Thailand Statistician
January 2021; 19(1): 162-177
http://statassoc.or.th
Contributed paper

Generalized Family of Copulas: Definition and Properties

Hakim Bekrizadeh
Department of Statistics, Payame-Noor University, Tehran, Iran.
Corresponding author; e-mail: H_bekri@yahoo.com

Received: 23 January 2019
Revised: 12 August 2019
Accepted: 25 August 2019

Abstract

In this paper, through limiting properties of generalized copula introduced, we propose a new family of copula, which includes some of known copulas. For this family, some general formulas for well-known association measures and concepts of dependence of the proposed model are obtained. This class highlights the usefulness of the asymmetric extended copula for modeling the interested variables whose marginal distribution effect on joint distribution is not the same. We also present a method to simulate from our generalized copula, and validate our method and its accuracy using the simulation results to recover the same dependency structure of the original data.

Keywords: Copula, measures of association.

1. Introduction

Sklar (1959) showed that if H is a bivariate distribution function with margins $F(X)$ and $G(Y)$, there must exist a copula C such that $H_\theta(X,Y) = C(F(X), G(Y);\theta)$, where θ is introduced as dependence parameter.

Copula is mostly defined as a function $C:[0,1]^2 \to [0,1]$ that satisfies boundary conditions:

(P1) $C(x,0) = C(0,x) = 0$ and $C(x,1) = C(1,x) = x, \forall x \in [0,1]$,

(P2) $\forall (s_1, s_2, t_1, t_2) \in [0,1]^4$, such that $s_1 \le s_2$ and $t_1 \le t_2$,

$C(s_2,t_2) - C(s_2,t_1) - C(s_1,t_2) + C(s_1,t_1) \ge 0.$

Eventually, for twice differentiable, 2-increasing property (P2) can be replaced by the condition

$$c(s,t) = \frac{\partial^2 C(s,t)}{\partial s \partial t} \ge 0, \tag{1}$$

where $c(s,t)$ is the so-called copula density. A copula C is symmetric if $C(s,t) = C(t,s)$, for every $(s,t) \in [0,1]^2$, otherwise C is asymmetric.

The copula is in comparison to correlation invariant under transformations of the risks. Correlation is a scalar measure of dependence; it does not tell us everything we would like to know about the dependence structure of risks. A copula determines the dependence relationship by joining the marginal distributions together to form a joint distribution. The scaling and the shape are entirely determined by the marginals. In contrast to correlation the copula function can be applied when risks are heavily tailed.

Many different copulas can be found in the literature, see, for instance Nelsen (2006). One of the most popular parametric families of copulas, which were studied by Farlie (1960), Gumbel (1960) and Morgenstern (1956), is the Farlie-Gumbel-Morgenstern (FGM) copula defined by

$$C^{FGM}(s,t) = st\left[1+\theta(1-s)(1-t)\right],\ \theta\in[-1,1],\ \forall(s,t)\in[0,1]^2,$$

where θ is called the association parameter. However, this copula has been shown to be somewhat limited. This limitation for the dependence parameter $\theta\in[-1,1]$, the Spearman's rho is $\rho_S = \theta/3 \in [-1/3,\ 1/3]$. Since the correlation range of FGM copula is limited, a large number of classes of more extensive copulas have been introduced with the aim of improving the correlation range. Huang and Kotz (1999) proposed two kernel extensions of FGM copulas are studied

$$C^{HK1}(s,t) = st\left[1+\theta(1-s)^{\alpha}(1-t)^{\alpha}\right],\ \forall\alpha\geq 1,$$

$$C^{HK2}(s,t) = st\left[1+\theta(1-s^{\alpha})(1-t^{\alpha})\right],\ \forall\alpha\geq 1/2.$$

It is shown that Spearman's rho can be increased up to approximately 0.39 while the lower bound remains 1/3. Another similar extension is

$$C^{LX}(s,t) = st\left[1+\theta s^{\alpha}(1-s)^{\beta}t^{\alpha}(1-t)^{\beta}\right],$$

where by introduced Lai and Xie (2000).

Copulas Huang and Kotz (1999) and Lai and Xie (2000) are particular cases of Bairamov-Kotz family defined by

$$C^{LX}(s,t) = st\left[1+\theta s^{\alpha}(1-s^{\beta})^{n}t^{\alpha}(1-t^{\beta})^{n}\right],$$

and with associated Spearman's rho $\rho_S \in [-0.48,\ 0.50]$.

Bekrizadeh et al. (2012) proposed a new class of generalized FGM copula by

$$C^{LX}(s,t) = st\left[1+\theta(1-s^{\alpha})(1-t^{\alpha})\right]^{n},$$

and showed that their generalization can improve the correlation domain of FGM copula. Pathak and Vellaisamy (2015, 2016) further extended the family given by Bekrizadeh et al. (2012). Recently, Bekrizadeh and Jamshidi (2017) proposed a new class generalized *FGM* family whose dependence is as follows

$$C_{\delta}^{\psi,p}(s,t) = st[1+\delta\psi(s,t)]^{p},\ \ p\in[1,\infty),\ \ \forall(s,t)\in[0,1]^2, \qquad (2)$$

where $\psi:[0,1]\times[0,1]\rightarrow[0,1]$, and $\delta\in\Theta\subseteq[-1,1]$.

The present article, we discuss an extension of bivariate copulas studied by Bekrizadeh and Jamshidi (2017) by methodology, which includes some extended copulas introduced in recent years and the properties of these copulas are studied. Also, explicit expressions for various measures of association are obtained. A main feature of this family is the capability of modeling a wider range of dependence that permits us to extend the range of potential applications of the class in various branches of sciences.

2. Methodology

In this section, a simple way to construct and define a new generalization of copula is presented, that the new generalization of copula requires a set of conditions. Then, using the mathematical method of Bekrizadeh and Jamshidi (2017), a new generalization of the copula different from that of other authors is presented. In the following, specific examples and modes of the generalization of the new copula are presented. Also, it has been shown that the new generalization covers some known copulas.

Suppose ψ is a function from $[0,1]^2$ into $[-1,1]$ be continuously differentiable functions on $(0,1)$, with additional properties:

- ψ is continuously differentiable on $(0,1)$,
- $s\psi_s(s,t) \to 0$, as $s \searrow 0$,
- $st\psi_{st}(s,t) \to 0$, as $(s,t) \searrow (0,0)$,
- $s\psi_s(s,t)$ converges as $s \nearrow 1$,
- $st\psi_{st}(s,t)$ converges as $(s,t) \nearrow (1,1)$.

As an example, the functions $\psi(s,t) = (1-s)(1-t)$ and $\psi(s,t) = \ln(s)\ln(t)$, for all $(s,t) \in [0,1]^2$ are satisfied in the above conditions.

By approaching p to infinity and properties of the function ψ, the limit of copula $C_\delta^{\psi,p}$ in (2), for $\delta = -\dfrac{\theta}{p}$, where $\theta \le p$ is as follows:

$$
\begin{aligned}
C_\theta^\psi(s,t) &= \lim_{p\to\infty} C_\delta^{\psi,p}(s,t),\\
&= \lim_{p\to\infty} st\left[1-\frac{\theta}{p}\psi(s,t)\right]^p,\\
&= st\lim_{p\to\infty}\left[1-\frac{\theta}{p}\psi(s,t)\right]^p,\\
&= st\exp[-\theta\,\psi(s,t)].
\end{aligned}
$$

Now, let us define

$$C_\theta^\psi(s,t) = st\exp[-\theta\,\psi(s,t)],\ \forall s,t \in [0,1],\ \text{for } \theta \in \Theta \subseteq [-1,1]. \qquad (3)$$

The following theorem gives sufficient and necessary conditions on ψ to ensure that C_θ^ψ in (3) is a bivariate copula. Note that the copula is limited to the range of $[0,1]$ and therefore, $\exp[-\theta\psi(s,t)]$ should be bounded on $[0,1]$.

Theorem 1 *Let ψ be differentiable, non-negative and monotonically decreasing or non-positive and monotonically increasing on* $[0,1]$. *If*

A1. $\psi(x,1) = \psi(1,x) = 0,\ \forall x \in [0,1]$,

A2. $|x\psi_x| \le 1,\ x \in [0,1]$, *and* $|st\psi_{st}| \le 1,\ \forall (s,t) \in [0,1]$, *where* $\psi_x = \partial\psi/\partial x$ *and* $\psi_{st} = \partial^2\psi/\partial s\partial t$.

(i) if $\psi_x,\ \forall x \in [0,1]$ *and* $\psi_{st},\ \forall (s,t) \in [0,1]^2$ *have the same sign, then* C_θ^ψ *is a copula, for* $\theta \in [0,1]$; *i.e.* $\Theta = [0,1]$.

(ii) if $\psi_x,\ \forall x \in [0,1]$ *and* $\psi_{st},\ \forall (s,t) \in [0,1]^2$ *do not have the same sign, then* C_θ^ψ *is a copula, for every* $\theta \in [-1,0]$; *i.e.* $\Theta = [-1,0]$.

Proof: The proof involves two steps:

First, it is clear that $C_\theta^\psi(x,1) = C_\theta^\psi(1,x) = x,\ \forall x \in [0,1] \Leftrightarrow$ (A1).

Second,

(i) since ψ_x and ψ_{st} are differentiable and monotone on $[0,1]$ and have the same sign, eventually for twice differentiable C_θ^ψ the 2-increasing property (1) can be replaced by the condition

$$c_\theta^\psi(s,t) = \partial^2 C_\theta^\psi(s,t)/\partial s\partial t,$$
$$= e^{-\theta\psi(s,t)}\left\{[1-\theta\, s\psi_s(s,t)][1-\theta\, t\psi_t(s,t)] - \theta\, st\psi_{st}(s,t)\right\}, \quad (4)$$

is nonnegative, if $|x\psi_x| \le 1$, $\forall x \in [0,1]$, $\theta \in [0,1]$. Because, let $\theta \in [0,1]$, then

$1-\theta\, s\psi_s(s,t) \ge 1$, $\forall s,t \in [0,1]$,

and

$1-\theta\, t\psi_t(s,t) \ge 1$, $\forall s,t \in [0,1]$,

with $|st\psi_{st}| \le 1$, $\forall (s,t) \in [0,1]$, we obtain

$|\theta\, st\psi_{st}(s,t)| \le 1$, $\forall s,t \in [0,1]$.

Therefore,

$[1-\theta s\psi_s(s,t)][1-\theta t\psi_t(s,t)] - \theta st\psi_{st}(s,t) \ge 0$, $\forall s,t \in [0,1]$.

(ii) if $\theta \in [-1,0]$ and ψ_x, ψ_{st} do not have the same sign. Then, similar to (i) with conditions $|x\psi_x| \le 1$, $\forall x \in [0,1]$, we have

$[1-\theta s\psi_s(s,t)][1-\theta t\psi_t(s,t)] \ge 1$, $\forall s,t \in [0,1]$,

with $|st\psi_{st}| \le 1$, $\forall (s,t) \in [0,1]$,

$|\theta\, st\psi_{st}(s,t)| \le 1$, $\forall s,t \in [0,1]$.

Therefore,

$[1-\theta s\psi_s(s,t)][1-\theta t\psi_t(s,t)] - \theta st\psi_{st}(s,t) \ge 0$, $\forall s,t \in [0,1]$.

So, the function C_θ^ψ satisfies 2-increasing property.

Example 1 Given $\psi(s,t) = (1-s)\ln(t)$, $\forall s,t \in [0,1]$.

Since $|s\psi_s| = |-s\ln(t)| \le 1$, $|t\psi_t| = |-(1-s)| \le 1$, and $|st\psi_{st}| = |-s| \le 1$, $\forall (s,t) \in [0,1]$, then by Theorem 1,

$C_\theta(s,t) = st\exp(-\theta(1-s)\ln(t))$,

is a bivariate copula for $\theta \in \Theta \subseteq [-1,1]$.

Example 2 Given $\psi(s,t) = (1-\sqrt{s})t(1-t)$, $\forall s,t \in [0,1]$.

Since $|s\psi_s| = \left|-\frac{t(1-t)\sqrt{s}}{2}\right| \le 1$, $|t\psi_t| = \left|(1-\sqrt{s})t(1-2t)\right| \le 1$, $|st\psi_{st}| = \left|-\frac{t(1-2t)\sqrt{s}}{2}\right| \le 1$,

$\forall (s,t) \in [0,1]$, then

$C_\theta(s,t) = st\exp(-\theta(1-\sqrt{s})t(1-t))$,

is a bivariate copula for $\theta \in \Theta \subseteq [-1,1]$.

Example 3 Given $\psi(s,t) = (1-s^\beta)(e^{-t}-e^{-1})$, $\forall s,t \in [0,1]$.

Since $|s\psi_s| = \left|-\beta s^\beta(e^{-t}-e^{-1})\right| \le 1$, $|t\psi_t| = \left|-t(1-s^\beta)e^{-t}\right| \le 1$, $|st\psi_{st}| = \left|\beta s^\beta t e^{-t}\right| \le 1$,

$\forall (s,t) \in [0,1]$, then

$C_\theta(s,t) = st\exp(-\theta 1-s^\beta)(e^{-t}-e^{-1}))$, is a bivariate copula for $\theta \in \Theta \subseteq [-1,1]$.

Corollary 1 *Let* $\psi(s,t) = f(s)f(t)$, *that the function* f *be a mapping from* $[0,1]$ *into* $[-1,1]$ *and differentiable and monotonic on* $[0,1]$, *and fulfilling the conditions*

A1. $f(1) = 0$,

A2. $|xf'(x)| \le 1,\ \ x \in [0,1]$,

then, the function $C_\theta^f(s,t) = st\exp[-\theta f(s)f(t)]$ *is a bivariate copula for* $\theta \in [0,1]$; *i.e.* $\Theta = [0,1]$.

Proof: Similar to Theorem 1.

Corollary 2 *Let* $\psi(s,t) = g_1(s)g_2(t)$, *that the functions* g_i *be a mapping from* $[0,1]$ *into* $[-1,1]$ *and differentiable and monotonic on* $[0,1]$, *and fulfilling the conditions*

A1. $\psi_i(1) = 0$,

A2. $|xg_i'(x)| \le 1,\ x \in [0,1]$, *for* $i = 1,2$, *where* $g_i'(x) = \partial g_i(x)/\partial x$.

Then,

(i) if g_i *have the same sign, then the function* $C_\theta^{g_1,g_2}(s,t) = st\exp[-\theta\, g_1(s)g_2(t)]$ *is a copula for* $\theta \in [0,1]$; *i.e.* $\Theta = [0,1]$.

(ii) if g_i *do not have the same sign, then the function* $C_\theta^{g_1,g_2}(s,t) = st\exp[-\theta g_1(s)g_2(t)]$ *is a copula for* $\theta \in [-1,0]$; *i.e.* $\Theta = [-1,0]$.

Proof: Similar to Theorem 1.

Remark 1 Note that θ is a parameter that shows the dependence structure of the class C_θ^ψ and $\theta = 0$, leads to the independence of S and T.

Remark 2 The class C_θ^ψ is an symmetric class of bivariate copula, if $\psi(s,t) = \psi(t,s),\ \forall (s,t) \in [0,1]^2$. Moreover, one of the advantages of the class C_θ^ψ is generating some sub–families of class $\{C_\theta^\psi\}$ through suitable choosing function ψ.

Remark 3 the concreted amount of the parameter space θ depends on the properties of the function ψ that has been investigated via (3) for every s and t in [0,1].

Remark 4 The family C_θ^ψ includes some known family of copulas introduced by researchers in recent years, which are as follows:

(i) if $\psi(s,t) = \ln(s)\ln(t),\ \ \forall s,t \in [0,1]$, the family C_θ^ψ reduce to the symmetric Gumbel-Barnett copula discussed by Hutchinson and Lai (1990).

(ii) if $\psi(s,t) = (1-s)(1-t),\ \ \forall s,t \in [0,1]$, the family C_θ^ψ reduce to the symmetric NC copula discussed by Cuadras (2009).

3. Measures of Dependence

Measures of dependence are common instruments to summarize a complicated dependence structure in the bivariate case. Correlation comes in trouble when the random variables are not elliptically distributed. The performance of the copula does not depend on the fact if you are dealing with elliptical distributions or not. Add the fact that copulas possess handy properties and the winner of the two is the copula.

Pearson's correlation coefficient is a common statistical measure. This linear correlation measure $(-1 \le \rho \le 1)$ is the most popular and well-known measure between pairwise random variables. Despite its simplicity and plain rationale, (Embrechts et al. 2001) note that ρ is simply a measure of the dependency of elliptical distributions, such as the binormal distribution (i.e. the marginals are normally distributed, linked by the Gaussian copula). Moreover, ρ measures a linear relationship itself and does not capture a non-linear one on its own, as noted in (Priest 2003). These properties constitute obvious limitations for modeling the dependency structure.

Regarding essential roles of copula in modeling dependency between variables of interest, it is very important to construct different kinds of copulas. Also, via this strategy, copulas could be useful to define nonparametric measures of dependence between random variables. One of the most important nonparametric measures of dependence is Spearman's rho (ρ_S). It is notable that the Spearman's rho coincides with correlation coefficient (ρ_S) between the marginal distributions. Since a copula is able to capture dependence structures regardless of the form of the margins, this approach is potentially very useful in modeling related variables especially for econometricians. For a historical review of measures of dependence, (See Joe 1997, Nelsen 2006).

This section sets out several fundamental measures of dependence for C_θ^ψ in (3). Let X and Y be continuous random variables whose copula is C. The Spearman's rho (ρ_S), Kendall's tau (τ_k), Gini's gamma (γ_C), and Spearman's footrule coefficient (δ_C) for X and Y are given by

$$\rho_S = 12\int_0^1\int_0^1 C(u,v)\,du\,dv - 3, \tag{5}$$

$$\tau_k = 4\int_0^1\int_0^1 c(u,v)C(u,v)\,du\,dv - 1, \tag{6}$$

$$\gamma_C = 4\int_0^1 [C(s,s) + C(s,1-s)]\,ds - 1, \tag{7}$$

$$\delta_C = 6\int_0^1 C(s,s)\,ds - 2, \tag{8}$$

respectively.

Since we cannot obtain formulas for measures of dependence in terms of elementary function ψ, the class C_θ^ψ is replaced by its Taylor expansion series with respect to ψ as

$$\begin{aligned} C_\theta^\psi(s,t) &= st\sum_{k=0}^{\infty}\frac{(-\theta)^k}{k!}\psi^k(s,t), \\ &= st + \sum_{k=1}^{\infty}\frac{(-\theta)^k}{k!}st\psi^k(s,t). \end{aligned} \tag{9}$$

In the case of a copula generated by (3), these measures rewrite only in terms of the function ψ.

Proposition 1 *Let* (X,Y) *be a pair of random variables with distribution belonging to the family* C_θ^ψ. *The Spearman's rho for the family* C_θ^ψ *is given by*

$$\rho_S = 12\sum_{k=1}^{\infty}\frac{(-\theta)^k}{k!}D(k), \tag{10}$$

where

$$D(k) = \int_0^1\int_0^1 st\psi^k(s,t)\,ds\,dt. \tag{11}$$

Proof: Replacing (9) in (5), the Spearman's rho for the copula C_θ^ψ can be expanded as

$$\begin{aligned}
\rho_S &= 12\int_0^1\int_0^1 C_\theta^\psi(s,t)\,ds\,dt - 3,\\
&= 12\int_0^1\int_0^1\left(st + \sum_{k=1}^{\infty}\frac{(-\theta)^k}{k!}st\psi^k(s,t)\right)dsdt - 3,\\
&= 12\int_0^1\int_0^1 st\,dsdt + 12\int_0^1\int_0^1\left(\sum_{k=1}^{\infty}\frac{(-\theta)^k}{k!}st\psi^k(s,t)\right)dsdt - 3,\\
&= 12\int_0^1\int_0^1\sum_{k=1}^{\infty}\frac{(-\theta)^k}{k!}st\psi^k(s,t)dsdt,\\
&= 12\sum_{k=1}^{\infty}\frac{(-\theta)^k}{k!}\int_0^1\int_0^1 st\psi^k(s,t)dsdt,\\
&= 12\sum_{k=1}^{\infty}\frac{(-\theta)^k}{k!}D(k).
\end{aligned}$$

Proposition 2 *Let* (X,Y) *be a pair of random variables with the copula* C_θ^ψ. *The Kendall's tau for the family* C_θ^ψ *is given by*

$$\tau_k = 4\sum_{k=0}^{\infty}\frac{(-2)^k}{k!}\theta^k\left\{D(k) + 4\frac{\theta}{k+1}\frac{k-2}{k+2}D(k+1) + \theta^2\frac{k+6}{k+2}\left(\frac{2}{k+1}\right)^2 D(k+1)\right\} - 1. \tag{12}$$

Proof: Using (3) and (4), we have

$$\begin{aligned}
c_\theta^\psi(s,t)C_\theta^\psi(s,t) &= st\exp(-2\theta\psi(s,t))\\
&\quad\times\left\{[1-\theta s\psi_s(s,t)][1-\theta t\psi_t(s,t)] - \theta st\psi_{st}(s,t)\right\}
\end{aligned}$$

The relation above may be written through Taylor expansion series with respect to ψ as:

$$\begin{aligned}
c_\theta^\psi(s,t)C_\theta^\psi(s,t) &= \sum_{k=0}^{\infty}\frac{(-2)^k}{k!}\theta^k st\psi^k(s,t)\left\{[1-\theta s\psi_s(s,t)][1-\theta t\psi_t(s,t)] - \theta st\psi_{st}(s,t)\right\},\\
&= \sum_{k=0}^{\infty}\frac{(-2)^k}{k!}\theta^k\left\{st\psi^k(s,t) - \theta st^2\psi_t(s,t)\psi^k(s,t) - \theta s^2 t\psi_s(s,t)\psi^k(s,t)\right.\\
&\qquad\left. + \theta^2 s^2t^2\psi_s(s,t)\psi_t(s,t)\psi^k(s,t) - \theta s^2t^2\psi_{st}(s,t)\psi^k(s,t)\right\}.
\end{aligned}$$

Using integration by parts, we have

$$\int_0^1 s^2\psi_s(s,t)\psi^k(s,t)ds = \frac{s^2}{k+1}\psi^{k+1}(s,t)\Big|_0^1 - \frac{2}{k+1}\int_0^1 s\psi^{k+1}(s,t)ds,$$
$$= -\frac{2}{k+1}\int_0^1 s\psi^{k+1}(s,t)ds.$$

Thus,

$$\int_0^1\int_0^1 s^2\, t\psi_s(s,t)\psi^k(s,t)dsdt = -\frac{2}{k+1}\int_0^1\int_0^1 st\psi^{k+1}(s,t)dsdt,$$
$$= -\frac{2}{k+1}D(k+1).$$

Similarly,

$$\int_0^1\int_0^1 s\,t^2\psi_t(s,t)\psi^k(s,t)dsdt = -\frac{2}{k+1}D(k+1).$$

Also,

$$\int_0^1 s^2\psi_{st}(s,t)\psi^k(s,t)ds = \frac{s^2}{k+1}\psi_t(s,t)\psi^{k+1}(s,t)\Big|_0^1 - \frac{2}{k+1}\int_0^1 s\psi_t(s,t)\psi^{k+1}(s,t)ds,$$
$$= -\frac{2}{k+1}\int_0^1 s\psi_t(s,t)\psi^{k+1}(s,t)ds.$$

Thus,

$$\int_0^1\int_0^1 s^2t^2\psi_{st}(s,t)\psi^k(s,t)dsdt = -\frac{2}{k+1}\int_0^1\int_0^1 st^2\psi_t(s,t)\psi^{k+1}(s,t)dsdt,$$
$$= \frac{4}{(k+1)(k+2)}D(k+1).$$

Moreover,

$$\int_0^1 s^2\psi_s(s,t)\psi_t(s,t)\psi^k(s,t)ds = \frac{s^2}{k+1}\psi_t(s,t)\psi^{k+1}(s,t)\Big|_0^1 - \frac{2}{k+1}\int_0^1 s\psi_t(s,t)\psi^{k+1}(s,t)ds$$
$$-\frac{1}{k+1}\int_0^1 s^2\psi_{st}(s,t)\psi^{k+1}(s,t)ds,$$
$$= -\frac{2}{k+1}\int_0^1 s\psi_t(s,t)\psi^{k+1}(s,t)ds - \frac{1}{k+1}\int_0^1 s^2\psi_{st}(s,t)\psi^{k+1}(s,t)ds.$$

Thus,

$$\int_0^1\int_0^1 s^2t^2\psi_s(s,t)\psi_t(s,t)\psi^k(s,t)dsdt = -\frac{2}{k+1}\int_0^1\int_0^1 st^2\psi_t(s,t)\psi^{k+1}(s,t)dsdt,$$
$$-\frac{1}{k+1}\int_0^1\int_0^1 s^2t^2\psi_{st}(s,t)\psi^{k+1}(s,t)dsdt,$$
$$= \left(\frac{2}{k+1}\right)^2 D(k+1) + \frac{4}{(k+1)^2(k+2)}D(k+1),$$
$$= \frac{k+6}{k+2}\left(\frac{2}{k+1}\right)^2 D(k+1).$$

By using (9), the Kendall's tau (τ_k) can be expanded as

$$\tau_k = 4\int_0^1\int_0^1 c_\theta^\psi(s,t)C_\theta^\psi(s,t)dsdt - 1,$$

$$= 4\sum_{k=0}^{\infty}\frac{(-2)^k}{k!}\theta^k\left\{\int_0^1\int_0^1 st\psi^k(s,t)dsdt - \theta\int_0^1\int_0^1 s\,t^2\psi_t(s,t)\psi^k(s,t)dsdt\right.$$

$$-\theta\int_0^1\int_0^1 s^2\,t\psi_s(s,t)\psi^k(s,t)dsdt + \theta^2\int_0^1\int_0^1 s^2t^2\psi_s(s,t)\psi_t(s,t)\psi^k(s,t)dsdt$$

$$\left.-\theta\int_0^1\int_0^1 s^2t^2\psi_{st}(s,t)\psi^k(s,t)dsdt\right\} - 1,$$

$$= 4\sum_{k=0}^{\infty}\frac{(-2)^k}{k!}\theta^k\left\{D(k) + 4\frac{\theta}{k+1}\frac{k-2}{k+2}D(k+1) + \theta^2\frac{k+6}{k+2}\left(\frac{2}{k+1}\right)^2 D(k+1)\right\} - 1.$$

Proposition 3 *Let* (X,Y) *be a pair of random variables with distribution belonging to the family* C_θ^ψ *and the density family* c_θ^ψ*; then the direction of equality between Gini's gamma* (γ_C) *and the Spearman's footrule* (δ_C) *is given by*

$$3\gamma_C - 2\delta_C = 12\sum_{k=1}^{\infty}\frac{(-\theta)^k}{k!}\int_0^1 s(1-s)\psi^k(s,1-s)ds - 1. \tag{13}$$

Proof: The proof is straightforward.

Corollary 3 *Let* $\psi(s,t) = f(s)f(t)$*, that* $f:[0,1]\to[-1,1]$*, differentiable and monotonic on* $[0,1]$*, and fulfilling the conditions Corollary 1, then*

$$D(k) = d^2(k),$$

where $d(k) = \int_0^1 x\,f^k(x)\,dx.$

Proof: Replacing $\psi(s,t) = f(s)f(t)$ in (11); $D(k)$ can be expanded as

$$D(k) = \int_0^1\int_0^1 st\psi^k(s,t)\,ds\,dt,$$

$$= \int_0^1\int_0^1 st\,f^k(s)\,f^k(t)\,ds\,dt,$$

$$= \left(\int_0^1 sf^k(s)\,ds\right)\left(\int_0^1 t\,f^k(t)\,dt\right),$$

$$= \left(\int_0^1 xf^k(x)\,dx\right)^2,$$

$$= d^2(k).$$

Corollary 4 Let $\psi(s,t) = g_1(s)\,g_2(t)$, that $g_i:[0,1]\to[-1,1]$, differentiable and monotonic on $[0,1]$, and fulfilling the conditions Corollary 2, then

$$D(k) = D_1(k)D_2(k),$$

where $D_i(k) = \int_0^1 x g_i^k(x)\,dx$, for $i = 1,2$.

Proof: Similar to Corollary 3.

Example 1 (Gumbel-Barnett copula): Let $\psi(s,t) = \ln(s)\ln(t), \forall s,t \in [0,1]$ or $f(x) = \ln(x), \forall x \in [0,1]$, then

$$d(k) = \frac{(-1)^k \Gamma(k+1)}{2^{k+1}},$$

and

$$\rho_S = 3\sum_{k=1}^{\infty} \Gamma(k+1)\left(-\frac{\theta}{4}\right)^k,$$

$$\tau_k = \sum_{k=0}^{\infty} \frac{(-\theta)^k \Gamma(k+1)}{2^k}\left\{1 + k\theta + \frac{1}{4}(k+1)^2\theta^2\right\} - 1.$$

Example 2 (Celebioglu-Cuadras copula; (Cuadras 2009)): Let $\psi(s,t) = (1-s)(1-t), \forall s,t \in [0,1]$ or $f(x) = 1-x, \forall x \in [0,1]$, then

$$d(k) = \frac{1}{(k+1)(k+3)},$$

and
$$\rho_S = \frac{3}{\theta^2}\left(12 + 8\theta - \theta^2\right) - \frac{12}{\theta^2}e^{\theta} - \frac{12}{\theta^2}(1-\theta)\left(Ei(1,-\theta) + \gamma + \ln(-\theta)\right),$$

$$\tau_k = \frac{1+4\theta}{\theta^2} - \frac{e^{2\theta}}{\theta^2} - \frac{(1-2\theta)\left(Ei(1,-2\theta) + \gamma + \ln(-2\theta)\right)}{\theta^2}$$
$$+ \frac{1}{12}\left(\frac{21 + 72\theta - 48\theta^2}{\theta^2} - \frac{(21-6\theta)e^{2\theta}}{\theta^2} - \frac{(12\theta^2 - 48\theta + 18)\left(Ei(1,-2\theta) + \gamma + \ln(-2\theta)\right)}{\theta^2}\right) - 1,$$

in which $\gamma = -0.5772156649$ and $Ei(t,z)$ is the exponential integrals by $Ei(t,z) = z^{t-1}\Gamma(1-t,z)$.

Example 3 Let $\psi(s,t) = (1-s)\ln(t)$ or $g_1(x) = 1-x$ and $g_2(x) = \ln(x)$, for every $x \in [0,1]$, then

$$D_1(k) = \frac{1}{(k+1)(k+2)} \text{ and } D_2(k) = \frac{(-1)^k \Gamma(k+1)}{2^{k+1}}.$$

Using (6),

$$\rho_S = \frac{12}{\theta^2}\left[(2+\theta)\ln\left(1 + \frac{\theta}{2}\right) - \theta\right] - 3.$$

4. The Tail Dependence

Tail dependence describes the amount of dependence in the upper tail or lower tail of a bivariate distribution and has been widely used in extreme value analysis and in quantitative risk management (Joe 1997; McNeil et al. 2005). Tail dependence is often studied by using the copula method, which is used to explore scale-invariant features for a joint distribution. A common theme is to study decay rates of joint tail probabilities of a random vector using same marginal tail scaling functions (Joe 1997), leading to the so called standard tail dependence. It is a concept that is relevant for the study of dependence between extreme values. It turns out that tail dependence between two continuous random variables X and Y is a copula property and hence the amount of tail dependence is invariant under strictly increasing transformations of X and Y (Joe 1997, Grobmab 2007, Coles et al. 2011).

For a bivariate copula C if

$$\lambda_U = \lim_{x\to 1}\frac{1-2x+C(x,x)}{1-x} \tag{14}$$

exists, then C has upper tail dependence if $\lambda_U \in (0,1]$, and upper tail independence if $\lambda_U = 0$. The measure is extensively used in extreme value theory.

The concept of lower tail dependence can be defined in a similar way. If the limit,

$$\lambda_L = \lim_{x\to 0}\frac{C(x,x)}{x} \tag{15}$$

exist, then C has lower tail dependence if $\lambda_L \in (0,1]$, and lower tail independence if $\lambda_L = 0$. A measure to quantify "dependence within tail independence" is suggested by Coles et al. (1999) who defines the weak upper tail dependence coefficient as

$$\chi_U = \lim_{x\to 1}\left(\frac{2\log(1-x)}{\log\left(1-2x+C(x,x)\right)}-1\right), \tag{16}$$

provided that limit in (16) exists. It can be shown that $-1 \le \chi_U \le 1$, $\chi_U = 1$ in the case of upper tail dependence (i.e. for $\lambda_U > 0$), $\chi_U = 0$ in the case of $C(u,v) = uv$ being the independence copula and for copulas with upper tail independence (i.e. with $\lambda_U = 0$), χ_U increases with the strength of dependence in the tail area. In the sequel, we speak of weak upper tail independence if $\chi_U = 0$, and of weak upper tail dependence if $\chi_U \ne 0$. It should be pointed out again that it is not necessary to calculate χ_U in the case of strong upper tail dependence, because $\chi_U = 1$ holds. Likewise, the weak lower tail dependence coefficient equals the limit of

$$\chi_L = \lim_{x\to 0}\left(\frac{2\log(x)}{\log\left(C(x,x)\right)}-1\right). \tag{17}$$

Proposition 4 *Let* (X,Y) *be a pair of random variables with the copula* C_θ^ψ. *Then for the copula* C_θ^ψ, *we have*

(i) $\lambda_U = 0,$

(ii) $\lambda_L = \begin{cases}\infty, & \text{if } -\theta\psi(x,x)\to+\infty\\ 0, & \text{otherwise},\end{cases}$

(iii) $\chi_U = 0,$

(iv) $\chi_L = -\left(1-\dfrac{2}{\theta}M\right)^{-1}$, *where* $M = \lim_{x\to 0}\dfrac{\log(x)}{\psi(x,x)}.$

Proof:

(i) Clearly, the upper tail dependence coefficients (λ_U) can be simplified by replacing (3) in (14) as

$$\lambda_U = \lim_{x\to 1}\left(\frac{1-2x+C_\theta^\psi(x,x)}{1-x}\right) = \lim_{x\to 1}\left(\frac{1-2x+x^2 e^{-\theta\psi(x,x)}}{1-x}\right).$$

Since $\psi(1,x)=\psi(x,1)=0$, $\forall x \in (0,1)$, we have

$$\begin{aligned}\lambda_U &= \lim_{x\to 1}\{ 2-2xe^{-\theta\psi(x,x)} \\ &\quad +\theta[\psi_{x,1}(x,x)\psi(x,x)+\psi_{x,2}(x,x)\psi(x,x)]x^2 e^{-\theta\psi(x,x)}\}, \\ &= 2-2\lim_{x\to 1} xe^{-\theta\psi(x,x)} \\ &\quad +\lim_{x\to 1}\theta[\psi_{x,1}(x,x)\psi(x,x)+\psi_{x,2}(x,x)\psi(x,x)]x^2 e^{-\theta\psi(x,x)}, \\ &= 0.\end{aligned}$$

(ii) The lower tail dependence coefficients (λ_L) can be simplified by replacing (3) in (15) as

$$\begin{aligned}\lambda_L &= \lim_{x\to 0}\frac{C_\theta^\psi(x,x)}{x}, \\ &= \lim_{x\to 0}\frac{x^2 e^{-\theta\psi(x,x)}}{x}, \\ &= \lim_{x\to 0} ue^{-\theta\psi(x,x)},\end{aligned}$$

where λ_L is indeterminate as $-\theta\psi(x,x)$ tends to $+\infty$, otherwise $\lambda_L = 0$.

(iii) The weak upper tail dependence coefficient (χ_U) can be expanded by replacing (3) in (16) as

$$\begin{aligned}\chi_U &= \lim_{x\to 1}\left(\frac{2\log(1-x)}{\log\left(1-2x+C_\theta^\psi(x,x)\right)}-1\right), \\ &= \lim_{x\to 1}\left(\frac{2\log(1-x)}{\log\left(1-2x+x^2 e^{-\theta\psi(x,x)}\right)}-1\right), \\ &= 0.\end{aligned}$$

(iv) The weak lower tail dependence coefficient (χ_L) can be expanded by replacing (3) in (17) as

$$\begin{aligned}\chi_L &= \lim_{x\to 0}\left(\frac{2\log(x)}{\log C_\theta^\psi(x,x)}-1\right) \\ &= \lim_{x\to 0}\left(\frac{2\log(u)}{\log(x^2 e^{-\theta\psi(x,x)})}-1\right) \\ &= \lim_{x\to 0}\left(\frac{\theta\psi(x,x)}{2\log(x)-\theta\psi(x,x)}\right) \\ &= \frac{1}{\frac{2}{\theta}\lim_{x\to 0}\frac{\log(x)}{\psi(x,x)}-1} \\ &= \frac{1}{\frac{2}{\theta}M-1} \\ &= -\left(1-\frac{2}{\theta}M\right)^{-1}.\end{aligned}$$

Example 1 Let $\psi(s,t)=\ln(s)\ln(t)$, $\forall s,t\in[0,1]$, then $\lambda_U=\lambda_L=\chi_U=0$ and $\chi_L=-1$.

Example 2 Let $\psi(s,t)=(1-s)(1-t)$, $\forall s,t\in[0,1]$, then $\lambda_U=\lambda_L=\chi_U=\chi_L=0$.

Example 3 Let $\psi(s,t)=(1-s)\ln(t)$, for every $s,t\in[0,1]$, then $\lambda_U=\chi_U=0$, $\lambda_L=0$ (as $0\le\theta\le1$), and $\chi_L=\dfrac{\theta}{2-\theta}$.

5. Data Analysis and Simulation

In this section, we apply mentioned sub-families of our presented asymmetric bivariate copula to some real dataset in medical science. According to the manual of R's package MASS, the US National Institute of Diabetes and Digestive and Kidney Diseases collected a data set from a population of women (at least 21 years old, of Pima Indian heritage and living near Phoenix, Arizona) whom were tested for diabetes according to World Health Organization criteria. This data set consists of 332 complete records after dropping the (mainly missing) data on serum insulin. This dataset was later reanalyzed by Li and Fang (2012).

There are 8 variables considered as risk factors about Diabetes among Pima Indians, where the Body Mass Index (BMI) and the Diabetes Pedigree Function (PED) were diagnosed as the most important of them. The role of Information is vital in analysis and management of body health tests. Specially, in case of our dataset, the most important part is the study of features frequency of BMI and PED, where we focus on modeling the dependence structure between them. Regarding the correlation and tail dependence of these two factors (Li and Fang 2012), some tools must be used to describe the relationship and analyze mutual impact; therefore it is necessary to determine the joint distribution of the two factors, where, in present paper, in accordance with low correlation coefficient and observed weak tail dependence, this determining is done through fitting mentioned sub-families of introduced new bivariate copula. For modeling, it is necessary to estimate the marginal distributions of BMI and PED and then to combine them by copula. In this paper, we apply Li and Fang (2012) illustration, where the dataset has been fitted by gamma distribution with shape parameter $\alpha_1=21.82$ and scale parameter $\beta_1=0.66$ for BMI data, and, shape parameter $\alpha_2=2.58$ and scale parameter $\beta_2=4.89$ for PED data. In sequel, in order to estimate parameter of copula, the log-likelihood function is computed, and then based on the AIC criterion, the best copula is selected. The results for different models of the introduced bivariate copula are presented in Table 1. This table shows that the class C_θ^ψ, via choosing different type of function ψ and estimating parameter θ is a flexible class of copula; accordingly, the class C_θ^ψ can better fit the interested medical data. Table 1 show that the model (III) with the parameter of $\hat{\theta}=0.243$ can determine the joint distribution of BMI and PED as

$$C(s,t)=st\ \exp[0.243\ (1-s)\ln(t)],$$

where $C(s,t)=H(F^{-1}(s),G^{-1}(t))$, and H is joint distribution of BMI and PED, $F\sim Gamma\ (21.82,0.66)$ and $G\sim Gamma\ (2.58,4.89)$ are marginal distributions of BMI and PED, respectively.

Table 1 Different models of the new bivariate copula

Model	ψ	Admissible range of θ	MLE of θ	AIC

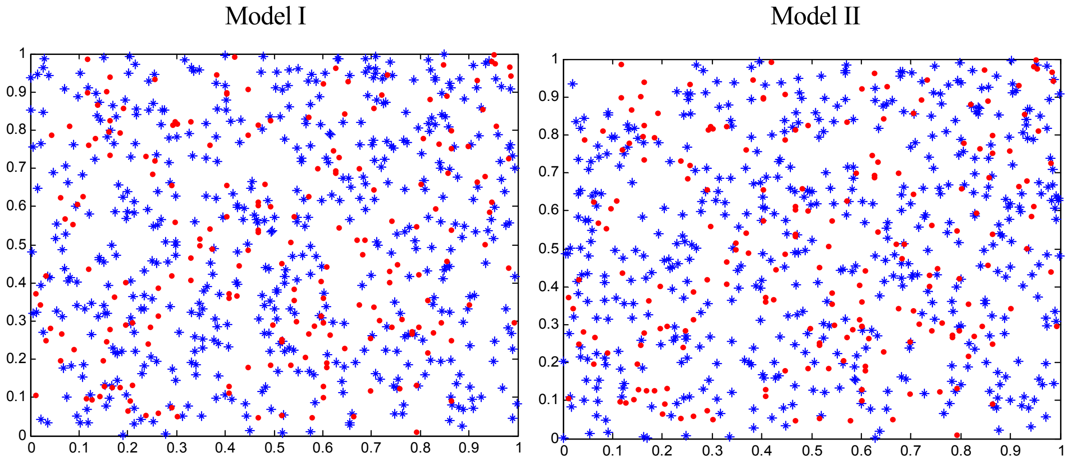
I*	$\ln(s)\ln(t)$	$\theta \in (0,1]$	0.2541	2349.198
II**	$(1-s)(1-t)$	$\theta \in [-1,1]$	-0.3050	2351.184
III	$(1-s)\ln(t)$	$\theta \in [-1,1]$	0.2430	2353.121

Note: *Gumbel-Barnett copula
**NC copula (See Cuadras 2009)

Li and Fang (2012) fitted Clayton, Franck and Sine copulas to this data. They chose Sine copula with AIC = 2342.125. Comparing with AICs of Table 1, we can say introduced copula is better than sine copula, and then the model (III) is more suitable for interested data set in comparison with Sine copula. In addition, these results are investigated using simulation. We now discuss the simulation of data from mentioned subfamilies of introduced bivariate copula and compare scatter plots and correlations of the observed data with the simulated data based on 1000 simulations, followed the simulation method proposed Johnson (1987) and later Nelson (2006).

Figure 1 illustrates the scatter plots of the transformed observed data (•) versus simulated samples of the CDFs of BMI and PED variables (∗) taken from the fitted new bivariate copula in Table 1. It can be seen that the simulated data and the original data have similar dependence patterns but the consistency amount between observed data and simulated data is not so clear in Figure 1.

To settle this concern, Table 2 shows the rank correlations between the BMI and PED variables calculated from the original observed data, and based on the simulated data of size 1000 taken from the fitted mentioned subfamilies of asymmetric bivariate copula. Comparing these correlations, we can conclude the consistency of the estimated correlations based on the mentioned subfamilies of asymmetric bivariate copula with the observed data. Specially, this consistence is more apparent about model (III).



Model III

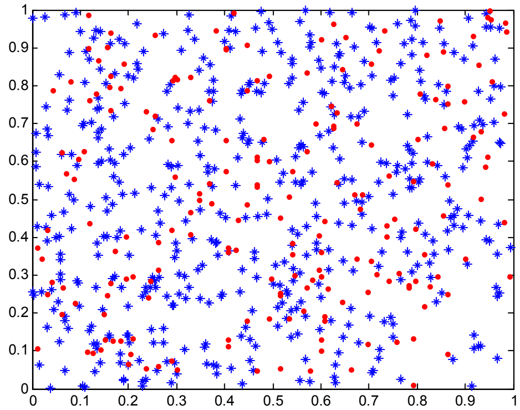


Figure 1 Scatter plots of the transformed observed values (•) versus simulated samples (∗) of BMI and PED variables from the generalized copula model

Table 2 Calculated correlations for different models of the generalized copula

Model	Correlation
Original data	0.1385
I*	0.1431
II**	0.1315
III	0.1381

Note: *Gumbel-Barnett copula
**NC copula (Cuadras 2009)

Acknowledgments

We would like to thank the editor, the associate editor and referees for many helpful comments that substantially improved the article.

References

Bekrizadeh H, Parham GA, Zadkarmi MR. The new generalization of Farlie-Gumbel-Morgenstern copulas. Appl Math Sci. 2012; 71(6): 3527-3533.

Bekrizadeh H, Jamshidi B. A new class of bivariate copulas: dependence measures and properties. Metron. 2017; 75(1): 31-50.

Coles S, Heffernan J, Tawn J. Dependence measures for extreme value analyses. Extremes. 1999; 2: 339-365.

Coles S, Currie J, Tawn J. Dependence measures for extreme value analyses. Extremes. 2011; 2: 339-365.

Cuadras CM. Constructing copula functions with weighted geometric means. J Stat Plan Inference. 2009; 139(11): 3766-3772.

Embrechts P, Lindskog F, McNeil A. Modelling dependence with copulas and applications to risk management. Rapport technique, Département de mathématiques, Institut Fédéral de Technologie de Zurich, Zurich, 14; 2001.

Farlie DGJ. The performance of some correlation coefficients for a general bivariate distribution. Biometrika. 1960; 47(3/4): 307-323.

Grobmab T. Copulae and tail dependence. Diploma [Thesis] Center for Applied Statistics and Economics, Berlin; 2007.

Gumbel EJ. Bivariate exponential distributions. Journal of the American Statistical Association.1960; 55(292): 698-707.

Huang JS, Kotz S. Modifications of the Farlie-Gumbel-Morgenstern distributions. A tough hill to climb. Metrika. 1999; 49(2): 135-145.

Hutchinson TP, Lai CD. Continuous bivariate distributions, emphasising applications. Adelaide: Rumsby Scientific Publishing; 1990.

Joe H. Multivariate models and dependence concepts. London: Chapman and Hall; 1997.

Johnson ME. Multivariate statistical simulation. New York: John Wiley & Sons; 1987.

Lai CD, Xie M. A new family of positive quadrant dependent bivariate distributions. Stat Prob Lett. 2000; 46(4): 359-364.

Li X, Fang R. A new family of bivariate copulas generated by univariate distributions. J Data Sci. 2012; 10: 1-17.

McNeil AJ, Frey R, Embrechts P. Quantitative risk management: concepts, techniques and tools, Princeton: Princeton University Press; 2005.

Morgenstern D. Einfache beispiele zweidimensionaler verteilungen Mitteilungsblatt für Mathematische. Statistik. 1956; 8: 234-235.

Nelsen, RB. An introduction to copulas. New York: Springer; 2006.

Pathak, AK, Vellaisamy P. Various measures of dependence of a new asymmetric generalized Farlie-Gumbel-Morgenstern copulas. Commun Stat - Theory Methods. 2015; 45(18): 5299-5317.

Pathak, AK, Vellaisamy P. A note on generalized Farlie-Gumbel-Morgenstern copulas. J Stat Theory Pract. 2016; 10: 40-58.

Priest C. Correlations: what they mean and more importantly what they don't mean. The Institute of Actuaries of Australia Biennial Convention; 2003.

Sklar A. Fonctions de répartition à n dimensions et leurs marges. Publications de L'Institute Statistical University Paris. 1959; 8: 229-231.