TWMS J. App. Eng. Math. V.1, N.1, 2011, pp. 86-97

ON GENERALIZED SARMANOV BIVARIATE DISTRIBUTIONS

I. BAIRAMOV[1], B. (YAĞCI) ALTINSOY[2], G. JAY KERNS[3] §

Abstract. A class of bivariate distributions which generalizes the Sarmanov class is introduced. This class possesses a simple analytical form and desirable dependence properties. The admissible range for association parameter for given bivariate distributions are derived and the range for correlation coefficients are also presented.

Keywords: Sarmanov class of bivariate distributions, range for association parameter, correlation coefficient.

AMS Subject Classification: 62H20, 62G30

1. Introduction

Let (X, Y) be a bivariate random vector. Sarmanov (1966) introduced a class of bivariate distributions for (X, Y) having a joint probability density function (pdf) of the form

$$h_\alpha(x, y) = f_X(x) f_Y(y) \{1 + \alpha\psi_1(x)\psi_2(y)\}, \quad (1)$$

where $f_X(x)$ and $f_Y(y)$ are the marginal pdf's of X and Y, respectively, $\psi_1(x)$ and $\psi_2(y)$ are bounded nonconstant functions such that

$$\int_{-\infty}^{\infty} f_X(t)\psi_1(t)dt = 0, \qquad \int_{-\infty}^{\infty} f_Y(t)\psi_2(t)dt = 0.$$

The association parameter α is a real number which satisfies the condition $1+\alpha\psi_1(x)\psi_2(y) \geq 0$ for all x and y. In this paper, we deal with the concept of a copula. A two-dimensional copula is a function $C(x, y)$ from $[0, 1] \times [0, 1]$ to $[0, 1]$ with the properties:

1. $C(x, 0) = 0 = C(0, y)$, $C(x, 1) = x$ and $C(1, y) = y$;
2. For every x_1, x_2 and y_1, y_2 such that $0 \leq x_1 < x_2 \leq 1$ and $0 \leq y_1 < y_2 \leq 1$, we have

$$C(x_2, y_2) - C(x_2, y_1) - C(x_1, y_2) + C(x_1, y_1) \geq 0.$$

[1] Izmir University of Economics, Department of Mathematics, 35330, Balcova, Izmir, Turkey,
e-mail: ismihan.bayramoglu@ieu.edu.tr

[2] Ministry of Transport and Communication of the Republic of Turkey, Department of Strategy and Development, Hakkı Turayliç Caddesi No:5 Emek, Ankara, Turkey,
e-mail: banuyagci@gmail.com

[3] Youngstown University, Department of Mathematics and Statistics, Youngstown, Ohio 44555-0002, USA,
e-mail: gkerns@ysu.edu

§ Manuscript received 15 May 2011.

According to Sklar's Theorem, if $F(x,y)$ is a joint distribution function with continuous marginals $F_X(x)$ and $F_Y(y)$, then there exists a unique copula C such that $F(x,y) = C(F_X(x), F_Y(y))$. Theory and applications of copulas are well documented in Nelsen (1998).

The copula generated by (1) has the form:

$$C(x,y) = xy + \int_0^x\int_0^y \psi_1^*(t)\psi_2^*(s)\, dt ds,$$

where $\psi_1^*(t) = \psi_1(F_X^{-1}(t))$ and $\psi_2^*(s) = \psi_2(F_Y^{-1}(s))$.

One may note in the special case $\psi_1(x) = 1 - 2F_X(x)$ and $\psi_2(y) = 1 - 2F_Y(y)$ that the classical Farlie-Gumbel-Morgenstern (FGM) distributions are recovered. The range of the correlation coefficient for the FGM copula is $-1/3 \leq \rho \leq 1/3$. Although the FGM model is an interesting family constructed from specified marginals, this model cannot be used to represent the joint distribution of two highly correlated variables. To increase dependence between variables, Huang and Kotz (1999) considered polynomial-type single parameter extensions of FGM (with uniform marginals):

$$C_\alpha(x,y) = xy\left\{1 + \alpha(1-x^p)(1-y^p)\right\} \ , p \geq 1 \ , 0 \leq x, y \leq 1 \tag{2}$$

and

$$C_\alpha^1(x,y) = xy\left\{1 + \alpha(1-x)^q(1-y)^q\right\} \ , q \geq 1 \ , 0 \leq x, y \leq 1. \tag{3}$$

The maximal positive correlation for (2), namely $\rho = 3/8$, is attained for $p = 2$, an improvement over the case $p = 1$ for which $\rho = 1/3$. Bairamov and Kotz (2001) provided several theorems characterizing symmetry and dependence properties of FGM distributions. They also proposed a modification by introducing additional parameters p and q:

$$F_{p,q,\alpha}(x,y) = xy\left\{1 + \alpha(1-x^p)^q(1-y^p)^q\right\} \ , q > 1 \ , p \geq 1 \ , 0 \leq x, y \leq 1. \tag{4}$$

For (4) the admissible range of α is

$$-\min\left\{\frac{1}{p^2}\left(\frac{1+pq}{p(q-1)}\right)^{2(q-1)}, 1\right\} \leq \alpha \leq \frac{1}{p}\left(\frac{1+pq}{p(q-1)}\right)^{q-1}.$$

The maximal and minimal values of ρ are within the range

$$-12t^2(q,p)\min\left\{\frac{1}{p^2}\left(\frac{1+pq}{pq-p}\right)^{2(q-1)}, 1\right\} \leq \rho \leq 12t^2(q,p)\frac{1}{p}\left(\frac{1+pq}{pq-p}\right)^{q-1}, \tag{5}$$

where $t(x,y) = \frac{\Gamma(x+1)\Gamma\left(\frac{2}{y}\right)}{y\Gamma\left(x+1+\frac{2}{y}\right)}$ and $\Gamma(x)$ is the Gamma function. In this case, the maximal positive correlation is $\rho_{\max} = 0.5021$ attained at $q = 1.496$ and $p = 3$. Hence, the extension (4) can achieve correlation greater than $1/2$ compared to the classical FGM where the correlation cannot be greater than $1/3$. Bairamov *et al.* (2001) considered a modification of the form

$$\begin{aligned} F_{p_1,p_2,q_1,q_2,\alpha}(x,y) &= xy\left[1 + \alpha(1-x^{p_1})^{q_1}(1-y^{p_2})^{q_2}\right] \ , \\ p_1, p_2 &\geq 1; \ q_1, q_2 > 1 \ ; \ 0 \leq x, y \leq 1 \end{aligned} \tag{6}$$

which yielded a slight improvement for the correlation coefficient over the case (4). They also derived recurrence relations between moments of concomitants of order statistics for this class of bivariate distributions.

Lee (1996) considered, for the Sarmanov model, "kernels" of the type $\psi_1(x) = x - \mu_X$ and $\psi_2(y) = y - \mu_Y$, where $\mu_X = E(X)$ and $\mu_Y = E(Y)$. Lee showed that the range of the correlation coefficient for this family of distributions is determined by both the marginals as

$$\max\left\{-\frac{1}{\mu_X\mu_Y}, \frac{1}{(1-\mu_X)(1-\mu_Y)}\right\} \leq \alpha \leq \min\left\{\frac{1}{\mu_X(1-\mu_Y)}, \frac{1}{(1-\mu_X)\mu_Y}\right\}. \tag{7}$$

For selected distributions, one can obtain from (7) the maximum positive (negative) correlation. However, for uniform marginals, the range for the correlation coefficient is the same as the FGM copula, *i.e.* $[-1/3,\ 1/3]$. Bairamov *et al.* (2001) considered kernels $\psi_1(x) = \psi_2(x) = x^p e^{-nx^q} - (1-x)^p e^{-n(1-x)^q}$, where n, p, and q are positive real numbers. In this case, the maximal correlation coefficient $\rho_{\max} = 0.59$ is achieved for $n = 2,\ p = 3$, and $q = 2$.

Yue *et al.* (2001) reviewed various bivariate models constructed from gamma marginals and applied them to hydrological frequency analysis. They noticed that many bivariate models have mainly remained in the form of theoretical developments and seldom succeeded in gaining popularity among practitioners in the field of hydrological frequency analysis. The main reason for this is that the mathematical expressions of some of these models are complex and therefore have computational limitations. Recently, Lin and Huang (2011) consider a generalized version of Sarmanov family introduced earlier by Bairamov et al. (2001) and show that unlike the traditional Sarmanov the generalized one always has a correlation approaching one regardless of the marginals, as long as the marginals are of the same type.

In this paper, we construct classes of bivariate distributions which are generalizations of Sarmanov and Sarmanov-Lee models. These distributions have a simple analytical form like the FGM models and, as in the "normal" case, the correlation coefficient ρ totally governs the dependence between the variables. Some dependence properties of these distributions are also discussed.

2. The generalized Sarmanov copula

Consider a class of bivariate distributions with absolutely continuous marginals F_X, F_Y given by the joint density

$$f_{X,Y}(x,y) = f_X(x)f_Y(y)\left\{1 + \alpha H[F(x), F_Y(y)]\right\}, \tag{8}$$

where $f_X(x), f_Y(y)$ are specified marginal densities, $H(x,y)$ is an integrable function defined on $[0,1]$ satisfying

$$\int_0^1 H(x,y)\,dx = 0,\ \int_0^1 H(x,y)dy = 0, \tag{9}$$

and α is a real number satisfying the condition that $1+\alpha H(x,y)\geq 0$ for all $0\leq x,y\leq 1$. The corresponding bivariate distribution function has the form

$$F_{X,Y}(x,y)=F_X(x)F_Y(y)+\alpha\int_0^{F_X(x)}\int_0^{F_Y(y)}H(t,s)dtds,\quad -\infty<x,y<\infty$$

and the copula is

$$C_{X,Y}(x,y)=xy+\alpha\int_0^x\int_0^y H(t,s)\,dtds,\ 0\leq x,y\leq 1. \tag{10}$$

Define the positive and negative domains D^+ and D^- by

$$D^+=\{(x,y):0\leq x,y\leq 1, H(x,y)>0\}$$

and

$$D^-=\{(x,y):\ 0\leq x,y\leq 1, H(x,y)<0\}\,,$$

respectively. The admissible range for α ensuring that (8) is a joint density can be expressed as

$$-\frac{1}{\max\{H(x,y)\,|\,(x,y)\in D^+\}}\leq\alpha\leq\frac{1}{\max\{-H(x,y)\,|\,(x,y)\in D^-\}}$$

$$-\frac{1}{\max\limits_{(x,y)\in D^+}(H(x,y))}\leq\alpha\leq\frac{1}{\max\limits_{(x,y)\in D^-}(-H(x,y))}. \tag{11}$$

The correlation coefficient of X and Y, if it exists, is given by

$$\rho=\frac{\alpha}{\sigma_X\sigma_Y}\int_0^1\int_0^1 F_X^{-1}(x)F_Y^{-1}(y)H(x,y)dxdy, \tag{12}$$

where $F_X^{-1}(x)=\inf\{y:\ F_X(x)\geq y\}$ is the inverse cumulative distribution function for F_X .

Therefore in family (8), as in the "normal" case, the correlation coefficient ρ totally governs the dependence. Using the Cauchy-Schwarz inequality, we have

$$|\rho|\leq|\alpha|\left(1+\frac{\mu_X^2}{\sigma_X^2}\right)^{1/2}\left(1+\frac{\mu_Y^2}{\sigma_Y^2}\right)^{1/2}\left(\int_0^1\int_0^1 H^2(x,y)\,dxdy\right)^{1/2}.$$

The main justification for the interest in the family (8) is the constructive approach involved in its definition and the transparent manner in which the dependence is created. It can be shown that the family possesses the TP_2 property, which is the strongest among the dependence properties; see, for example, Shaked and Shantikumar (1994).

The pair of variables (X,Y) is said to be *totally positive dependent* (TP_2) if its joint density $f(x,y)$ satisfies the condition

$$f(x_1,y_1)f(x_2,y_2)-f(x_2,y_1)f(x_1,y_2)\geq 0$$

for all $x_1<x_2$ and $y_1<y_2$. It is well known that – for more details on dependence properties see Barlow and Proschan (1981) and Joe (1990) – the TP_2 property implies the following important dependence properties:

(1) Stochastically increasing positive dependence, SI$(X \mid Y)$: (X is stochastically increasing in Y if $P\{X > x \mid Y = y\}$ is a nondecreasing function of y for all x);
(2) Right-tail increasing - RTI$(X \mid Y)$ (X is right-tail increasing in Y if $P\{X > x \mid Y > y\}$ is increasing in y for each x);
(3) $A(X, Y)-$association: (the random vector (X, Y) is associated if the inequality

$$E\left[g_1(X,Y)g_2(X,Y)\right] \geq E\left[g_1(X,Y)\right] E\left[g_2(X,Y)\right]$$

holds for all real-valued functions g_1, g_2 which are increasing in each component and such that the expectations exist);
(4) Positive quadrant dependence - PQD (the random vector (X, Y) is positive quadrant dependent if $P\{X \leq x, Y \leq y\} \geq P\{X \leq x\} P\{Y \leq y\}$).

For the points $0 \leq x_1 < x_2 \leq 1$ and $0 \leq y_1 < y_2 \leq 1$, denote $A(x_1, x_2, y_1, y_2) = H(x_2, y_2) - H(x_2, y_1) - H(x_1, y_2) + H(x_1, y_1)$ and $B(x_1, x_2, y_1, y_2) = H(x_1, y_1) - H(x_2, y_2) - H(x_1, y_2) + H(x_2, y_1)$.

Theorem 2.1. *The following holds:*

(1) *For $\alpha \geq 0$, if $A(x_1, x_2, y_1, y_2) \geq 0$ and $B(x_1, x_2, y_1, y_2) \geq 0$, then copula (10) satisfies the TP_2 property;*
(2) *For $\alpha \leq 0$, if $A(x_1, x_2, y_1, y_2) \leq 0$ and $B(x_1, x_2, y_1, y_2) \geq 0$, then copula (10) satisfies the TP_2 property.*

Proof. The assertion of the theorem is a direct consequence of the equalities

$$\begin{aligned} & f(x_1,y_1)f(x_2,y_2) - f(x_2,y_1)f(x_1,y_2) \\ = \; & \left[1 + \alpha H(x_1,y_1)\right]\left[1 + \alpha H(x_2,y_2)\right] - \left[1 + \alpha H(x_1,y_2)\right]\left[1 + \alpha H(x_2,y_1)\right] \\ = \; & \alpha A(x_1,x_2,y_1,y_2) + \alpha^2 B(x_1,x_2,y_1,y_2). \end{aligned}$$

□

Example 2.1. *Consider the function*

$$H(x,y) = x^p + y^p - (p+1)x^p y^p - \frac{1}{p+1}, \; 0 \leq x, y \leq 1, \; p > 0. \tag{13}$$

It is easy to verify that $H(x, y)$ satisfies conditions (9). The corresponding bivariate density is of the form

$$f(x,y) = 1 + \alpha \left\{ x^p + y^p - (p+1)x^p y^p - \frac{1}{p+1} \right\}, \; 0 \leq x, y \leq 1, \; p > 0. \tag{14}$$

Lemma 2.1. *For $p \geq 1$ and for α satisfying*

$$-\frac{p+1}{p} \leq \alpha \leq \frac{p+1}{2(p+1) - (p+1)^2 - 1},$$

and for $0 < p < 1$ and for α satisfying

$$-\frac{p+1}{p} \leq \alpha \leq p+1 \; ,$$

(13) represents a joint probability density of two uniform random variables.

Proof. We will use the inequalities in (11). It can be verified that

$$\frac{\partial H(x,y)}{\partial x} = px^{p-1} - (p+1)px^{p-1}y^p, \ \frac{\partial H(x,y)}{\partial y} = py^{p-1} - (p+1)px^p y^{p-1},$$

$$H_{xx} = \frac{\partial^2 H(x,y)}{\partial x^2} = p(p-1)x^{p-2} - (p+1)p(p-1)x^{p-2}y^p,$$

$$H_{xy} = \frac{\partial^2 H(x,y)}{\partial x \partial y} = -(p+1)p^2 x^{p-1} y^{p-1},$$

$$H_{yy} = \frac{\partial^2 H(x,y)}{\partial y^2} = p(p-1)y^{p-2} - (p+1)p(p-1)x^p y^{p-2}.$$

From the equations $\frac{\partial H(x,y)}{\partial x} = 0$ and $\frac{\partial H(x,y)}{\partial y} = 0$, we find the critical points of $H(x,y)$ to be $x_* = \frac{1}{(p+1)^{1/p}}$ and $y_* = \frac{1}{(p+1)^{1/p}}$. Since

$$H_{xx}(x_*,y_*)H_{yy}(x_*,y_*) - [H_{xy}(x_*,y_*)]^2 = -\frac{p^4}{(p+1)^{2(p-2)/p}} < 0,$$

(x_*,y_*) is a saddle point of $H(x,y)$. It is also true that $f(x_*,y) = f(x,y_*) = f(x_*,y_*) = 1$ for every $0 \le x, y \le 1$. The points $(0,1)$ and $(1,0)$ are local maxima of $H(x,y)$ in D^+ and

$$H(0,1) = H(1,0) = \frac{p}{p+1}. \tag{15}$$

The local minima of $H(x,y)$ in D^- are $(0,0)$ and $(1,1)$. For $p \ge 1$,

$$H(1,1) = 2 - \frac{(p+1)^2+1}{(p+1)} < -\frac{1}{p+1} = H(0,0). \tag{16a}$$

For $0 < p < 1$,

$$H(1,1) = 2 - \frac{(p+1)^2+1}{(p+1)} > -\frac{1}{p+1} = H(0,0). \tag{17}$$

The proof is completed by using (11), (15), (16a), and (17). □

The corresponding copula can be written as follows:

$$C_1(x,y) = xy + \frac{\alpha}{p+1}\left\{x^{p+1}y - x^{p+1}y^{p+1} - xy + xy^{p+1}\right\}, \ 0 \le x, y \le 1. \tag{18}$$

From (12), we then have

$$\rho = 12\alpha \int_0^1\!\!\int_0^1 xy\left(x^p + y^p - (p+1)x^p y^p - \frac{1}{p+1}\right)dxdy = -\frac{3\alpha p^2}{(p+1)(p+2)^2}.$$

For $p \ge 1$, the range for the correlation coefficient is

$$-\frac{3}{(p+2)^2} \le \rho \le \frac{3p}{(p+2)^2};$$

for $0 < p < 1$, the range for the correlation coefficient is

$$-\frac{p^2}{(p+2)^2} \le \rho \le \frac{3p}{(p+2)^2}.$$

The maximum positive correlation is $3/8 = 0.375$ which is achieved at the point $p = 2$. The maximum negative correlation is $-1/3$ and it is achieved at the point $p = 1$.

Corollary 2.1. *For $\alpha < 0$, copula (18) has the TP_2 property.*

Proof. For copula (18),

$$A(x_1, x_2, y_1, y_2) = (p+1)(y_2^p - y_1^p)(x_1^p - x_2^p)$$

and

$$\begin{aligned} B(x_1, x_2, y_1, y_2) &= \left(x_1^p + y_1^p - (p+1)x_1^p y_1^p - \frac{1}{p+1}\right)\left(x_2^p + y_2^p - (p+1)x_2^p y_2^p - \frac{1}{p+1}\right) \\ &\quad - \left(x_1^p + y_2^p - (p+1)x_1^p y_2^p - \frac{1}{p+1}\right)\left(x_2^p + y_1^p - (p+1)x_2^p y_1^p - \frac{1}{p+1}\right) \\ &= 0. \end{aligned}$$

And since $(y_2^p - y_1^p)(x_1^p - x_2^p) \leq 0$, we have $A(x_1, x_2, y_1, y_2) \leq 0$. Therefore, it follows from Theorem 1 that if $\alpha < 0$, (18) has the TP_2 property. □

Remark 2.1. *One may consider the kernel*

$$H_1(x, y) = -H(x, y).$$

Then the admissible range for the association parameter α *is*

$$\frac{-(p+1)}{2(p+1) - (p+1)^2 - 1} \leq \alpha \leq \frac{p+1}{p} \text{ for } p \geq 1,$$

and

$$-(p+1) \leq \alpha \leq \frac{p+1}{p} \text{ for } 0 < p < 1.$$

The corresponding copula is

$$C_2(x, y) = xy - \frac{\alpha}{p+1}\left\{x^{p+1}y - x^{p+1}y^{p+1} - xy + xy^{p+1}\right\},\ 0 \leq x, y \leq 1. \tag{19}$$

The range for the correlation coefficient is

$$-\frac{3p}{(p+2)^2} \leq \rho \leq \frac{3}{(p+2)^2} \text{ for } p \geq 1$$

and

$$-\frac{3p}{(p+2)^2} \leq \rho \leq \frac{p^2}{(p+2)^2} \text{ for } 0 < p < 1.$$

The maximum negative correlation is -0.375 *and the maximum positive correlation is* $1/3$.

Corollary 2.2. *For* $\alpha > 0$, *copula (19) has the* TP_2 *property.*

Example 2.2. *Consider the kernel*

$$H_2(x, y) = px^{p-1} + py^{p-1} - p^2x^{p-1}y^{p-1} - 1,\ 0 \leq x, y \leq 1.$$

The corresponding bivariate pdf is

$$f_2(x, y) = 1 + \alpha\left\{px^{p-1} + py^{p-1} - p^2x^{p-1}y^{p-1} - 1\right\}, 0 \leq x, y \leq 1.$$

A similar analysis shows that the admissible range for α *is*

$$-\frac{1}{p-1} \leq \alpha \leq 1 \text{ for } 1 < p \leq 2, \tag{20}$$

and

$$-\frac{1}{p-1} \leq \alpha \leq \frac{1}{(p-1)^2} \text{ for } p > 2. \tag{21}$$

The corresponding copula is

$$C_3(x,y) = xy\left\{1+\alpha\left[x^{p-1}+y^{p-1}-x^{p-1}y^{p-1}-1\right]\right\},\ 0\le x,y\le 1. \tag{22}$$

The range for the correlation coefficient is

$$-\frac{3(p-1)^2}{(p+1)^2}\le\rho\le\frac{3(p-1)}{(p+1)^2}\ \textit{for}\ 1<p\le 2 \tag{23}$$

and

$$-\frac{3}{(p+1)^2}\le\rho\le\frac{3(p-1)}{(p+1)^2}\ \textit{for}\ p>2. \tag{24}$$

Remark 2.2. *As before, one may consider the kernel $H_3(x,y)=-H_2(x,y)$. The copula has the form*

$$C_4(x,y) = xy\left\{1-\alpha\left[x^{p-1}+y^{p-1}-x^{p-1}y^{p-1}-1\right]\right\},\ 0\le x,y\le 1. \tag{25}$$

Then the inequalities (20) and (21) for α and the inequalities (23) and (24) for the correlation coefficient will be reversed.

Corollary 2.3. *For $\alpha<0$, copula (22) has the TP_2 property.*

Corollary 2.4. *For $\alpha>0$, copula (25) has the TP_2 property.*

The proofs are similar to the proof of Corollary 2.1.

Example 2.3. *Now consider the following kernel which is different from those considered previously:*

$$H_4(x,y)=\min(x,y)-x-y+\frac{3}{2}x^2y^2+\frac{1}{2},\ 0\le x,y\le 1. \tag{26}$$

It can be verified that the function (26) satisfies conditions (9). It is evident that

$$\max_{(x,y)\in D^-}\left[-H_4(x,y)\right]=-\min_{(x,y)\in D^-}H_4(x,y)=-H_4(1,0)=-H_4(0,1)=\frac{1}{2}, \tag{27}$$

$$\max_{(x,y)\in D^+}H_4(x,y)=H_4(1,1)=1. \tag{28}$$

The inequalities (11) along with equations (27) and (28) give the admissible range for the association parameter α as

$$-1\le\alpha\le 2.$$

From formula (12), the correlation coefficient is given by the quantity

$$\rho=12\alpha\int_0^1\int_0^1 xy\left\{\min(x,y)-x-y+\frac{3}{2}x^2y^2+\frac{1}{2}\right\}\ dx\ dy=\frac{9}{40}\alpha$$

and the range for the correlation coefficient is

$$-\frac{9}{40}\le\rho\le\frac{9}{20}$$

(or $-0.225 \le \rho \le 0.45$). The corresponding copula has the form

$$C_5(x,y) = xy + \alpha\left\{\frac{xy}{2} + \frac{x^3y^3}{6} - \frac{xy}{2}\max(x,y) - \frac{\{\min(x,y)\}^3}{6}\right\},$$

$$= \begin{cases} xy + \alpha\left\{\frac{xy}{2} + \frac{x^3y^3}{6} - \frac{xy^2}{2} - \frac{x^3}{6}\right\}, & x < y \\ xy + \alpha\left\{\frac{xy}{2} + \frac{x^3y^3}{6} - \frac{x^2y}{2} - \frac{y^3}{6}\right\}, & x \ge y. \end{cases}$$

Remark 2.3. *Just as before, one may consider the kernel $H_5(x,y) = -H_4(x,y)$. Then the admissible range for the association parameter α is $-2 \le \alpha \le 1$ and the range for the correlation coefficient is*

$$-\frac{9}{20} \le \rho \le \frac{9}{40}.$$

The simple analytical form of the presented copulas is sufficiently robust to allow applications in many practical problems. But even more flexibility may be added by introducing additional parameters to model (26).

3. Multivariate extension

One can extend the generalized Sarmanov family of copulas to the multivariate case. Let $H_k(x_1, x_2, ..., x_k), k = 2, 3, ..., n,$ be a family of symmetric functions which have the following recurrence properties:

$$\int_0^1 H_k(x_{i_1}, x_{i_2}, ..., x_{i_{k-1}}, x_{i_k})\, x_{i_k} = H_{k-1}(x_{i_1}, x_{i_2}, ..., x_{i_{k-1}}), \tag{29}$$

and

$$\int_0^1\int_0^1 \overset{(n-1)}{\cdots} \int_0^1 H_n\left(x_{i_1}, x_{i_2}, ..., x_{i_n}\right)\, dx_{i_1} dx_{i_2} \cdots dx_{i_{n-1}} = 0, \tag{30}$$

for any $k = 2, 3, ..., n$; $i_1, i_2, ..., i_k \in \{1, 2, ..., k\}$ and $0 \le x_{i_1}, x_{i_2}, ..., x_{i_{k-1}} \le 1$. Define now the following $n-$variate pdf with marginal distribution functions $F_i(x_i)$ and marginal pdfs $f_i(x_i)$:

$$\begin{aligned} f_\alpha(x_{i_1}, x_{i_2}, ..., x_{i_n}) &= f_{i_1}(x_{i_1}) f_{i_2}(x_{i_2}) ... f_{i_n}(x_{i_n}) \\ &\times \left\{1 + \alpha_n H_n\left(F_{i_1}(x_{i_2}), F_{i_2}(x_{i_2}), ..., F_{i_n}(x_{i_n})\right)\right\}, \end{aligned} \tag{31}$$

where the association parameter α_n satisfies the condition $1 + \alpha_n H_n\left(x_1, x_2, ..., x_n\right) \ge 0$ for any $0 \le x_1, x_2, ..., x_n \le 1$.

Denote

$$D_n^+ = \{(x_1, x_2, ..., x_n) : H_n(x_1, x_2, ..., x_n) \ge 0\}$$

and

$$D_n^- = \{(x_1,\ x_2, ..., x_n) : H_n(x_1, x_2, ..., x_n) \le 0\}.$$

Then, as in the bivariate case, the function given in (31) is a pdf of $n-$variate random vector $(X_1, X_2, ..., X_n)$ if

$$-\frac{1}{\max\limits_{(x_1, x_2, ..., x_n) \in D_n^+} H_n(x_1, x_2, ..., x_n)} \le \alpha_n \le \frac{1}{\max\limits_{(x_1, x_2, ..., x_n) \in D_n^-} \left(-H_n(x_1, x_2, ..., x_n)\right)}. \tag{32}$$

For uniform marginals, we have

$$c_\alpha(x_1, x_2, ..., x_n) = 1 + \alpha_n H_n(x_1, x_2, ..., x_n),\ 0 \le x_1, x_2, ..., x_n \le 1.$$

The corresponding $n-$variate copula is

$$C(x_1, x_2, ..., x_n) = x_1 x_2 ... x_n + \alpha_n \int_0^{x_n} \int_0^{x_{n-1}} \cdots \int_0^{x_1} H_n(t_1, t_2, ..., t_n)\, dt_1 dt_2 \cdots dt_n,$$

$$0 \le x_1, x_2, ..., x_n \le 1.$$

Consider for example a kernel

$$H_n(x_1, x_2, ..., x_n) = x_1^p + x_2^p + \cdots + x_n^p - (p+1)^{n-1} x_1^p x_2^p \cdots x_n^p - \frac{n-1}{p+1}. \tag{33}$$

It can be checked easily that the kernel satisfies conditions (29) and (30). Therefore, one can consider the $n-$ variate pdf

$$f_n(x_1, x_2, ..., x_n) = f_1(x_1) f_2(x_2) \cdots f_n(x_n)$$
$$\times \{1 + \alpha_n [F_1^p(x_1) + F_2^p(x_2) + \cdots + F_n^p(x_n)$$
$$-(p+1)^{n-1} (F_1(x_1) F_2(x_2) \cdots F_n(x_n))^p - \frac{n-1}{p+1}\Big]\Big\},$$
$$-\infty < x_1, x_2, ..., x_n < \infty.$$

For uniform $(0, 1)$ marginals, we have

$$c_\alpha(x_1, x_2, ..., x_n) = 1 + \alpha_n \left\{ x_1^p + x_2^p + \cdots + x_n^p - (p+1)^{n-1} x_1^p x_2^p \cdots x_n^p - \frac{n-1}{p+1} \right\}$$

and the $n-$ variate copula is

$$C_\alpha(x_1, x_2, ..., x_n) = x_1 x_2 \cdots x_n \left\{ 1 + \frac{\alpha_n}{p+1} \left[\sum_{i=1}^n x_i^p - \prod_{i=1}^n x_i^p - (n-1) \right] \right\}. \tag{34}$$

Theorem 3.1. *The admissible range for α_n allowing (34) to be $n-$variate copula is*

$$-\frac{p+1}{(n-1)p} \le \alpha_n \le \frac{p+1}{(p+1)^n - n(p+1) + (n-1)} \quad \text{for } p \ge n^{\frac{1}{n-1}} - 1,$$

and

$$-\frac{p+1}{(n-1)p} \le \alpha_n \le \frac{p+1}{n-1} \quad \text{for } 0 < p < n^{\frac{1}{n-1}} - 1.$$

Proof. We will use (32). It can be seen that in D_n^- the function (33) takes its minimum value at the point $x_1 = x_2 = ... = x_n = 1$, *i.e.*,

$$H_n(1, 1, ..., 1) = n - (p+1)^{n-1} - \frac{n-1}{p+1}. \tag{35}$$

In fact, if at least one of the coordinates is equal to 0 and the others are equal to 1, then the second term in (33) is equal to 0. In this case the minimum value is obtained when one of the x_i 's is equal to 1, and all others are equal to 0, *i.e.*,

$$H_n(0, 0, ..., 0, 1) = 1 - \frac{n-1}{p+1}. \tag{36}$$

Compare (35) with (36). If $n-(p+1)^{n-1} < 1$, *i.e.* $p > (n-1)^{\frac{1}{n-1}} - 1$, then $H_n(1,1,...,1) < H_n(0,0,...,0,1)$. Now when all of x_i 's are equal to zero, then

$$H_n(0,0,....,0) = -\frac{n-1}{p+1}. \tag{37}$$

Compare (37) with (35). If $n-(p+1)^{n-1} < 0$, *i.e.* $p > n^{\frac{1}{n-1}} - 1$, then $H_n(1,1,...,1) < H_n(0,0,...,0)$. Therefore, if

$$p > n^{\frac{1}{n-1}} - 1,$$

then

$$\alpha_n \leq \frac{p+1}{(p+1)^n - n(p+1) + (n-1)}.$$

Now in D_n^+ the function (33) has its maximum when one of the x_i 's is equal to 0 and all others are equal to 1. This value is $H(1,1,...,1,0) = \frac{(n-1)p}{p+1}$. Therefore, if $p > n^{\frac{1}{n-1}} - 1$, then

$$-\frac{p+1}{(n-1)p} \leq \alpha_n \leq \frac{p+1}{(p+1)^n - n(p+1) + (n-1)}.$$

□

References

[1] Barlow R. and Proschan, F., (1981), Statistical Theory of Reliability and Life Testing: Probability Models, to begin with, Silver Spring, MD.

[2] Bairamov, I.G., Kotz, S. and Bekçi, M., (2001), New generalized Farlie-Gumbel-Morgenstern distributions and concomitants of order statistics, Journal of Applied Statistics, 28(5), 521-536.

[3] Bairamov, I.G. and Kotz, S. (2002) Dependence structure and symmetry of Huang-Kotz FGM distributions and their extensions. Metrika. 56 ,1, 55-72.

[4] Bairamov, I. G., Kotz, S. and Gebizlioglu O.L. (2001) The Sarmanov family and its generalization. South African Statistical Journal. 35, 205-224.

[5] Huang, J.S. and Kotz, S., (1999), Modifications of the Farlie-Gumbel-Morgenstern distributions, A tough hill to climb, Metrika, 49, 135-145.

[6] Joe H., (1997), Multivariate Models and Dependence Concepts, Chapman and Hall, London.

[7] Lee, M.-L. T., (1996), Properties and applications of the Sarmanov family of bivariate distributions, Communications in Statistics. -Theory Meth., 25(6), 1207-1222.

[8] Lin, G.D. and Huang, J.S., (2011), Maximum correlation for the generalized Sarmanov bivariate distributions, Journal of Statistical Planning and Inference, 141, 2738-2749.

[9] Nelsen, R.B., (1998), An Introduction to Copulas, Springer-Verlag, New York.

[10] Sarmanov, O.V., (1966), Generalized normal correlation and two-dimensional Frechet classes, Doklady (Sovyet Mathematics), Tom 168, 596-599.

[11] Shaked, M., Shanthikumar, J.G., (1994), Stochastic Orders and Their Applications, Academic Press. Boston.

[12] Yu, S., Ouarda, T.B.M.J. and Bobée, B., (2001), A rewiew of bivariate gamma distributions for hydrological application, Journal of Hydrology, 246, 1-18.

Ismihan Bairamov is presently Professor of Mathematics and Statistics and Dean of Faculty of Arts and Sciences, Izmir University of Economics, Turkey. He graduated from Faculty of Applied Mathematics of Azerbaijan State University and received his PhD degree in probability and statistics from Kiev University, Ukraine in 1988. He worked in Institute of Cybernetics of Academy of Sciences of Azerbaijan, Azerbaijan State University and Azerbaijan Aerospace Agency as a researcher and lecturer. Between 1993 and 2001 he worked as a professor teaching probability and statistic courses in Ankara University and from 2001 he teaches in department of Mathematics of Izmir University of Economics. His research interests focus mainly in theory of statistics, probability, theory order statistics, nonparametric statistics and multivariate distributions and copulas.

Banu (Yağcı) Altınsoy is presently Statistician at Ministry of Transport and Communication, Department of Strategy and Development. She graduated from Faculty of Science, Department of Statistics, Ankara University and received her PhD degree in Statistics from Ankara University, Turkey in 2009. She worked in Department of Statistics, Faculty of Science, Ankara University as assistant between 2002 and 2009. In June, 2009, she started to work at Ministry of Transport and Communication, Republic of Turkey. Her research interests focus mainly in theory order statistics, nonparametric statistics, copulas, applied statistics and environmental statistics. She is the member of the environmental studies at the ministry and also preparing annual Turkey National Inventory Report of UNFCCC (United Nations Framework Convention on Climate Change).

G. Jay Kerns is presently Associate Professor at Department of Mathematics and Statistics Youngstown University, USA. He has B.A.Ed. Degree in Mathematics from Glenville State College (1999), M.A. and Ph.D. degrees in Statistics from Bowling Green State University (2000, 2004). His research interests include Statistical Computing, Exchangeability in Statistics, Exchangeable Measures and the Primary Teaching Areas are Statistics and Probability. He has published research papers dealing with De Finetti's Theorem for Abstract Finite Exchangeable Sequences, Calibration Experiments in Engineering Quality Control, Applications of Statistics in Medical Systems and Ecology.