# Another generalization of the bivariate FGM distribution with two-dimensional extensions 

Carles M. Cuadras and Walter Díaz


#### Abstract

The Farlie-Gumbel-Morgenstern family of bivariate distributions with given marginals is frequently used in theory and applications and has been generalized in many ways. With the help of two auxiliary distributions, we propose another generalization and study its properties. After defining the dimension of a distribution as the cardinal of the set of canonical correlations, we prove that some well-known distributions are practically two-dimensional. Then we introduce an extended FGM family in two dimensions and study how to approximate any distribution to this family.


## 1. Introduction

Constructing dependence models by means of copulas has interest in statistics, probability, econometrics, informatics, insurance, biometry, physics, etc. A copula function is a bivariate cumulative distribution function (cdf) with uniform $[0,1]$ marginals, that captures the dependence properties of two random variables (r.v., in short). Many copulas and bivariate families of distributions have been described in Hutchinson and Lai (1991), Joe (1997), Drouet-Mari and Kotz (2001) and Nelsen (2006). Among others, the so-called Farlie-Gumbel-Morgenstern (FGM) bivariate family is frequently used in theory and applications. This motivated Huang and Kotz (1999), Amblard and Girard (2002), Nelsen et al. (1997), Rodríguez-Lallena and Úbeda-Flores (2004), Durante (2007), Cuadras (2009) and Cuadras et al. (2000) to study proper extensions.

Throughout this paper, $x, y$ in $H(x, y), F(x), G(y)$, as well as $u, v$ in $C(u, v)$, will be suppressed, unless it is strictly necessary. We write $H \in$

[^0]$\mathcal{F}(F, G)$, where $\mathcal{F}(F, G)$ is the Fréchet family of distribution functions with marginals $F, G$. A member of $\mathcal{F}(F, G)$ is the FGM family
$$
H_{\theta}=F G[1+\theta(1-F)(1-G)], \quad-1 \leq \theta \leq 1
$$

We show that this family is one-dimensional and propose a two-dimensional extension in a geometric sense.

As in Nelsen et al. (1997), who obtained second-order approximations, the goal of this paper is to prove that some well-known distributions can be approximated by the extended family proposed in Section 3.

## 2. Geometric dimensionality

Suppose that the Radon-Nikodým derivative $d H / d F d G$ exists. A global measure of dependence is Pearson contingency coefficient $\phi^{2}$ defined by

$$
\phi^{2}=\int_{a}^{b} \int_{c}^{d}\left(\frac{d H}{d F d G}-1\right)^{2} d F d G
$$

We have $\phi^{2} \geq 0$ and $\phi^{2}=0$ if and only if there is stochastic independence between $X$ and $Y$.

If $\phi^{2}$ is finite, then there exists the diagonal expansion (Lancaster, 1958)

$$
\begin{equation*}
d H=d F d G+\sum_{n \geq 1} \rho_{n} A_{n} B_{n} d F d G \tag{1}
\end{equation*}
$$

where $A_{n}, B_{n}$ are unitary functions in $L^{2}([a, b])$ and $L^{2}([c, d])$ on $F$ and $G$ respectively, in the sense that $E\left[A_{n}(X)\right]=E\left[B_{n}(Y)\right]=0$ and $E\left[A_{n}^{2}(X)\right]=$ $E\left[B_{n}^{2}(Y)\right]=1$. Then $A_{n}(X)$ and $B_{n}(Y)$ are the canonical variables. The sequence of canonical correlations is $\rho_{1} \geq \rho_{2} \geq \cdots \geq 0$. It can be proved from expansion (1) that Pearson contingency coefficient can be expressed in terms of this sequence:

$$
\phi^{2}=\sum_{n \geq 1} \rho_{n}^{2}
$$

The first canonical correlation $\rho_{1}=\sup \operatorname{corr}(A(X), B(Y))$, where corr means correlation coefficient, is the maximal correlation between a function of $X$ and a function of $Y$. The sequence of canonical correlations captures the full dependence between $X$ and $Y$ and the coefficient $\phi^{2}$ is an overall measure of dependence, often presented as the ratio $\phi^{2} /\left(1+\phi^{2}\right)$.

Definition 1. The geometric dimension of $H \in \mathcal{F}(F, G)$ such that the canonical expansion (1) exists, is the cardinal of the set $\left(\rho_{n}\right)$.

The cdf $H$ with dimension $N \leq \infty$, can be approximated by another cdf $H_{D}$ of smaller dimension $D<N$, where

$$
d H_{D}=d F d G+\sum_{n=1}^{D} \rho_{n} A_{n} B_{n} d F d G
$$

The "proportion of dependence" of $H$ accounted by $H_{D}$ is:

$$
P_{D}=\frac{\sum_{n=1}^{D} \rho_{n}^{2}}{\phi^{2}} .
$$

This proportion is common in some methods of multivariate analysis. For instance, it is used in correspondence analysis to measure the quality of the graphical representation of a contingency table with respect to the chi-square distance, see Greenacre (1984) and Cuadras et al. (2000).

On the other hand, from the canonical expansion (1), we can also consider the following nested family

$$
d H_{\lambda}=d F d G+\sum_{i=1}^{k} \lambda_{i} A_{i} B_{i} d F d G
$$

depending on the parameters $\lambda_{1}, \lambda_{2}, \ldots, \lambda_{k}$, where $\left|\lambda_{i}\right| \leq \rho_{i}$, for $i=1, \ldots, k$.
Next, let us consider the FGM copula. Recall that a bivariate copula is a function $C: I^{2} \rightarrow I$, with $I=[0,1]$. such that $C(u, 0)=C(0, v)=0$, $C(u, 1)=u, C(1, v)=v$, and for $0 \leq u_{1} \leq u_{2} \leq 1$ and $0 \leq v_{1} \leq v_{2} \leq 1$ satisfies $C\left(u_{2}, v_{2}\right)-C\left(u_{2}, v_{1}\right)-C\left(u_{1}, v_{2}\right)+C\left(u_{1}, v_{1}\right) \geq 0$.

Copulas are important because of Sklar's theorem. Let $H(x, y)$ be the bivariate cdf of $(X, Y)$, with univariate marginal distribution functions $F(x)$, $G(y)$ and supports $[a, b],[c, d]$, respectively. Then this bivariate cdf can be expressed as $H(x, y)=C(F(x), G(y))$, where $C$ is a copula related to $H$. Thus modeling copulas is an interesting task.

The related copula for the FGM distribution is

$$
C_{\theta}=u v[1+\theta(1-u)(1-v)], \quad-1 \leq \theta \leq 1 .
$$

The density is $c(u, v)=1+\theta(1-2 u)(1-2 v)$. Since the only canonical correlation is $\rho_{1}=|\theta| / 3$, clearly this distribution has geometric dimension one. In general, the dimension is higher.

We are interested in two-dimensional distributions as possible approximations to a given distribution. For example, for the following copulas (see Nelsen, 2006, Çelebioğlu, 1997, Cuadras, 2009):

$$
\begin{array}{lrr}
\text { Ali-Mikhail-Haq: } & u v /[1-\theta(1-u)(1-v)], & -1 \leq \theta \leq 1, \\
\text { Gumbel-Barnett: } & u v \exp (-\theta \ln u \ln v), & 0 \leq \theta \leq 1, \\
\text { Çelebioğlu-Cuadras: } & u v \exp [\theta(1-u)(1-v)], & -1 \leq \theta \leq 1 .
\end{array}
$$

The classical FGM family is the first order approximation from a Taylor expansion. However, the full dimension in these three copulas is countable (i.e., $\varkappa_{0}$ ), whereas the approximation using only two dimensions is very accurate, as the proportion of dependence is $P_{2} \simeq 0.99$.

The extension of the FGM family proposed below is two-dimensional in the above geometric sense.

## 3. Two-dimensional extension

Recall that the supports of $F, G$ are $[a, b],[c, d]$. Let $\Phi, \Psi$ be univariate distribution functions with the same supports. An extension of FGM is the bivariate family

$$
\begin{align*}
& H(x, y)=F(x) G(y)+\lambda_{1}[F(x)-\Phi(x)][G(y)-\Psi(y)] \\
& \quad+\lambda_{2}\left[\frac{1}{2} F^{2}(x)+\left(F_{\Phi}-\frac{1}{2}\right) F(x)-F_{\Phi}(x)\right]\left[\frac{1}{2} G^{2}(y)+\left(G_{\Psi}-\frac{1}{2}\right) G(y)-G_{\Psi}(y)\right], \tag{2}
\end{align*}
$$

where

$$
F_{\Phi}(x)=\int_{a}^{x} \Phi(t) d F(t), G_{\Psi}(y)=\int_{c}^{y} \Psi(t) d G(t), \quad F_{\Phi}=F_{\Phi}(b), G_{\Psi}=G_{\Psi}(d) .
$$

The density function (Lebesgue measure) is

$$
\begin{equation*}
h=f g+\lambda_{1} f\left(1-\varphi f^{-1}\right) g\left(1-\psi g^{-1}\right)+\lambda_{2} f(F-\Phi-\gamma) g(G-\Psi-\delta), \tag{3}
\end{equation*}
$$

where

$$
\gamma=\int_{a}^{b}(F-\Phi) d F=1 / 2-F_{\Phi}, \quad \delta=\int_{c}^{d}(G-\Psi) d G=1 / 2-G_{\Psi} .
$$

Of course, (2) reduces to FGM for $\Phi=F^{2}, \Psi=G^{2}$ and $\lambda_{2}=0$.
This family is diagonal and the canonical correlations are

$$
\begin{aligned}
\rho_{1}= & \lambda_{1} \sqrt{(\alpha-1)(\beta-1)}, \\
\rho_{2}= & \lambda_{2} \sqrt{1 / 3+F_{\Phi^{2}}-\left(1 / 2-F_{\Phi}\right)^{2}-2 \int_{a}^{b} F \Phi d F} \\
& \times \sqrt{1 / 3+G_{\Psi^{2}}-\left(1 / 2-G_{\Psi}\right)^{2}-2 \int_{a}^{b} G \Psi d G},
\end{aligned}
$$

where

$$
\alpha=\int_{a}^{b}\left(\frac{d \Phi}{d F}\right)^{2} d F, \quad \beta=\int_{c}^{d}\left(\frac{d \Psi}{d G}\right)^{2} d G
$$

and $F_{\Phi^{2}}=\int_{a}^{b} \Phi^{2} d F, G_{\Psi^{2}}=\int_{c}^{d} \Psi^{2} d G$.
To prove this write (2) as

$$
d H=d F d G+\lambda_{1} a_{1} d F b_{1} d G+\lambda_{2} a_{2} d F b_{2} d G
$$

where $a_{1}=1-d \Phi / d F, b_{1}=1-d \Psi / d G, a_{2}=(F-\Phi-\gamma)$ and $b_{2}=(G-\Psi-\delta)$.
It is readily proved that $E\left(a_{1}\right)=E\left(a_{2}\right)=0$ and $E\left(b_{1}\right)=E\left(b_{2}\right)=0$.
Note that

$$
\begin{aligned}
\int_{a}^{b}(1-d \Phi / d F)(F-\Phi-\gamma) d F & =0 \\
\int_{a}^{b}(F-\Phi-\gamma)^{2} d F & =1 / 3+F_{\Phi^{2}}-\left(1 / 2-F_{\Phi}\right)^{2}-2 \int_{a}^{b} F \Phi d F \\
\int_{c}^{d}(G-\Psi-\delta)^{2} d G & =1 / 3+G_{\Psi^{2}}-\left(1 / 2-G_{\Psi}\right)^{2}-2 \int_{c}^{d} G \Psi d G .
\end{aligned}
$$

Hence $E\left(a_{1} a_{2}\right)=\int_{a}^{b} a_{1} a_{2} d F=0$ and similarly $E\left(b_{1} b_{2}\right)=0$.

The other covariances are:

$$
\begin{aligned}
\operatorname{cov}\left(a_{1}, b_{1}\right) & =\int_{a}^{b} \int_{c}^{d} a_{1} b_{1}(d H-d F d G) \\
& =\lambda_{1} \int_{a}^{b} a_{1}^{2} d F \int_{c}^{d} b_{1}^{2} d G+\lambda_{2} \int_{a}^{b} a_{1} a_{2} d F \int_{c}^{d} b_{1} b_{2} d G \\
& =\lambda_{1}(\alpha-1)(\beta-1), \\
\operatorname{cov}\left(a_{1}, b_{2}\right) & =\lambda_{1} \int_{a}^{b} a_{1}^{2} d F \int_{c}^{d} b_{1} b_{2} d G+\lambda_{2} \int_{a}^{b} a_{1} a_{2} d F \int_{c}^{d} b_{2}^{2} d G \\
& =0,
\end{aligned}
$$

and similarly $\operatorname{cov}\left(a_{2}, b_{1}\right)=0$. Also

$$
\begin{aligned}
\operatorname{cov}\left(a_{2}, b_{2}\right)= & 0+\lambda_{2} \int_{a}^{b} a_{2}^{2} d F \int_{c}^{d} b_{2}^{2} d G \\
= & \lambda_{2}\left[1 / 3+F_{\Phi^{2}}-\left(1 / 2-F_{\Phi}\right)^{2}-2 \int_{a}^{b} F \Phi d F\right] \\
& \times\left[1 / 3+G_{\Psi^{2}}-\left(1 / 2-G_{\Psi}\right)^{2}-2 \int_{c}^{d} G \Psi d G\right] .
\end{aligned}
$$

The variances are

$$
\begin{aligned}
& E\left(a_{1}^{2}\right)=\alpha-1 \\
& E\left(a_{2}^{2}\right)=1 / 3+F_{\Phi^{2}}-\left(1 / 2-F_{\Phi}\right)^{2}-2 \int_{a}^{b} F \Phi d F, \text { etc. }
\end{aligned}
$$

Next, we study the corresponding copula. First note that $Q=\Phi \circ F^{-1}$ and $R=\Psi \circ G^{-1}$ are distribution functions with support in $[0,1]$.

Since $F_{\Phi}=\int_{a}^{b} \Phi d F=\int_{0}^{1} Q(t) d t=1-\mu_{Q}$, where $\mu_{Q}$ is the mean of the r.v. with cdf $Q$, and similarly $\mu_{R}$, the copula corresponding to (2) is

$$
\begin{align*}
C(u, v)= & u v+\lambda_{1}[u-Q(u)][v-R(v)] \\
& +\lambda_{2}\left[\frac{1}{2} u^{2}+\left(\frac{1}{2}-\mu_{Q}\right) u-\int_{0}^{u} Q(t) d t\right]\left[\frac{1}{2} v^{2}\right.  \tag{4}\\
& \left.+\left(\frac{1}{2}-\mu_{R}\right) v-\int_{0}^{v} R(t) d t\right] .
\end{align*}
$$

For $\Phi=F^{k}, \Psi=G^{l}$, family (2) has the copula

$$
\begin{aligned}
C(u, v) & =u v+\lambda_{1} u\left(1-u^{k-1}\right) v\left(1-v^{l-1}\right) \\
& +\lambda_{2}\left[\frac{1}{2} u^{2}-u(k-1) /(2(k+1))-u^{k+1} /(k+1)\right][\text { similar term in } v, l]
\end{aligned}
$$

and Spearman's correlation $\rho_{S}=12 \int_{I^{2}} C d u d v-3$ is

$$
\rho_{S}=\frac{(k-1)(l-1)}{(k+1)(l+1)}\left[3 \lambda_{1}+\frac{\lambda_{2}(k-2)(l-2)}{12(k+2)(l+2)}\right] .
$$

In particular, for $\Phi=F^{2}, \Psi=G^{2}$, the copula is

$$
\begin{equation*}
C=u v+u v(1-u)(1-v)\left[\lambda_{1}+\bar{\lambda}_{2}(1-2 u)(1-2 v)\right], \tag{5}
\end{equation*}
$$

where $\bar{\lambda}_{2}=\lambda_{2} / 36$.
Nelsen et al. (1997) studied the copula (5) as a member of the family of symmetric copulas with cubic sections in both $u$ and $v$, providing Spearman's rho, Kendall's tau, certain dependence properties and showing that they are second-degree Maclaurin approximations to members of the Frank and Plackett copula families.

## 4. Approximating a bivariate cdf for another one

Let us consider $H_{a} \in \mathcal{F}(F, G)$ and suppose that the canonical expansion $d H_{a}=d F d G+\sum_{n \geq 1} \rho_{n} A_{n} d F B_{n} d G$ exists, where $A_{n}, B_{n}$ are unitary canonical functions. Let $H_{t} \in \mathcal{F}(F, G)$ be the "true" cdf of two observable random variables $(X, Y)$. We are interested in approximating $H_{t}$ by means of a finite linear combination of canonical functions obtained from $H_{a}$ :

$$
d H_{t} \simeq d F d G+\sum_{i=1}^{k} \lambda_{i} A_{i} d F B_{i} d G
$$

In a more precise way, we seek for the approximation

$$
\frac{d H_{t}}{d F d G} \simeq 1+\sum_{i=1}^{k} \lambda_{i} A_{i} B_{i}
$$

where $\lambda_{1}, \ldots, \lambda_{k}$ are real coefficients such that

$$
\begin{equation*}
\int_{a}^{b} \int_{c}^{d}\left(\frac{d H_{t}-d F d G}{d F d G}-\sum_{i=1}^{k} \lambda_{i} A_{i} B_{i}\right)^{2} d F d G \tag{6}
\end{equation*}
$$

is minimized. If the densities $h_{t}, f, g$ exist, then $h_{t}$ is approximated by

$$
\widehat{h}_{t}=f g\left(1+\sum_{i=1}^{k} \lambda_{i} A_{i} B_{i}\right)
$$

Theorem 1. Suppose $(X, Y) \sim H_{t} \in \mathcal{F}(F, G)$. The coefficients minimizing (6) are $\lambda_{i}=\rho_{i}$, where

$$
\rho_{i}=\operatorname{corr}\left(A_{i}(X), B_{i}(Y)\right), i=1, \ldots, k
$$

Proof. Write $z=\left(d H_{t}-d F d G\right) / d F d G$. Then

$$
\begin{aligned}
\int_{a}^{b} \int_{c}^{d}\left(z-\sum_{i=1}^{k} \lambda_{i} A_{i} B_{i}\right)^{2} d F d G & =\phi_{t}^{2}+\int_{a}^{b} \int_{c}^{d} \sum_{i=1}^{k} \lambda_{i}^{2} A_{i}^{2} B_{i}^{2} d F d G \\
& -2 \int_{a}^{b} \int_{c}^{d}\left(\sum_{i=1}^{k} \lambda_{i} A_{i} B_{i}\right)\left(d H_{t}-d F d G\right) \\
& +\sum_{i \neq j=1}^{k} \lambda_{i} \lambda_{j} \int_{a}^{b} A_{i} d F \int_{c}^{d} B_{j} d G \\
& =\phi_{t}^{2}+\sum_{i=1}^{k} \lambda_{i}^{2}-2 \sum_{i=1}^{k} \lambda_{i} \rho_{i}
\end{aligned}
$$

where $\phi_{t}^{2}$ is the Pearson contingency coefficient. Taking the partial derivative with respect to $\lambda_{i}$ on the left hand side of this equation, and equating to zero, we obtain $\lambda_{i}=\rho_{i}, i=1, \ldots, k$.

Note that each $\rho_{i}$ is the correlation between the canonical variables $A_{i}, B_{i}$ obtained from $H_{a}$, but the correlation is taken with respect to the "true" cdf $H_{t}$. That is, the observed data used in computing $\rho_{i}$ comes from $(X, Y) \sim$ $H_{t}$. This result is useful when the canonical functions of $H_{a}$ are known, as it occurs in the above generalized FGM.

## 5. Examples

5.1. Ali-Mikhail-Haq copula. We first study the approximation of the Ali-Mikhail-Haq (AMH) copula (Nelsen, 2006)

$$
C_{t}=\frac{u v}{1-\theta(1-u)(1-v)}, \quad-1 \leq \theta \leq 1,
$$

to the copula (5). The density for this generalized FGM copula is

$$
c=1+\frac{\lambda_{1}}{3} \sqrt{3}(1-2 u) \sqrt{3}(1-2 v)+\frac{\bar{\lambda}_{2}}{5} \sqrt{5}\left(6 u^{2}-6 u+1\right) \sqrt{5}\left(6 v^{2}-6 v+1\right) .
$$

The canonical functions are $A_{1}=\sqrt{3}(1-2 u), A_{2}=\sqrt{5}\left(6 u^{2}-6 u+1\right)$ and similarly $B_{1}, B_{2}$. To be sure that $c$ is a density, the canonical correlations should belong to the region $\mathcal{R}=\left\{\left(\rho_{1}, \rho_{2}\right) \mid c \geq 0\right\}$, see Figure 1 .


Figure 1. Region of the correlations (parameters) for which the density is positive.

We should calculate the correlations

$$
\rho_{1}=\operatorname{corr}(U, V), \quad \rho_{2}=\operatorname{corr}\left(U^{2}-U, V^{2}-V\right),
$$

where $(U, V) \sim C$. By using the following formula

$$
\operatorname{cov}(\nu(U), \xi(V))=\int_{I^{2}}[C(u, v)-u v] d \nu(u) d \xi(v)
$$

which provides the covariance between functions of $U$ and $V$ (Cuadras, 2002), we have

$$
\begin{aligned}
& \rho_{1}=12 \int_{I^{2}} C(u, v) d u d v-3 \\
& \rho_{2}=180 \int_{I^{2}} C(u, v)(4 u v-2 u-2 v+1) d u d v-5 .
\end{aligned}
$$

Of course, $\rho_{1}$ is Spearman's rho correlation.
Expanding the AMH copula $C_{t}$, which plays the role of "true" copula,

$$
\frac{u v}{1-\theta(1-u)(1-v)}=u v\left[1+\sum_{i=1}^{\infty} \theta^{i}(1-u)^{i}(1-v)^{i}\right],
$$

we obtain

$$
\rho_{1}=12 \sum_{i=1}^{\infty} \theta^{i} B(2, i+1)^{2},
$$

and

$$
\rho_{2}=180 \sum_{i=1}^{\infty} \theta^{i}\left[4 B(3, i+1)^{2}-4 B(2, i+1) B(3, i+1)+B(2, i+1)^{2}\right],
$$

where $B(\cdot, \cdot)$ is the beta function. All pairs $\left(\rho_{1}, \rho_{2}\right)$ belong to $\mathcal{R}$ and the AMH copula can be approximated by

$$
C_{2}=u v+\rho_{1} 3 u(1-u) v(1-v)+\rho_{2} 5\left(2 u^{3}-3 u^{2}+u\right)\left(2 v^{3}-3 v^{2}+v\right) .
$$

Kendall's tau $\tau=4 \int_{I^{2}} C d C-1$ for copula $C_{2}$ is (Nelsen et al., 1997):

$$
\tau\left(C_{2}\right)=\frac{2}{3} \rho_{1}+\frac{2}{15} \rho_{1} \rho_{2} .
$$

See Nelsen (2006) for the exact expressions for $\tau$ and $\rho_{S}$ in the AMH family.
If $M=\min \{u, v\}$ and $W=\max \{u+v-1,0\}$ are the Fréchet-Hoeffding upper and lower bounds, two measures of fit are

$$
\begin{aligned}
& \eta_{1}=\max _{u, v \in I}\left|C_{t}(u, v)-C_{2}(u, v)\right|, \\
& \eta_{2}=D\left(C_{t}, C_{2}\right) / D(M, W),
\end{aligned}
$$

where

$$
D\left(C_{t}, C_{2}\right)=\int_{I^{2}}\left(C_{t}-C_{2}\right)^{2} d u d v
$$

which satisfies $D\left(C_{t}, C_{2}\right)<D(M, W)=1 / 24$. Thus $0<\eta_{i}<1, i=1,2$.
Table 1. Canonical correlations and fit for the AMH copula.

| $\theta$ | $\rho_{1}$ | $\rho_{2}$ | $\eta_{1}$ | $\eta_{2}$ | $\rho_{S}(A M H)$ | $\tau(A M H)$ | $\tau\left(C_{2}\right)$ |
| ---: | ---: | :---: | ---: | :---: | ---: | ---: | ---: |
| -1 | -0.2711 | 0.0217 | 0.0055 | 0.0002 | -0.2710 | -0.1817 | -0.1815 |
| -.5 | -0.1489 | 0.0080 | 0.0017 | 0.0000 | -0.1489 | -0.0995 | -0.0995 |
| .5 | 0.1924 | 0.0223 | 0.0032 | 0.0001 | 0.1924 | 0.1288 | 0.1286 |
| 1 | 0.4783 | 0.2323 | 0.0261 | 0.0029 | 0.4784 | 0.3333 | 0.3335 |

Table 1 reports a numerical illustration, showing that the fit is quite good, practically preserving Spearman's rho and Kendall's tau.
5.2. Clayton-Oakes copula. We consider the Clayton-Oakes copula (Nelsen, 2006):

$$
C_{t}=\left[\max \left(u^{-\theta}+v^{-\theta}-1,0\right)\right]^{-1 / \theta}, \quad-1 \leq \theta<\infty .
$$

The correlations $\rho_{1}$ and $\rho_{2}$ have been obtained numerically. However, now the results can provide FGM approximations which are not copulas, i.e., the density $c(u, v)$ is negative for some values of $u, v$. Then we take $\left(\rho_{1}^{*}, \rho_{2}^{*}\right) \in \mathcal{R}$ with the smallest Euclidean distance to ( $\rho_{1}, \rho_{2}$ ). Thus the Clayton-Oakes copula can be approximated by

$$
C_{2}=u v+\rho_{1}^{*} 3 u(1-u) v(1-v)+\rho_{2}^{*} 5\left(2 u^{3}-3 u^{2}+u\right)\left(2 v^{3}-3 v^{2}+v\right) .
$$

Table 2. Canonical correlations and fit for the ClaytonOakes copula.

| $\theta$ | $\rho_{1}$ | $\rho_{2}$ | $\rho_{1}^{*}$ | $\rho_{2}^{*}$ | $\eta_{1}$ | $\eta_{2}$ | $\lambda_{L}(C O)$ |
| ---: | ---: | ---: | :---: | :---: | :---: | :---: | ---: |
| -1 | -1 | 1 | -0.5085 | 0.2950 | 0.0625 | 0.0060 | - |
| -0.5 | -0.4667 | 0.1997 | -0.4665 | 0.1995 | 0.0263 | 0.0036 | - |
| 0.5 | 0.2950 | 0.1150 | 0.2950 | 0.1150 | 0.0162 | 0.0014 | 0.2500 |
| 1 | 0.4784 | 0.2337 | 0.4785 | 0.2335 | 0.0261 | 0.0029 | 0.5000 |
| 2 | 0.6822 | 0.4104 | 0.5495 | 0.2620 | 0.0349 | 0.0035 | 0.7071 |
| 5 | 0.8846 | 0.6809 | 0.5190 | 0.2875 | 0.0385 | 0.0030 | 0.8706 |
| 10 | 0.9582 | 0.8470 | 0.5190 | 0.2875 | 0.0387 | 0.0037 | 0.9330 |

The fit is acceptably good for intermediate (positive) values of the parameter $\theta$, see Table 2. However, there are differences in the upper and lower tail dependence parameters $\lambda_{L}, \lambda_{U}$. In both families $\lambda_{U}=0$, whereas $\lambda_{L}=0$ for $C_{2}$. The Clayton-Oakes copula has $\lambda_{L}=2^{-1 / \theta}$ if $\theta \geq 0$, but there is no mass in the lower-left corner if $\theta<0$, so $\lambda_{L}$ does not exist if the parameter is negative.

## References

Amblard, C., and Girard, S. (2002), Symmetry and dependence properties within a semiparametric family of bivariate copulas, J. Nonparametr. Stat. 14, 715-727.
Çelebioğlu, S. (1997), A way of generating comprehensive copulas, J. Inst. Sci. Tech. Gazi Univ. 10, 57-61.
Cuadras, C. M. (2002), On the covariance between functions, J. Multivariate Anal. 81, 19-27.
Cuadras, C. M. (2009), Constructing copula functions with weighted geometric means, J. Statist. Plann. Inference 139, 3766-3772.

Cuadras, C. M., Fortiana, J., and Greenacre, M. J. (2000), Continuous extensions of matrix formulations in Correspondence Analysis, with applications to the FGM family of distributions; In: Innovations in Multivariate Statistical Analysis, R. D. H. Heijmans, D. S. G. Pollock and A. Satorra (eds.), Kluwer Academic Publishers, Dordrecht, pp. 101-116.
Druet-Mari, D., and Kotz, S. (2001), Correlation and Dependence, Imperial College Press, London.
Durante, F. (2007), A new family of symmetric bivariate copulas, C. R. Acad. Sci. Paris Ser. I 344, 195-198.
Greenacre, M. J. (1984), Theory and Applications of Correspondence Analysis, Academic Press, London.
Huang, J. S., and Kotz, S. (1999), Modifications of the Farlie-Gumbel-Morgenstern distributions. A tough hill to climb, Metrika 49, 135-145.
Hutchinson, T.P., and Lai, C.D. (1991), The Engineering Statistician's Guide to Continuous Bivariate Distributions, Rumsby Scientific Publishing, Adelaide.
Joe, H. (1997), Multivariate Models and Dependence Concepts, Chapman and Hall, London.
Lancaster, H. O. (1958), The structure of bivariate distributions, Ann. Math. Statist. 29, 719-736.
Nelsen, R. B. (2006), An Introduction to Copulas. 2nd Ed., Springer, New York.
Nelsen, R. B., Quesada-Molina, J. J., and Rodríguez-Lallena, J. A. (1997), Bivariate copulas with cubic sections, J. Nonparametr. Stat. 7, 205-220
Rodríguez-Lallena, J. A., and Úbeda-Flores, M. (2004), A new class of bivariate copulas, Statist. Probab. Lett. 66, 315-325.

Department of statistics, Faculty of Biology, University of Barcelona, Barcelona, Spain

E-mail address: ccuadras@ub.edu
Department of Mathematics, Faculty of Economics, University of Antioquia, Medellin, Colombia

E-mail address: wdiaz@udea.edu.co


[^0]:    Received October 19, 2011.
    2010 Mathematics Subject Classification. Primary 62H20; Secondary 60E05.
    Key words and phrases. Farlie-Gumbel-Morgenstern distribution, bivariate copulas, stochastic dependence, Pearson contingency coefficient, dimension of a distribution.

    Work supported in part by grant Ministerio de Educación y Ciencia, MTM2008-00642.

