{X,Y}\left(x, y\right)$ with the level curves. In our plots we have used the program system Maple, and graph just the correlation function.

Examples

Here we illustrate the information one may visually get from the mapping of the 3-d graphs for the considered

above copula. We plot their correlation function $R_{i,j}(u,v)$, with $(u,v) \in [0,1] \times [0,1]$. Since the pictures are the same for any pair of indices i, j, we use notations without any indices.

1) The correlation functions for FGM copula (continued)

The two-dimensional marginal distributions of the multivariate FGM copula are given by one and the same equation, and we drop the subscripts:

$$C(u,v) = uv(1 + \theta(1-u)(1-v)),$$

where $\theta \in [-1,1]$.

The correlation function is respectively

$$R(u,v) = \theta\sqrt{u(1-u)v(1-v)},$$

and satisfies the inequalities $-\dfrac{\theta}{4} \leq R(u,v) \leq \dfrac{\theta}{4}$. On **Figure 1** we see the 3-dimensional graph of the surface $z = R(u,v)$, for the cases $\theta = -1$, and $\theta = 1$.

2) Clayton Archimedean copula (continued)

The two-dimensional marginal distributions of the multivariate Clayton copula also are given by the same equations, and we drop the subscripts:

The copula is given by the equation

$$C(u,v) = \left(\max\left(u^{-\theta} + v^{-\theta} - 1, 0\right) \right)^{-\frac{1}{\theta}},$$

and in the multivariate case it is required that $\theta > 0$. However the Clayton copula in the two-dimensional case is defined for all $\theta \geq -1$, and $\theta \neq 0$.

The respective correlation function is

$$R(u,v) = \left(\max\left(\frac{1}{\left(u^\theta + v^\theta - u^\theta v^\theta\right)^{\frac{1}{\theta}}}, 0 \right) - 1 \right) \times \sqrt{\frac{u}{1-u}\frac{v}{1-v}},$$

and this relation is valid for all $u, v \in [0,1]$.

To have a base for comparison (with the pictures in [8] p. 120, Figure 4.2), we consider the case $\theta = 4$, and present several level curves on the 3-d graph of $R(u,v)$. Actually, the level curves $R(u,v) = \alpha$ are visible on the surface from which we see that α is between 0 and 1.0. Results are pictured of **Figure 2**.

The 3-d graphs for this particular measure shown on **Figure 2(b)** of dependence between the two components of the copula closely imitate the situations given in the scatter-plots in the reference book [8], but are more impressive and information giving.

Together with the surface plots, the correlation function and its level curves are quite more informative than the simulated scatter-plots used in copula.
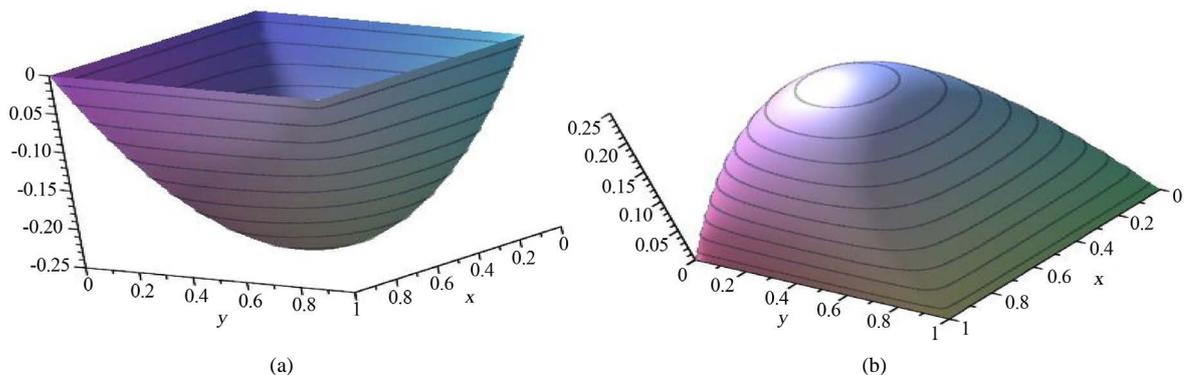

(a)   (b)

Figure 1. Correlation function for the Farlie-Gumbel-Morgenstern copula with level curves, $\theta = \pm 1$. (a) FGM-copula correlation function surface, $\theta = -1$; (b) FGM copula correlation function surface, $\theta = 1$.
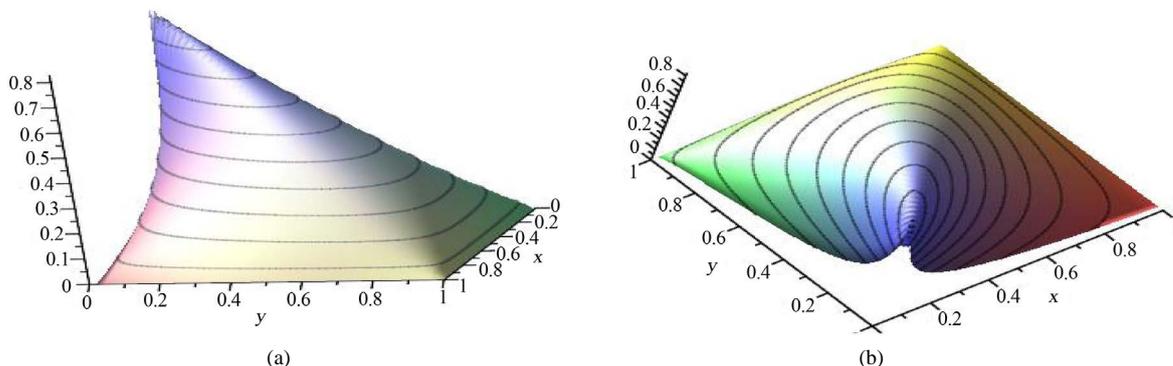
Figure 2. Correlation function for the Clayton copula, $\theta = 4$. (a) Clayton copula correlation function surface-a look from (1,1) corner; (b) Clayton copula correlation function surface-a look from (9,0) corner.

7. On the Use of the New Measures

At this point we believe that the introduced local measures of dependence can be used in similar ways as the covariance and correlation matrices are used in the multivariate analysis. However, because we think that the ideas developed here are so natural and simple, there will be no wonder if what we propose here is already in use. But, with the best of our knowledge, we have never seen anything similar, and we dare to say it under the risk we are not the first who offers such approach.

The results of Section 3 for random events may suit analysis of non-numerical, categorical, and finite discrete cases of random data similar to the ways as the numeric data are treated.

The results of Section 4 show that we are getting an opportunity to treat mixed discrete and continuous variables equally, using the cumulative distribution functions (marginal, and the two-dimensional).

If one treats the induces $i, j, \cdots$ of the events A_i, A_j as time variables (time is $i, j = 1, 2, \cdots$), the above approach may offer interesting applications in the areas of *Time Series Analysis*, as well as in the studies of sequences of random variables.

Obtained results may have impact on various side effect questions, and we leave it for other's comments. We insist on main practical orientation of our results-same and similar as everything else build on the use of the covariance and correlation matrices in the multivariate analysis.

8. Further Expansion of the Idea of Transforming Measures

In addition, the new measures may be used in studying the local structures of dependence within rectangular areas of the sample spaces when the sets of random events are defined as type of Cartesian products

$$\mathcal{X} := \left\{ \omega; \ \prod_{i=0}^{n} \left\{ x_i^{(1)} \leq X_i(\omega) \leq x_i^{(2)} \right\} \right\},$$

and by making use of the main idea of the transfer from the Introduction.

Another extension that may allow the involvement of the probability density functions (p.d.f.) could be obtained as in Section 4, by considering the elementary events

$$A_i(x_i) = \left[x_i < X_i(\omega) \leq x_i + dx_i \right], \quad i = 1, \cdots, n,$$

whose probabilities are

$$P\left\{ A_i(x_i) \right\} = f_i(x_i) dx_i,$$

$$P\left\{ A_i(x_i) \cap A_j(x_j) \right\} = f_{ij}(x_i, x_j) dx_i dx_j.$$

Here $f_i(x_i), f_j(x_j)$ and $f_{ij}(x_i, x_j)$ are the marginal and the joint probability densities of the r.v.'s X_i, X_j and of the pair (X_i, X_j). Since the densities do not have meaning of probabilities, for now we cannot use them in similar constructions. However, the picture changes in case of discrete variables.

As an illustration of this approach in the study of local dependence structure we consider the discrete case.

The r.v.'s X and Y have joint distribution $P(X = k, Y = m) = p(k,m)$ and marginals $P(x = k) = p_X(k)$ and $P(Y = m) = p_Y(m)$, $k, m = 0, 1, \cdots$. Then the covariance function $\Sigma_{X,Y}(k,m)$ is defined by the equalities

$$\Sigma_{X,Y}(k,m) = p(k,m) - p_X(k) \cdot p_Y(m), \quad m = 0, 1, 2, \cdots.$$

The respective correlation function will be defined by

$$R_{X,Y}(k,m) = \frac{p(k,m) - p_X(k) \cdot p_Y(m)}{\sqrt{p_X(k)[1 - p_X(k)] p_Y(m)[1 - p_Y(m)]}} \quad k, m = 0, 1, 2, \cdots.$$

Graphing these functions, one can see the local dependence structure between the two variables at any point (k,m) and to evaluate its magnitude, and in addition will see the entire picture of this mutual dependence in the field-the range of the pair (X,Y).

As an illustration of this suggestion we consider the simplest Bivariate Poisson distribution constructed in the following way, borrowed from [9]. There are three Poisson distributed r.v.'s T_i of parameters $\lambda_i > 0$, $i = 1, 2, 3$. Then it is supposed $X = T_1 + T_3$, $Y = T_2 + T_3$. The dependence between X and Y comes from the involvement of T_3 in both components. Their marginal distributions are Poisson with parameters $\lambda_1 + \lambda_3$ and $\lambda_2 + \lambda_3$ correspondingly. The joint distribution is given by

$$P(X = k, Y = m) = \mathrm{e}^{-\lambda_1 - \lambda_2 - \lambda_3} \frac{\lambda_1^k}{k!} \frac{\lambda_2^m}{m!} \times \sum_{n=0}^{\min(k,m)} \binom{k}{n} \binom{m}{n} n! \frac{\lambda_3}{\lambda_1 \cdot \lambda_2} \quad k, m = 0, 1, 2, \cdots.$$

After substitution of this form of the joint distribution and the marginal Poisson distributions in the above expressions, we get the covariance and correlation functions and the study of the local dependence can start. We omit detailed analytical expressions for these functions which is not a problem to get but are too cumbersome to write here.

On **Figure 3** we give an illustration of these functions by their 3-d graphs in the case when $\lambda_1 = 3$, $\lambda_2 = 2$, and $\lambda_3 = 5$. The magnitude of dependence is not strong (the local correlation coefficient is too small). However, we observe that on the square $[0,5] \times [0,5]$ the association is positive, when on the rectangle $[6,13] \times [3,14]$ this dependence is negative and also is a bit stronger. Outside these areas the two variables are almost independent.

More exotic spots of dependences can be obtained when in Definition 2 the events $A_i(X_i)$ are chosen as

$$A_i(X_i) = \{\omega; X_i(\omega) \in B_i\},$$

where B_i are any Borel sets on the real line, with $P\{A_i(X_i)\} \in (0,1)$.

In our opinion, there is a number of follow up questions in the multivariate dependence analysis, and more illustrations on real statistical data will trace the utility of the offered approach.

9. Conclusions

In this paper we discuss the possibility to transfer conventional global measures of dependence onto measures of
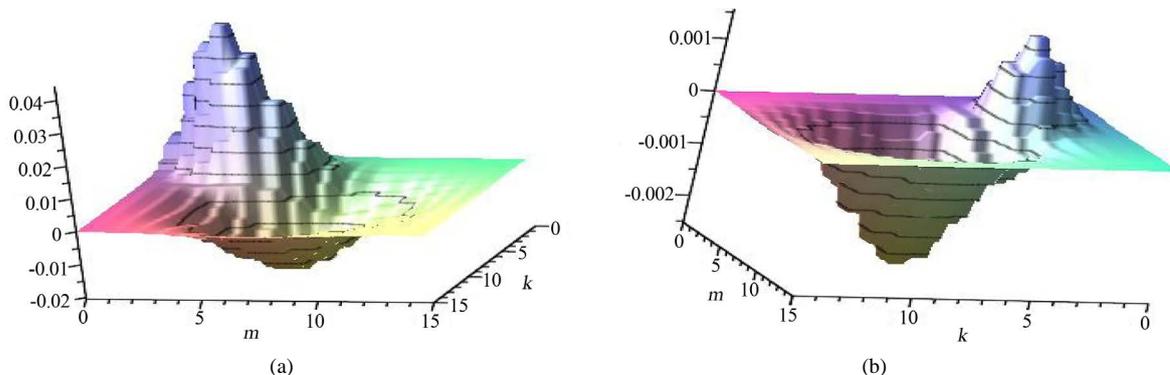

Figure 3. Surfaces explain local dependence structure in the Bivariate Poisson distribution. (a) Dependent Poisson-Covariance function; (b) Dependent Poisson-Correlation function.

dependence between random events, and then using another transfer to turn them into cumulative local measures of dependence between random variables.

We illustrate the method on the example of covariance and correlation matrices between random variables. Important results can be found in the discussion on some specific properties of the new measures, and in the qualitative information that may be derived from the meaning of these measures.

The behavior of these measures under monotonic transformations is established. It shows full coincidence between the pictures in the sample space of the original random variables and their respective copulas.

We found that locally the random variables may have areas of positive dependence and areas of negative dependence, as well as areas of local independence. This seems something interesting from informational point of view compared to the global measures of dependence.

Our suggestion is that the local measures of dependence offer important topological pictures to study. Our examples illustrate the new opportunities in this field.

Acknowledgements

References

[1] Cooke, R.M., Morales, O. and Kurowitcka, D. (2007) Vine in Overview. *Proceedings of* 3*rd Brazilian Conference on Statist. Modeling in Insurance and Finance*, 26-30 March 2007, 2-20.

[2] Dimitrovm B. (2010) Some Obreshkov Measures of Dependence and Their Use. *Compte Rendus de l'Academie Bulgare des Sciences*, **63**, 15-18.

[3] Bradley, R.C. (2005) Basic Properties of Strong Mixing Conditions. A Survey and Some Open Questions. *Probability Surveys*, **2**, 107-104.

[4] Genest, C. and Boies, J. (2003) Testing Dependence with Kendall Plots. *The American Statistician*, **44**, 275-284. http://dx.doi.org/10.1198/0003130032431

[5] Obreshkov, N. (1963) Probability Theory. Nauka i Izkustvo, Sofia (in Bulgarian).

[6] Dimitrov, B., Kolev, N. and Goncalves, M. (2007) Some Measures of Dependence between Two Random Variables. Work Paper.

[7] Renyi, A. (1969) Foundations of Probability Theory. Holden-Day, San Francisco.

[8] Nelsen, R. (1999) An Introduction to Copulas. Springer, New York. http://dx.doi.org/10.1007/978-1-4757-3076-0

[9] Kocherlakota, S. (1988) On the Compounded Bivariate Poisson Distribution: A Unified Treatment. *Annals of the Institute of Statistical Mathematics*, **49**, 61-70. http://dx.doi.org/10.1007/BF00053955