Sayed Mohsen Mirhosseini; Mohammad Amini; Ali Dolati

On a general structure of the bivariate FGM type distributions

Applications of Mathematics, Vol. 60 (2015), No. 1, 91–108

Persistent URL: http://dml.cz/dmlcz/144095

ON A GENERAL STRUCTURE OF THE BIVARIATE FGM TYPE DISTRIBUTIONS

Sayed Mohsen Mirhosseini, Mohammad Amini, Mashhad,
Ali Dolati, Yazd

(Received May 5, 2013)

Abstract. In this paper, we study a general structure for the so-called Farlie-Gumbel-Morgenstern (FGM) family of bivariate distributions. Through examples we show how to use the proposed structure to study dependence properties of the FGM type distributions by a general approach.

Keywords: copula; dependence; FGM family; measure of association

MSC 2010: 62E15, 62H10

1. Introduction

The construction of multivariate distributions with given marginals has been a problem of interest to statisticians for many years. For any pair of univariate cumulative distribution functions (cdf) F_1 and F_2, one of the most popular bivariate distributions is the so-called (bivariate) Farlie-Gumbel-Morgenstern (FGM) class of distributions defined by

$$H(x,y) = F_1(x)F_2(y)\{1+\alpha\overline{F}_1(x)\overline{F}_2(y)\}, \quad -1 \leqslant \alpha \leqslant 1, \tag{1.1}$$

originally considered by Farlie [10], Gumbel [12] and Morgenstern [18]. A well-known limitation to this family is that it does not allow the modeling of large dependencies since, for example, Spearman's rho of this family is limited to $[-1/3, 1/3]$. For this reason and its simple analytical form, various generalizations of this family

We acknowledge the helpful comments of the reviewers that led to several improvements in this paper. This research was supported by a grant from Ferdowsi University of Mashhad (No. MS92298AMI).

have been introduced in the literature, e.g., [2], [4], [5], [3], [6], [9], [11], [14], [15], [16], [17], [20]. Lai and Xie [16] considered a representation of the FGM bivariate distribution possessing positive quadrant dependence property of the form

$$H(x,y) = F_1(x)F_2(y) + W(x,y) \quad \text{for all } x, y, \tag{1.2}$$

with nonnegative $W(x,y)$ satisfying certain regularity conditions ensuring that $H(x,y)$ is a bivariate distribution function. Recently, Han [13] studied the above representation of the FGM family possessing negative quadrant dependence property. In this paper we consider a generalization of the above model defined by

$$H(x,y) = F_1(x)F_2(y) + W\{F_1(x), F_2(y)\} \quad \text{for all } x, y, \tag{1.3}$$

to obtain a general extension for the FGM family of distributions. Our modelling approach starts on the path of Lai and Xie [16], but then rapidly diverges, because the marginal cdfs of our model are of the form

$$H_i(x) = F_i(x) + W\{F_i(x), 1\}, \quad i = 1, 2,$$

rather than F_1 and F_2. Through examples we show how to use the proposed structure to analyze different dependence properties of the already existent and new extensions of the FGM model. The paper is organized as follows. We discuss the general form of the proposed model in Section 2. In Section 3 we study different dependence properties of the general model. Several new extensions of the FGM family are given in Section 4. We conclude the paper in Section 5.

2. The proposed model

Let F_1 and F_2 be two continuous cdfs with the density functions f_1 and f_2 and suppose that $I = [0,1]$. Let $W\colon I^2 \to I$ and consider the function $H(x,y)$ defined by

$$H(x,y) = F_1(x)F_2(y) + W\{F_1(x), F_2(y)\}. \tag{2.1}$$

The following proposition shows under which conditions on W the function H defines a bivariate cdf.

Proposition 1. *The function H defined by* (2.1) *is a bivariate distribution function with the univariate marginal distribution functions*

$$H_1(x) = F_1(x) + W(F_1(x), 1) \quad \text{and} \quad H_2(x) = F_2(y) + W(1, F_2(y)), \tag{2.2}$$

provided that the kernel W satisfies

(i) $W(u, 0) = W(0, v) = W(1, 1) = 0$ *for all* $u, v \in I$;

(ii) $uv + W(u, v)$ *is 2-increasing on* I^2.

Proof. We recall that H is a bivariate cdf if and only if

(P1) $H(x, -\infty) = H(-\infty, y) = 0$, $H(x, \infty) = H_1(x)$, and $H(\infty, y) = H_2(y)$ for all x, y;

(P2) for every x_1, x_2, y_1, y_2 such that $x_1 \leqslant x_2$ and $y_1 \leqslant y_2$,

$$\Delta_{x_1}^{x_2} \Delta_{y_1}^{y_2} H(x, y) = H(x_2, y_2) - H(x_2, y_1) - H(x_1, y_2) + H(x_1, y_1) \geqslant 0.$$

It is clear that (i) ensures (P1). For condition (ii) we notice that (P2) holds if and only if

$$\begin{aligned}\Delta_{x_1}^{x_2} \Delta_{y_1}^{y_2} H(x, y) = {}& [F_1(x_2) - F_1(x_1)][F_2(y_2) - F_2(y_1)] + W(F_1(x_2), F_2(y_2)) \\ & - W(F_1(x_1), F_2(y_2)) - W(F_1(x_2), F_2(y_1)) \\ & + W(F_1(x_1), F_2(y_1)) \geqslant 0,\end{aligned}$$

or equivalently,

$$\begin{aligned}\Delta_{u_1}^{u_2} \Delta_{v_1}^{v_2} H(u, v) = {}& (u_2 - u_1)(v_2 - v_1) + W(u_2, v_2) - W(u_1, v_2) \\ & - W(u_2, v_1) + W(u_1, v_1) \geqslant 0,\end{aligned}$$

for all $u_1, u_2, v_1, v_2 \in I$ with $u_1 \leqslant u_2$ and $v_1 \leqslant v_2$, which is the 2-increasingness property of the function $uv + W(u, v)$, $u, v \in I$. Note that if $uv + W(u, v)$ is 2-increasing, then the functions $u + W(u, 1)$ and $v + W(1, v)$ are increasing. This property together with the condition (i) ensures that H_1 and H_2 are univariate cdfs. □

If $W(u, v)$ is a twice-differentiable function defined on I^2, then the following result shows under which conditions on W the function H defines an absolutely continuous bivariate cdf.

Proposition 2. *The function H defined by* (2.1) *is an absolutely continuous bivariate distribution function with the univariate marginal cdfs H_1 and H_2 if the kernel W satisfies*

(i) $W(u,0) = W(0,v) = 0$, $W(1,1) = 0$;

(ii) $W^{12}(u,v) \geqslant -1$, $W^1(u,1) \geqslant -1$, $W^2(1,v) \geqslant -1$;

(iii) $\int_0^1 W^1(u,1)\,\mathrm{d}u = 0$, $\int_0^1 W^2(1,v)\,\mathrm{d}v = 0$, $\int_0^1\int_0^1 W^{12}(u,v)\,\mathrm{d}u\,\mathrm{d}v = 0$, *where*

$$W^{12}(u,v) = \frac{\partial^2}{\partial u \partial v} W(u,v), \quad W^1(u,1) = \frac{\partial}{\partial u} W(u,1), \quad \text{and } W^2(1,v) = \frac{\partial}{\partial v} W(1,v).$$

P r o o f. Under conditions (i) and (ii) the functions h, h_1, and h_2 given by

$$h(x,y) = \frac{\partial^2}{\partial x \partial y} H(x,y) = f_1(x) f_2(y)\{1 + W^{12}(F_1(x), F_2(y))\}, \tag{2.3}$$

$$\begin{aligned} h_1(x) &= \int_{-\infty}^{\infty} h(x,y)\,\mathrm{d}y \\ &= f_1(x) \int_{-\infty}^{\infty} f_2(y)\{1 + W^{12}(F_1(x), F_2(y))\}\,\mathrm{d}y \\ &= f_1(x) \int_0^1 \{1 + W^{12}(F_1(x), v)\}\,\mathrm{d}v \\ &= f_1(x)\{1 + W^1(F_1(x), 1)\} \\ &= \frac{\partial}{\partial x} H_1(x), \end{aligned} \tag{2.4}$$

and similarly

$$h_2(y) = \int_{-\infty}^{\infty} h(x,y)\,\mathrm{d}x = f_2(y)\{1 + W^2(1, F_2(y))\} = \frac{\partial}{\partial y} H_2(y), \tag{2.5}$$

satisfy $h(x,y) \geqslant 0$, $h_1(x) \geqslant 0$, $h_2(y) \geqslant 0$, for all x, y and under condition (iii) we have

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{-\infty}^{\infty} h_1(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} h_2(y)\,\mathrm{d}y = 1;$$

that is, the function H defined by (2.1) is an absolutely continuous bivariate cdf with the joint density function (2.3) and univariate marginal density functions (2.4) and (2.5). □

R e m a r k 1. For the special case $W(u,1) = W(1,u) = 0$, for all $u \in I$, the marginal distribution functions of (2.1) are F_1 and F_2. Distributions of this form are well studied in [11], [15], [20].

R e m a r k 2. Note that every convex combination of two bivariate distribution functions of the form (2.1), with the kernels satisfying conditions of Proposition 1, is still a distribution function of the form (2.1). Let

$$H_1(x,y) = F_1(x)F_2(y) + W_1\{F_1(x), F_2(y)\},$$

and

$$H_2(x,y) = F_1(x)F_2(y) + W_2\{F_1(x), F_2(y)\},$$

where W_1 and W_2 satisfy conditions of Proposition 1. For every $\lambda \in I$, let

$$H(x,y) = \lambda H_1(x,y) + (1-\lambda)H_2(x,y).$$

Then

$$H(x,y) = F_1(x)F_2(y) + W\{F_1(x), F_2(y)\},$$

where $W(u,v) = \lambda W_1(u,v) + (1-\lambda)W_2(u,v)$, $u, v \in I$.

In view of Sklar's Theorem [21] from (2.1) and (2.2), solving the equation

$$C\{H_1(x), H_2(y)\} = H(x,y)$$

for the function $C\colon I^2 \to I$ yields the underlying copula associated with the bivariate cdf H defined by (2.1) as

$$C(u,v) = \psi_1^{-1}(u)\psi_2^{-1}(v) + W\{\psi_1^{-1}(u), \psi_2^{-1}(v)\}, \tag{2.6}$$

for all $u, v \in I$, where $\psi_i\colon I \to I$, $i = 1,2$, are two (distribution) functions given by $\psi_1(t) = t + W(t,1)$ and $\psi_2(t) = t + W(1,t)$. For $0 < p < 1$, the function $\psi_i^{-1}(p) = \inf\{t\colon \psi_i(t) \geqslant p\}$, denotes the inverse (quantile function) of ψ_i, $i = 1,2$. Note that the equation (2.6) could be written as

$$C(\psi_1(u), \psi_2(v)) = uv + W(u,v). \tag{2.7}$$

Let (X_1, X_2) be a pair of random variables distributed as (2.1). If $F_1(x) = F_2(x)$ for all x, then (X_1, X_2) is exchangeable, that is, (X_1, X_2) and (X_2, X_1) have the same joint cdf, if and only if $W(u,v) = W(v,u)$ for all $u, v \in I$.

The joint survival function associated with (2.1) is given by

$$\begin{aligned} \overline{H}(x,y) &= P[X > x, Y > y] \\ &= 1 - H_1(x) - H_2(y) + H(x,y) \\ &= \overline{F_1}(x)\overline{F_2}(y) - W(F_1(x),1) - W(1,F_2(y)) + W(F_1(x),F_2(y)). \end{aligned} \tag{2.8}$$

Recall that a pair (X_1, X_2) with the joint cdf H and marginal cdfs H_1 and H_2 is radially symmetric (see [19]) if the pairs (X_1, X_2) and $(-X_1, -X_2)$ have the same joint cdf or equivalently $H(x, y) = \overline{H}(-x, -y)$ for all x, y. In such a case the marginal cdfs satisfy $H_1(x) = \overline{H}_1(-x)$ and $H_2(y) = \overline{H}_2(-y)$ for all x, y. For the joint cdf (2.1) we have the following result.

Proposition 3. *Let H be the joint distribution function defined by* (2.1) *with $F_1(x) = \overline{F}_1(-x)$ and $F_2(y) = \overline{F}_2(-y)$ for all x, y. Then H is radially symmetric if and only if for all $u, v \in I$,*

$$W(u, v) = W(1 - u, 1 - v) - W(1 - u, 1) - W(1, 1 - v). \tag{2.9}$$

Proof. The result follows from (2.8) and the assumption $F_i(-x) = 1 - F_i(x)$, $i = 1, 2$. □

3. Dependence properties

A useful way of formalizing properties of a bivariate distribution is to examine the form of its stochastic dependence. In this section we study some results on dependence properties of the bivariate distribution defined by (2.1). Four forms of stochastic dependence between two random variables X and Y are the quadrant dependence, tail monotonicity, regression dependence and the likelihood ratio dependence; see [19] for detail. Let X and Y be two random variables with the joint distribution function H, the joint density function h and the associated copula C. Recall that X an Y are said to be positively quadrant dependent (PQD) if and only if $C(u, v) \geqslant uv$ for all $u, v \in I$. Similarly, X and Y are said to be negatively quadrant dependent (NQD) if and only if $C(u, v) \leqslant uv$ for all $u, v \in I$. The random variable Y is said to be left tail decreasing in X (briefly, LTD($Y|X$)) if and only if $C(u, v)/u$ is decreasing in u, for all v. The random variable Y is right tail increasing in X (briefly, RTI($Y|X$)) if and only if $(v - C(u, v))/(1 - u)$ is increasing in u for all v. Y is said to be positive regression dependent in X (briefly, PRD($Y|X$)) if and only if $C(u, v)$ is concave in u for every v. The random vector (X, Y) is said to be positive likelihood ratio dependent (PLRD) if $h(x, y)$ is total positive of order two (TP2), that is,

$$h(x_1, y_1)h(x_2, y_2) \geqslant h(x_1, y_2)h(x_2, y_1), \tag{3.1}$$

for all $x_1 \leqslant x_2$ and $y_1 \leqslant y_2$. The negative likelihood ratio dependence (NLRD) is defined by reversing the inequality (3.1).

The following propositions provide some results for the bivariate distribution defined by (2.1).

Proposition 4. *Let* (X_1, X_2) *be a vector of continuous random variables distributed as* (2.1). *Then* (X_1, X_2) *is* PQD (NQD) *if for all* $u, v \in I$,

$$W(u,v) \geqslant uW(1,v) + vW(u,1) + W(u,1)W(1,v)$$
$$(W(u,v) \leqslant uW(1,v) + vW(u,1) + W(u,1)W(1,v)).$$

P r o o f. The proof follows from (2.6). □

Proposition 5. *Let* (X_1, X_2) *be a vector of continuous random variables distributed as* (2.1). *Then* (X_1, X_2) *is* PLRD *if for all* $u, v \in I$, *the function* $1+W^{12}(u,v)$ *is* TP2.

P r o o f. From (2.3), the inequality (3.1) holds if and only if

$$\begin{aligned}&[1+W^{12}(F_1(x_1), F_2(y_1))][1+W^{12}(F_1(x_2), F_2(y_2))] \\ &\qquad \geqslant [1+W^{12}(F_1(x_1), F_2(y_2))][1+W^{12}(F_1(x_2), F_2(y_1))],\end{aligned}$$

for all $x_1 \leqslant x_2$ and $y_1 \leqslant y_2$, or equivalently,

$$[1+W^{12}(u_1,v_1)][1+W^{12}(u_2,v_2)] \geqslant [1+W^{12}(u_1,v_2)][1+W^{12}(u_2,v_1)],$$

for all $0 \leqslant u_1 \leqslant u_2 \leqslant 1$ and $0 \leqslant v_1 \leqslant v_2 \leqslant 1$, which is the TP2 property of $1+W^{12}(u,v)$. □

R e m a r k 3. The following chain of implications is known; see, e.g., [19].

$$\text{PLRD} \Rightarrow \text{PRD} \Rightarrow \begin{matrix}\text{LTD}\\ \text{RTI}\end{matrix} \Rightarrow \text{PQD}. \tag{3.2}$$

Thus under the condition of Proposition 5, a random vector (X_1, X_2) distributed as (2.1) has PRD, LTD, RTI, and the weaker dependence property PQD.

The population versions of two of the most common nonparametric measures of association between the components of a continuous random pair (X_1, X_2) are *Kendall's tau* (τ) and *Spearman's rho* (ϱ) which depend only on the copula C of the pair (X_1, X_2), and are given by

$$\begin{aligned}\tau(X_1,X_2) &= 4\int_0^1\int_0^1 C(u,v)\,\mathrm{d}C(u,v) - 1 \\ &= 1 - 4\int_0^1\int_0^1 \frac{\partial}{\partial u}C(u,v)\frac{\partial}{\partial v}C(u,v)\,\mathrm{d}u\,\mathrm{d}v,\end{aligned} \tag{3.3}$$

and

$$\varrho(X_1, X_2) = 12 \int_0^1 \int_0^1 C(u,v)\,\mathrm{d}u\,\mathrm{d}v - 3. \tag{3.4}$$

See [19] for detail. The following result provides expressions for these measures associated with a vector (X_1, X_2) distributed as (2.1).

Proposition 6. *Let* (X_1, X_2) *be a random vector distributed as* (2.1). *Then*

$$\tau(X_1, X_2) = -4 \int_0^1 \int_0^1 [vW^2(u,v) + uW^1(u,v) + W^1(u,v)W^2(u,v)]\,\mathrm{d}u\,\mathrm{d}v \tag{3.5}$$

and

$$\varrho(X_1, X_2) = 12 \int_0^1 \int_0^1 [uv + W(u,v))(1 + W^1(u,1))(1 + W^2(1,v))]\,\mathrm{d}u\,\mathrm{d}v - 3. \tag{3.6}$$

P r o o f. Applying (2.6), the proof is a straightforward calculation using (3.3) and (3.4). □

The stress-strength parameter (i.e. $R = P(X_1 < X_2)$) is useful for data analysis purposes [7]. The following proposition gives a convenient form for the stress-strength parameter of the proposed model. Let X_1 and X_2 be independent and identically distributed continuous random variables. Then $P(X_1 < X_2) = \frac{1}{2}$. For a random vector (X_1, X_2) distributed as (2.1) we have the following result.

Proposition 7. *Let* (X_1, X_2) *be a vector of continuous random variables distributed as* (2.1). *Then*

$$P(X_2 \leqslant X_1) = E[F_2 \circ F_1^{-1}(U) + W^1(U, F_2 \circ F_1^{-1}(U))], \tag{3.7}$$

where U *is a uniform* $(0,1)$ *random variable and* $F_i^{-1}(u) = sup\{x | F_i(x) \leqslant u\}$ *for* $i = 1, 2$.

P r o o f. By using (2.2) and the conditional density function of $(X_2 | X_1 = x)$

$$h(y|x) = \frac{1 + W^{12}(F_1(x), F_2(y))}{1 + W^1(F_1(x), 1)} f_2(y), \tag{3.8}$$

we obtain

$$\begin{aligned}P(X_2 \leqslant X_1) &= \int_{-\infty}^{\infty} P(X_2 \leqslant t|X_1 = t]\,\mathrm{d}H_1(t) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{t} h(y|t)\,\mathrm{d}y\,\mathrm{d}H_1(t) \\ &= \int_{-\infty}^{\infty} \frac{F_2(t) + W^1(F_1(t), F_2(t))}{1 + W^1(F_1(t), 1)}\,\mathrm{d}H_1(t).\end{aligned}$$

Now the transformation $F_1(t) = u$ gives the required result. □

Note that under the assumption $F_1(x) = F_2(x)$ for all x we have

$$P(X_2 \leqslant X_1) = \frac{1}{2} + \int_0^1 W^1(u, u)\,\mathrm{d}u,$$

where U is a uniform (0,1) random variable.

In the following we provide an expression for the expected value of the left truncated model; i.e., $(X_1|X_1 \geqslant X_2)$ of a random vector (X_1, X_2) distributed as (2.1).

Proposition 8. *Let (X_1, X_2) be a vector of continuous random variables distributed as* (2.1) *with $F_1(x) = F_2(x) = x$, $0 < x < 1$. Then*

$$E(X_1|X_1 \geqslant X_2) = \frac{2}{3} - 2\int_0^1 W(x, x)\,\mathrm{d}x. \tag{3.9}$$

Proof. First we note that the conditional distribution of X_1 given $X_2 \leqslant X_1$ is

$$\begin{aligned}P(X_1 \leqslant x|X_2 \leqslant X_1) &= \frac{P(X_1 \leqslant x, X_2 \leqslant X_1)}{P(X_2 \leqslant X_1)} = \frac{P(X_2 \leqslant X_1 \leqslant x)}{P(X_2 \leqslant X_1)} \\ &= \frac{1}{P(X_2 \leqslant X_1)}\int_0^x\int_0^u h(u, v)\,\mathrm{d}v\,\mathrm{d}u,\end{aligned}$$

where h is the joint density function of X_1 and X_2. The conditional density function of X_1 given $X_2 \leqslant X_1$ is then

$$h_{X_1|X_2 \leqslant X_1}(x) = \frac{1}{P(X_2 \leqslant X_1)}\int_0^x h(x, v)\,\mathrm{d}v,$$

and then

$$E(X_1|X_2 \leqslant X_1) = \frac{1}{P(X_2 \leqslant X_1)}\int_0^{\infty} x\left\{\int_0^x h(x, v)\,\mathrm{d}v\right\}\mathrm{d}x.$$

Using (2.3) and letting $F_1(x) = F_2(x)$, we have $P(X_2 \leqslant X_1) = \frac{1}{2}$ and hence

$$E(X_1|X_2 \leqslant X_1) = 2E\{F_1^{-1}(U)(U + W^1(U, U))\},$$

where U is the uniform $(0, 1)$ random variable. If $F_1(x) = x, 0 < x < 1$, then

$$E(X_1|X_2 \leqslant X_1) = \frac{2}{3} + 2E\{UW^1(U, U)\} = \frac{2}{3} - 2\int_0^1 W(u, u)\,\mathrm{d}u,$$

which completes the proof. □

Note that for independent and identically distributed random variables X_1 and X_2 having uniform (0,1) distribution we have

$$E\{X_1|X_1 \geqslant X_2\} = \frac{2}{3}.$$

Thus for two random variables X_1 and X_2 distributed as

$$H(u, v) = uv + W(u, v),$$

the value

$$\kappa(X_1, X_2) = \left|E(X_1|X_1 \geqslant X_2) - \frac{2}{3}\right| = 2\left|\int_0^1 W(u, u)\,\mathrm{d}u\right| \tag{3.10}$$

provides a natural measure of dependence. This kind of dependence measures was proposed in [3] for FGM type distributions.

4. Subfamilies and several new examples

For $\theta \in [-1, 1]$, the kernel $W(u, v) = \theta uv(1 - u)(1 - v)$ generates the standard FGM family of distributions defined by (1.1). An extension of the FGM family is introduced in [20] with $W(u, v) = f(u)g(v)$, where f and g are two continuous functions on I with $f(0) = f(1) = g(0) = g(1) = 0$ and satisfying the Lipschitz condition, i.e.,

$$|f(u_2) - f(u_1)| \leqslant |u_2 - u_1| \quad \text{and} \quad |g(v_2) - g(v_1)| \leqslant |v_2 - v_1|,$$

for all $u_2 \leqslant u_1$ and $v_2 \leqslant v_1$ in I. This class of copulas generalizes several families; see, e.g., [1], [2], [5], [3], [15], [16]. Under the boundary conditions on f and g, the joint distribution constructed via

$$H(x, y) = F_1(x)F_2(y) + f(F_1(x))g(F_2(y))$$

has univariate marginals F_1 and F_2 and the kernel W satisfies $W(u,1) = W(1,v) = 0$.

In the following we introduce several new FGM type distributions of the form (2.1), where the condition $W(u,1) = W(1,v) = 0$ is cancelled and thus the marginal distributions of H are not F_1 and F_2. The kernel W is constructed based on the joint distributions of order statistics and record values of the uniform (0,1) random samples of size two.

Example 1. Let X_1, X_2, Y_1, and Y_2 be independent uniform (0,1) random variables. Let $X_{(1)}, X_{(2)}$ and $Y_{(1)}, Y_{(2)}$ be their corresponding order statistics. For $-1 \leqslant \theta \leqslant 1$, consider the random pair $(V_1, V_2) = (X_{(1)}, Y_{(1)})$ with the probability $\frac{1}{2}(1+\theta)$ and $(V_1, V_2) = (X_{(2)}, Y_{(2)})$ with the probability $\frac{1}{2}(1-\theta)$. Then the joint distribution function of (V_1, V_2) is given by

$$F_{V_1,V_2}(u,v) = uv + uv[\theta(2-u-v) + (1-u)(1-v)], \quad u,v \in I.$$

It is clear that the above family of distributions complies with (2.1) for

$$W(u,v) = uv[\theta(2-u-v) + (1-u)(1-v)]. \tag{4.1}$$

Thus for every pair of univariate continuous distribution functions F_1 and F_2 and each $\theta \in [-1,1]$, the function

$$H^{+}(x,y) = F_1(x)F_2(y)\{1 + \theta(\overline{F}_1(x) + \overline{F}_2(y)) + \overline{F}_1(x)\overline{F}_2(y)\} \tag{4.2}$$

is a bivariate distribution with the univariate marginal distribution functions

$$H_1(x) = F_1(x)\{1 + \theta\overline{F}_1(x)\} \quad \text{and} \quad H_2(y) = F_2(y)\{1 + \theta\overline{F}_2(y)\}.$$

Alternatively, putting $(V_1, V_2) = (X_{(1)}, Y_{(2)})$ with the probability $\frac{1}{2}(1+\theta)$ and $(V_1, V_2) = (X_{(2)}, Y_{(1)})$ with the probability $\frac{1}{2}(1-\theta)$, gives

$$F_{V_1,V_2}(u,v) = uv + uv[\theta(v-u) - (1-u)(1-v)], \quad u,v \in I,$$

with the associated kernel

$$W(u,v) = uv[\theta(v-u) - (1-u)(1-v)].$$

For every $\theta \in [-1,1]$, and fixed univariate distributions F_1 and F_2, the generated bivariate distribution is given by

$$H^{-}(x,y) = F_1(x)F_2(y) + F_1(x)F_2(y)\{\theta(\overline{F}_1(x) - \overline{F}_2(y)) - \overline{F}_1(x)\overline{F}_2(y)\}, \tag{4.3}$$

with the univariate marginal distributions

$$H_1(x) = F_1(x)\{1 + \theta\overline{F}_1(x)\} \quad \text{and} \quad H_2(y) = F_2(y)\{1 - \theta\overline{F}_2(y)\}.$$

The following result shows the dependence properties of the joint cdf's defined by (4.2) and (4.3).

Proposition 9. *A random vector* (X_1, X_2) *with the joint distribution function defined by* (4.2) *(or* (4.3)*) has* PLRD *(or* NLRD*) property.*

P r o o f. We only prove the PLRD property of (4.2), the proof of NLRD for (4.3) is similar. Note that for (4.2) we have

$$W^{12}(u, v) = 2\theta(1 - u - v) - 2(u - v + 2uv) + 1.$$

Using Proposition 4, we have

$$\begin{aligned}[1 + W^{12}(u_1, v_1)][1 + W^{12}(u_2, v_2)] - [1 + W^{12}(u_1, v_2)][1 + W^{12}(u_2, v_1)]\\ = 4\theta^2(u_2 - u_1)(v_2 - v_1) \geqslant 0,\end{aligned}$$

for all $u_1 \leqslant u_2$ and $v_1 \leqslant v_2$, which is the TP2 property of the density function of H^+. □

The following result provides the measures of association Kendall's tau and Spearman's rho for the joint distribution functions (4.2) and (4.3).

Proposition 10. *The Spearman's rho and Kendall's tau and the measure of dependence* (3.10) *associated with the families of distributions* H^+ *and* H^- *defined by* (4.2) *and* (4.3) *are given by*

$$\varrho(H^+) = \frac{1 - \theta^2}{3}, \quad \tau(H^+) = \frac{2(1 - \theta^2)}{9},$$
$$\varrho(H^-) = -\frac{1 - \theta^2}{3}, \quad \tau(H^-) = -\frac{2(1 - \theta^2)}{9},$$

and

$$\kappa(H^+) = \left|\frac{1 + 5\theta}{15}\right|, \quad \kappa(H^-) = \frac{1}{15}.$$

P r o o f. The proof follows from expressions (3.5), (3.6), and (3.10). □

R e m a r k 4. Note that for both families (4.2) and (4.3) we have the equality $2\varrho = 3\tau$.

R e m a r k 5. Since the family of distributions defined by (4.2) has PLRD property, it is suitable to describe the positive dependence of a random pair (X, Y). In contrast, the family of distributions defined by (4.3) with NLRD property is suitable for modeling a negative dependence. However, it is very simple to consider a distribution with both positive and negative dependence property using a convex combination of these models. For example, let $\alpha, \beta \in [0, 1]$, with $\alpha + \beta = 1$, and define

$$H^{\pm}(x, y) = \alpha H^{+}(x, y) + \beta H^{-}(x, y). \tag{4.4}$$

Then, $H^{\pm}$ is a bivariate distribution with the univariate marginal distributions

$$H_1^{\pm}(x) = F_1(x)\{1 + \theta \overline{F}_1(x)\}, \quad H_2^{\pm}(y) = F_2(y)\{1 + \theta(\alpha - \beta)\overline{F}_2(y)\}.$$

The following example provides some extensions of the FGM family based on the joint distribution of the record statistics.

E x a m p l e 2. Let X_1 and X_2 be independent random variables with the common distribution functions F. It is known that the distribution functions of the lower record statistics (X_{L_2}) and upper record statistics (X_{U_2}) are given by [8]

$$F_{X_{L_2}}(x) = F(x) - F(x)\ln(F(x)) \quad \text{and} \quad F_{X_{U_2}}(x) = F(x) + \overline{F}(x)\ln(\overline{F}(x)).$$

Let X_1, X_2, Y_2, and Y_2 be independent random variables from the univariate distributions F_1 (of X_i) and F_2 (of Y_i). For $\theta \in [0, 1]$, choose the random vector $(V_1, V_2) = (X_{L_2}, Y_{L_2})$ with the probability θ and $(V_1, V_2) = (X_1, Y_1)$ with the probability $1 - \theta$. Then,

$$F_{V_1,V_2}(u, v) = uv + \theta uv[\ln(u)\ln(v) - \ln(uv)], \quad u, v \in (0, 1),$$

and

$$W(u, v) = \theta uv[\ln(u)\ln(v) - \ln(uv)].$$

For every $\theta \in [0, 1]$ and univariate distribution functions F_1 and F_2, the corresponding generated distribution function is given by

$$G^{+}(x, y) = F_1(x)F_2(y)\{1 + \theta[\ln(F_1(x))\ln(F_2(y)) - \ln(F_1(x)F_2(y))]\}, \tag{4.5}$$

whose univariate marginal distribution functions are

$$G_1(x) = F_1(x)\{1 - \theta\ln(F_1(x))\} \quad \text{and} \quad G_2(y) = F_2(y)\{1 - \theta\ln(F_2(y))\}.$$

Alternatively, we can take $(V_1, V_2) = (X_{L_2}, Y_{U_2})$ with the probability θ and $(V_1, V_2) = (X_1, Y_1)$ with the probability $1-\theta$. Then, for $\theta \in [0,1]$,

$$F_{V_1,V_2}(u,v) = uv + \theta u[(1-v)(1-\ln(u))\ln(1-v) - v\ln(u)], \quad u,v \in (0,1).$$

Thus with the kernel

$$W(u,v) = \theta u[(1-v)(1-\ln(u))\ln(1-v) - v\ln(u)]$$

we have a joint distribution function of the form

$$G^-(x,y) = F_1(x)F_2(y) + \theta F_1(x)[\overline{F}_2(y)\ln(\overline{F}_2(y))(1-\ln(F_1(x))) - F_2(y)\ln(F_1(x))], \tag{4.6}$$

with univariate marginal distribution functions

$$G_1(x) = F_1(x)\{1-\theta\ln(F_1(x))\} \quad \text{and} \quad G_2(y) = F_2(y) + \theta\overline{F}_2(y)\ln(\overline{F}_2(y)).$$

Proposition 11. *The joint distribution functions defined by* (4.5) *and* (4.6) *have* PLRD *and* NLRD *property, respectively.*

Proposition 12. *The Kendall's tau, Spearman's rho and measure of dependence* (3.10) *associated with the families of distributions defined by* (4.5) *and* (4.6) *are given by*

$$\tau(G^+) = \frac{1}{2}\theta(1-\theta), \quad \varrho(G^+) = \frac{3}{4}\theta(1-\theta),$$
$$\tau(G^-) = -\frac{1}{2}\theta(1-\theta), \quad \varrho(G^-) = -\frac{3}{4}\theta(1-\theta),$$

and

$$\kappa(G^+) = \frac{16\theta}{27}, \quad \kappa(G^-) = \frac{(40-3\pi^2)\theta}{54}.$$

Remark 6. Note that for $\alpha, \beta \in [0,1]$, with $\alpha+\beta=1$, the convex combination

$$G^{\pm}(x,y) = \alpha G^+(x,y) + \beta G^-(x,y) \tag{4.7}$$

is a bivariate cdf which has both positive and negative dependence property.

E x a m p l e 3. Let (X_1, Y_1) and (X_2, Y_2) be two independent random vectors having a common bivariate cdf F and univariate marginal cdfs F_1 (of X_i) and F_2 (of Y_i), $i = 1, 2$. Let $X_{(1)}, X_{(2)}$ and $Y_{(1)}, Y_{(2)}$ be their corresponding order statistics. For $-1 \leqslant \lambda \leqslant 1$, consider the random pair $(V_1, V_2) = (X_{(1)}, Y_{(1)})$ with the probability $\frac{1}{2}(1+\lambda)$ and $(V_1, V_2) = (X_{(2)}, Y_{(2)})$ with the probability $\frac{1}{2}(1-\lambda)$. Then, it is straightforward to verify that (V_1, V_2) has the joint cdf

$$G_\lambda(x, y) = (1+\lambda)(F_1(x)F_2(y) + F(x, y)\overline{F}(x, y)) - \lambda F^2(x, y), \tag{4.8}$$

with the univariate marginal distribution functions

$$G_1(x) = F_1(x)\{1 + \lambda\overline{F}_1(x)\} \quad \text{and} \quad G_2(y) = F_2(y)\{1 + \lambda\overline{F}_2(y)\}. \tag{4.9}$$

The resulting distribution is of the form (1.3) with

$$W(u, v) = uv + D(u, v)\overline{D}(u, v) + \lambda\{uv + D(u, v)\overline{D}(u, v) - D^2(u, v)\},$$

where D is the copula of the baseline cdf F. For the special case when F is equal to FGM distribution (1.1) one gets a bivariate distribution of the form

$$\begin{aligned} G(x, y) = F_1(x)F_2(y)\big(&F_1(x)F_2(y)\overline{F}_1^2(x)\overline{F}_2^2(y)\alpha^2 \\ &- \overline{F}_1(x)\overline{F}_2(y)((1+\lambda-2F_2(y))F_1(x) \\ &- \overline{F}_2^2(y)(1+\lambda))\alpha + F_1(x)(\lambda+\overline{F}_2(y)) + (1+\lambda)(1-\overline{F}_2(y))\big). \end{aligned} \tag{4.10}$$

The associated Spearman's rho with this distribution is given by

$$\varrho = \frac{1}{3} + \alpha\Big(\frac{1}{75}\alpha + \frac{1}{6}\Big) + \lambda^2\Big(\frac{11}{150}\alpha - \frac{1}{3}\Big).$$

Note that $\varrho \in [-0.227, 0.513]$ which is larger than the Spearman's rho associated with usual FGM family. The maximum value of the Spearman's ϱ is equal to 0.513 which happens at $(\lambda, \alpha) = (0, 1)$.

In the general structure (4.8), if one uses the extended FGM introduced by Huang and Kotz [15], given by

$$F(x, y) = F_1(x)F_2(y)(1 + \alpha(1 - F_1^2(x))(1 - F_2^2(y))), \ -\frac{1}{2} \leqslant \alpha \leqslant \frac{1}{2},$$

the resulting distribution has $\varrho \in [-0.313, 0.539]$, which is a good progress in extending the dependence of the usual FGM family. The maximum value of the Spearman's ϱ is equal to 0.539 which happens at $(\lambda, \alpha) = (0, 1)$. Table 1 shows the detailed range of Spearman's rho in the proposed models.

		Baseline cdf $F(x,y)$									
ϱ		FGM					Extended FGM				
	$\alpha \rightarrow$	-1	-0.5	0	0.5	1	-0.5	-0.25	0	0.25	0.5
$\lambda \downarrow$	-1	**-0.227**	-0.117	0	0.123	0.253	**-0.313**	-0.162	0	0.172	0.354
	-0.5	0.078	0.161	0.250	0.346	0.448	0.031	0.136	0.250	0.374	0.507
	0	0.180	0.253	0.333	0.420	**0.513**	0.162	0.244	0.333	0.432	**0.539**
	0.5	0.078	0.161	0.250	0.346	0.448	0.084	0.163	0.250	0.345	0.447
	1	-0.227	-0.117	0	0.123	0.253	-0.199	-0.103	0	0.110	0.228
		-0.333	-0.167	0	0.167	0.333	-0.375	-0.188	0	0.188	0.375

Table 1. Range of Spearman's rho for models proposed in Example 3.

5. Concluding remarks

We presented a general new structure for the so-called FGM family of bivariate distributions. We presented different dependence properties of the proposed model. Several new extensions of the FGM family are discussed. Although the development here is for the bivariate case, the generalization to higher dimensions is straightforward. In the following we present an example. By considering univariate distribution functions F_i, $i = 1, \dots, p$, a multivariate generalization of the family (4.2) defined in Example 1 could be

$$H^{+}(x_1, \dots, x_p) = \prod_{i=1}^{p} F_i(x_i) + W(F_1(x_1), \dots, F_p(x_p)).$$

The kernel W is then

$$\begin{aligned} W(u_1, \dots, u_p) &= F_{V_1,\dots,V_p}(u_1, \dots, u_p) - u_1 u_2 \dots u_p \\ &= \prod_{i=1}^{p} u_i \left\{ \frac{1+\theta}{2} \prod_{i=1}^{p} (2 - u_i) + \frac{1-\theta}{2} \prod_{i=1}^{p} u_i - 1 \right\}, \end{aligned}$$

where $F_{V_1,\dots,V_p}$ is the joint distribution function of the vector

$$(V_1, \dots, V_p) = (\min(X_{11}, X_{12}), \dots, \min(X_{p1}, X_{p2}))$$

with the probability $\frac{1}{2}(1+\theta)$ and

$$(V_1, \dots, V_p) = (\max(X_{11}, X_{12}), \dots, \max(X_{p1}, X_{p2}))$$

with the probability $\frac{1}{2}(1-\theta)$, where X_{i1} and X_{i2}, $i = 1, 2, \dots, p$ are independent uniform $(0, 1)$ random variables.

The functions $W(\cdot,\cdot)$ that we used here are constructed based on the joint distribution functions of the order statistics and record values of a sample of size two. A natural question arises how to find methods for constructing a large and diverse collection of functions $W(\cdot,\cdot)$ to generate extensions with stronger dependence than the usual FGM family.

References

[1] *C. Amblard, S. Girard*: Symmetry and dependence properties within a semiparametric family of bivariate copulas. J. Nonparametric Stat. *14* (2002), 715–727.

[2] *C. Amblard, S. Girard*: A new extension of bivariate FGM copulas. Metrika *70* (2009), 1–17.

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

[4] *I. Bairamov, S. Kotz, M. Bekçi*: New generalized Farlie-Gumbel-Morgenstern distributions and concomitants of order statistics. J. Appl. Stat. *28* (2001), 521–536.

[5] *I. Bairamov, S. Kotz, O. L. Gebizlioglu*: The Sarmanov family and its generalization. S. Afr. Stat. J. *35* (2001), 205–224.

[6] *R. Baker*: An order-statistics-based method for constructing multivariate distributions with fixed marginals. J. Multivariate Anal. *99* (2008), 2312–2327.

[7] *J. D. Church, B. Harris*: The estimation of reliability from stress-strength relationships. Technometrics *12* (1970), 49–54.

[8] *H. A. David, H. N. Nagaraja*: Order Statistics. Wiley Series in Probability and Statistics, John Wiley & Sons, Chichester, 2003.

[9] *D. Drouet-Mari, S. Kotz*: Correlation and Dependence. Imperial College Press, London, 2001.

[10] *D. J. G. Farlie*: The performance of some correlation coefficients for a general bivariate distribution. Biometrika *47* (1960), 307–323.

[11] *M. Fisher, I. Klein*: Constructing generalized FGM copulas by means of certain univariate distributions. Metrika *65* (2007), 243–260.

[12] *E. J. Gumbel*: Bivariate exponential distributions. J. Am. Stat. Assoc. *55* (1960), 698–707.

[13] *K. H. Han*: A new family of negative quadrant dependent bivariate distributions with continuous marginals. Journal of the Chungcheong Mathematical Society *24* (2011), 795–805.

[14] *D. Hlubinka, S. Kotz*: The generalized FGM distribution and its application to stereology of extremes. Appl. Math., Praha *55* (2010), 495–512.

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

[16] *C. D. Lai, M. Xie*: A new family of positive quadrant dependent bivariate distributions. Stat. Probab. Lett. *46* (2000), 359–364.

[17] *S. M. Mirhoseini, A. Dolati, M. Amini*: On a class of distributions generated by stochastic mixture of the extreme order statistics of a sample of size two. J. Stat. Theory Appl. *10* (2011), 455–468.

[18] *D. Morgenstern*: Einfache Beispiele zweidimensionaler Verteilungen. Mitt.-Bl. Math. Statistik *8* (1956), 234–235. (In German.)

[19] *R. B. Nelsen*: An Introduction to Copulas. Springer Series in Statistics, Springer, New York, 2006.

[20] *J. A. Rodríguez-Lallena, M. Úbeda-Flores*: A new class of bivariate copulas. Stat. Probab. Lett. *66* (2004), 315–325.
[21] *M. Sklar*: Fonctions de répartition à n dimensions et leurs marges. Publ. Inst. Stat. Univ. Paris *8* (1960), 229–231. (In French.)

Authors' addresses: *Sayed Mohsen Mirhosseini*, Department of Statistics, Yazd University, Yazd, 89195-741, Iran, e-mail: `mmirhoseini@yazd.ac.ir`; *Mohammad Amini*, Department of Statistics, Ordered and Spatial Data Center of Excellence, Ferdowsi University of Mashhad, P. O. Box 1159, Mashhad 91775, Iran, e-mail: `m-amini@um.ac.ir`; *Ali Dolati*, Department of Statistics, Yazd University, Yazd, 89195-741, Iran, e-mail: `adolati@yazd.ac.ir`.