Global Journal of Science Frontier Research: F
Mathematics and Decision Sciences
Volume 17 Issue 1 Version 1.0 Year 2017
Type : Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 2249-4626 & Print ISSN: 0975-5896

# On Characterizing Generalized Cambanis Family of Bivariate Distributions

By N. Unnikrishnan Nair, Johny Scaria & Sithara Mohan
*Nirmala College Muvattupuzha*

*Abstract-* In this work we present characterizations of a generalized version of Cambanis family of bivariate distributions. This family contains extensions of the Farlie-Gumbel-Morgenstern system as special cases. The characterizations are by properties of P(X>Y), regression functions and E(XjX > Y) which were found to be useful in many applications.

*Keywords: cambanis family, FGM system, characterization, regression functions, conditional expectations.*

*GJSFR-F Classification: MSC 2010: 62H10*

ONCHARACTERIZINGGENERALIZEDCAMBANISFAMILYOFBIVARIATEDISTRIBUTIONS

*Strictly as per the compliance and regulations of :*

# On Characterizing Generalized Cambanis Family of Bivariate Distributions

N. Unnikrishnan Nair [α], Johny Scaria [σ] & Sithara Mohan [ρ]

*Abstract-* In this work we present characterizations of a generalized version of Cambanis family of bivariate distributions. This family contains extensions of the Farlie-Gumbel-Morgenstern system as special cases. The characterizations are by properties of P(X>Y), regression functions and E(XjX > Y) which were found to be useful in many applications.

*Keywords: cambanis family, FGM system, characterization, regression functions, conditional expectations.*

## I. Introduction

In the present work we consider a generalized version of a family of bivariate distributions specified by an absolutely continuous distribution function of the form

$$F(x,y) = F_X(x)F_Y(y)[1 + \alpha_1 A(F_X(x)) + \alpha_2 B(F_Y(y)) + \alpha_3 A(F_X(x))B(F_Y(y))], \qquad (1.1)$$

of a random vector $(X,Y)$. The kernels $A(.)$ and $B(.)$ in the model are differentiable over [0,1], satisfies the conditions

$$A(1) = 0 = B(1) \quad \text{and} \quad A(0) = 1 = B(0)$$

and are chosen in such a way that (1.1) is a distribution function with absolutely continuous marginal distributions. The family subsumes several distributions of potential interest in distribution theory as well as in modelling problems associated with other disciplines. These include the extended Farlie-Gumbel-Morgenstern (FGM) system

$$F_1(x,y) = F_X(x)F_Y(y)[1 + \alpha_3 A(F_X(x))B(F_Y(y))] \qquad (1.2)$$

considered in Bairamov and Kotz (2002) and several particular cases obtained by giving different forms for $A(.)$ and $B(.)$ like the classical FGM when $A(F)$=1-$F_X(x)$ and $B(F)$=1-$F_Y(y)$ and others discussed in Huang and Kotz (1984, 1999), Bairamov et al. (2001), Amblard and Girard (2009) and Carles et al. (2012) and the references therein.

A somewhat different special case of (1.1) is the Cambanis (1977) model specified by

$$\begin{aligned} F_2(x,y) &= F_X(x)F_Y(y)[1 + \alpha_1(1 - F_X(x)) + \alpha_2(1 - F_Y(y)) \\ &+ \alpha_3(1 - F_X(x))(1 - F_Y(y))], \end{aligned} \qquad (1.3)$$

*Author α: Cochin University of Science and Technology. e-mail: unnikrishnannair4@gmail.com,*
*Author σ ρ: Department of Statistics, Nirmala College Muvattupuzha- 686661, India. e-mails: johny.kakkattil@gmail.com, sithara.mohan3@gmail.com*

$-\infty < x, y < \infty$, $(1+\alpha_1+\alpha_2+\alpha_3) \geq 0$, $(1+\alpha_1-\alpha_2-\alpha_3) \geq 0$, $(1-\alpha_1+\alpha_2-\alpha_3) \geq 0$ and $(1-\alpha_1-\alpha_2+\alpha_3) \geq 0$.

The major difference here is that unlike the FGM, the marginals of (1.3) are not $F_X(x)$ and $F_Y(y)$ but they are uniquely determined by $F_X(x)$ and $F_Y(y)$. The distributional aspects, dependence structure and applications of (1.3) are discussed in Nair et al. (2016). Various forms of $A(.)$ and $B(.)$ used to extend the FGM can also be applied in (1.1) to generate new families of bivariate models. In view of the wide variety of distributions generated from (1.1), it is important to study its properties.

The objective of the present work is to attempt characterizations of $F(x,y)$ through properties of $P(X < Y)$, $E(X|X > Y)$ and the regressions functions of $(X, Y)$. The former is well known in stress-strength modelling. When $X$ represents the stress and $Y$, the strength, $P(X < Y)$ indicates the reliability of the material, while the latter is the average stress at which it exceeds the strength. More details about these aspects are discussed in the next section where the characterizations are considered.

Global Journal of Science Frontier Research ( F ) Volume XVII Issue I Version I Year 2017



Notes

## II. Characterizations

First we consider $P(X < Y)$ when $(X, Y)$ follows the distribution (1.1). Apart from the stress-strength interpretation $P(X < Y)$ is suited to other variables in different fields of study such as quality control, genetics, psychology, economics and clinical trial. For details we refer to Kotz et al. (2003).

When bivariate distributions are used to model stress-strength data some sort of dependence is assumed between $X$ and $Y$. Among various bivariate cases considered in literature in this context, one of particular interest to the present work is Domma and Giordano (2013) in which the FGM copula is considered.

Let $f(x,y)$, $f_X(x)$ and $f_Y(y)$ denote the probability density functions of $(X, Y)$, $X$ and $Y$ respectively. Then

$$\begin{aligned} \alpha = P(X < Y) &= \int_{-\infty}^{\infty} P(X < y|Y = y) f_Y(y) dy \\ &= \int_{-\infty}^{\infty} \int_{0}^{y} f(\nu, y) d\nu dy. \end{aligned} \tag{2.1}$$

Since $P(X < Y){=}P(F_X(x) < F_Y(y))$, for calculation purposes it is enough to consider the uniform distribution in (1.1),

$$F_4(x, y) = xy[1 + \alpha_1 A(x) + \alpha_2 B(y) + \alpha_3 A(x)B(y)],$$

and the corresponding probability density function

$$f_4(x, y) = 1 + \alpha_1 \frac{d}{dx} xA(x) + \alpha_2 \frac{d}{dy} yB(y) + \alpha_3 \frac{d}{dx} xA(x) \frac{d}{dy} yB(y). \tag{2.2}$$

In that case

$$\alpha = \int_0^1 \int_0^y f_4(\nu, y) d\nu dy. \tag{2.3}$$

Following Bairamov and Kotz (2002), we have the first theorem concerning the nature of $P(X<Y)$ in the model (1.1).

*Theorem 2.1*

Let $(X,Y)$ be a continuous random variable with distribution function (1.1). Then $P(X<Y)$ is independent of $\alpha_3$ if and only if $A(x)$=$B(x)$ for all $x$ in [0,1] provided that

Notes

$$A(x)B^{'}(x)-A^{'}(x)B(x) \text{ is } \geq 0 \text{ or } \leq 0 \text{ for all } x. \tag{2.4}$$

In this case,

$$P(X<Y)=\frac{1}{2}+(\alpha_1-\alpha_2)\int_0^1 xA(x)dx. \tag{2.5}$$

*Proof:* From (2.2) and (2.3) using $A(1)$=0=$B(1)$,

$$\begin{aligned} \alpha &= \int_0^1 [y+\alpha_1 yA(y)+\alpha_2 y\frac{d}{dy}yB(y)+\alpha_3 yA(y)\{B(y)+yB^{'}(y)\}]dy \\ &= \frac{1}{2}+\alpha_1\int_0^1 yA(y)dy-\alpha_2\int_0^1 yB(y)dy+\alpha_3\int_0^1\{yA(y)B(y)+y^2A(y)B^{'}(y)\}dy. \end{aligned} \tag{2.6}$$

Since

$$\frac{d}{dy}y^2A(y)B(y)=2yA(y)B(y)+y^2\{A(y)B^{'}(y)+A^{'}(y)B(y)\}$$

(2.6) takes the form,

$$\begin{aligned} \alpha &= \frac{1}{2}+\alpha_1\int_0^1 yA(y)dy-\alpha_2\int_0^1 yB(y)dy+\frac{\alpha_3}{2}\int_0^1\frac{d}{dy}y^2A(y)B(y)dy \\ &+ \frac{\alpha_3}{2}\int_0^1 y^2\{A(y)B^{'}(y)-A^{'}(y)B(y)\}dy. \end{aligned} \tag{2.7}$$

Since $\alpha$ is independent of $\alpha_3$ and $\int_0^1\frac{d}{dy}y^2A(y)B(y)dy$=0, one must have

$$\int_0^1 y^2\{A(y)B^{'}(y)-A^{'}(y)B(y)\}dy=0,$$

which means that $A(y)B^{'}(y)-A^{'}(y)B(y)$=0 by virtue of (2.4). Hence $A(y)$=$C\ B(y)$ in which $C$=1 from $A(0)$=$B(0)$=1. Thus $A(x)=B(x)$ and (2.5) holds. Conversely if $A(x)=B(x)$, (2.7) shows that $\alpha$ does not contain $\alpha_3$.

*Remark 2.1*

For the theorem to be true, the representation (1.1) must yield a distribution function for some $A(.)$ and $B(.)$. In the case of the Cambanis family (1.3), $A(x)=1-x$, $B(y)=1-y$ so that we have a distribution function belonging to (1.1) satisfying (2.4). Further, $A(x)=B(x)$ and so

$$P(X<Y) = \frac{1}{2}+(\alpha_1-\alpha_2)\int_0^1 x(1-x)dx = \frac{1}{2}+\frac{\alpha_1-\alpha_2}{6}.$$

*Remark 2.2*

When $\alpha_3$=0, the conditions on the parameters $\alpha_1$ and $\alpha_2$ are satisfied by a convex set containing $|\alpha_i| \leq \frac{1}{3}$, $i$=2. Thus $P(X<Y)$ lies in the range $(\frac{7}{18}, \frac{11}{18})$ which is more flexible than that of FGM for which $P(X<Y)=\frac{1}{2}$. The flexibility can be further increased with other choices of $A(x)$, for example $A(x)=1-x^2=B(x)$ as in the Huang-Kotz modification.

Notes

From Theorem 2.1, a modification of Theorem in Bairamov and Kotz (2002) is evident.

*Theorem 2.2*

Let *(X, Y)* be a continuous random vector with distribution function (1.1) with $\alpha_1$=$\alpha_2$. Then $P(X<Y)=\frac{1}{2}$ if and only if $A(x)=B(x)$ for all $x$, provided that (2.4) is satisfied.

Our second characterization is based on the conditional expectations $E(X|X>Y)$ and $E(Y|Y>X)$. In a reliability frame work these have interpretations and applications. Suppose that *(X, Y)* represents the lifetimes of a two-component system. Then $E(X|X>Y)$ and $E(Y|Y>X)$ denotes the average lifetime of the longest living component and is an important information about the system. The probability density function of $X$ given $X>Y$ is

$$f(x|X>Y)=\frac{1}{P(X>Y)}\int_{-\infty}^{x} f(x,\nu)d\nu$$

and hence

$$P(X>Y)\ \ E(X|X>Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{x} xf(x,\nu)d\nu dx.$$

Specializing to the uniform case

$$P(X>Y)\ \ E(X|X>Y)=\int_0^1\int_0^x xf_4(x,\nu)d\nu dx. \tag{2.8}$$

Similarly

$$P(Y>X)\ \ E(Y|Y>X)=\int_0^1\int_0^y yf_4(\nu,y)d\nu dy. \tag{2.9}$$

*Theorem 2.3*

Let *(X, Y)* be a random vector specified by $f_4(x,y)$ satisfying $A(1)=0=B(1)$ and $A(0)=1=B(0)$. Then $P(X>Y)$ $E(X|X>Y)$ - $P(Y>X)$ $E(Y|Y>X)$ is independent of $\alpha_3$ if and only if $A(x)=B(x)$ for all $x$ provided that (2.4) is satisfied.

*Proof:* From equations (2.8) and (2.2),

$$P(X>Y)E(X|X>Y)=\int_0^1\int_0^x x[1+\alpha_1\frac{d}{dx}xA(x)+\alpha_2\frac{d}{d\nu}\nu B(\nu)+\alpha_3\frac{d}{dx}xA(x)\frac{d}{d\nu}\nu B(\nu)]d\nu dx$$

$$\begin{aligned} &= \int_0^1 x[x+\alpha_1 x\frac{d}{dx}xA(x)+\alpha_2 xB(x)+\alpha_3(\frac{d}{dx}xA(x))(xB(x))]dx \\ &= \frac{1}{3}-2\alpha_1\int_0^1 x^2A(x)dx+\alpha_2\int_0^1 x^2B(x)dx \\ &- \alpha_3\int_0^1[2x^2A(x)B(x)+x^3A(x)B^{'}(x)]dx. \end{aligned}$$

Notes

Similarly from (2.9),

$$\begin{aligned} P(Y>X)E(Y|Y>X) &= \frac{1}{3}+\alpha_1\int_0^1 x^2A(x)dx-2\alpha_2\int_0^1 x^2B(x)dx \\ &- \alpha_3\int_0^1[2x^2A(x)B(x)+x^3A^{'}(x)B(x)]dx. \end{aligned}$$

Thus

$$\begin{aligned} P(X>Y)E(X|X>Y)-P(Y>X)E(Y|Y>X) = -3\alpha_1\int_0^1 x^2A(x)dx+3\alpha_2\int_0^1 x^2B(x)dx \\ - \alpha_3\int_0^1 x^3(A(x)B^{'}(x)-A^{'}(x)B(x))dx. \quad (2.10) \end{aligned}$$

Now assume that $A(x)=B(x)$. Then obviously (2.10) is independent of $\alpha_3$. Conversely (2.10) is independent of $\alpha_3$, then $A(x)=B(x)$ using the arguments in Theorem 2.1 and the proof is complete.

Notice that in the above case,

$$\begin{aligned} P(X>Y)E(X|X>Y) &= \frac{1}{3}-2\alpha_1\int_0^1 x^2A(x)dx+\alpha_2\int_0^1 x^2A(x)dx \\ &- \alpha_3\int_0^1 x^2A(x)[xA^{'}(x)+2A(x)]dx. \end{aligned}$$

The last integral is

$$\begin{aligned} \int_0^1[x^3A(x)A^{'}(x)+2x^2A^2(x)]dx &= \int_0^1[2x^2A^2(x)+\frac{1}{2}x^3\frac{d}{dx}A^2(x)-\frac{3}{2}x^2A^2(x)]dx \\ &= \frac{1}{2}\int_0^1 x^2A^2(x)dx. \end{aligned}$$

Thus

$$P(X>Y)E(X|X>Y)=\frac{1}{3}+(\alpha_2-2\alpha_1)\int_0^1 x^2A(x)dx-\frac{\alpha_3}{2}\int_0^1 x^2A^2(x)dx.$$

and finally

$$E(X|X>Y)=\frac{\frac{1}{3}+(\alpha_2-2\alpha_1)\int_0^1 x^2A(x)dx-\frac{\alpha_3}{2}\int_0^1 x^2A^2(x)dx}{\frac{1}{2}-(\alpha_1-\alpha_2)\int_0^1 xA(x)dx}$$

using (2.5). As an example, for the Cambanis family, with uniform marginals,

$$F(x,y)=xy[1+\alpha_1(1-x)+\alpha_2(1-y)+\alpha_3(1-x)(1-y)],\qquad 0\leq x,y\leq 1,$$

the conditions of the Theorem 2.3 are satisfied. Accordingly

$$P(X>Y)=\frac{1}{2}+\frac{\alpha_2-\alpha_1}{6}$$

and

$$E(X|X>Y)=\frac{20+5(\alpha_2-2\alpha_1)-\alpha_3}{30+10(\alpha_2-\alpha_1)}.$$

Further in the light of the above discussions Theorem 2 in Bairamov and Kotz (2002) can be modified as follows.

*Theorem 2.4*

Let $(X,Y)$ be a bivariate random vector with distribution function

$$F(x,y)=xy[1+\alpha_1A(x)+\alpha_2B(y)+\alpha_3A(x)B(y)] \tag{2.11}$$

satisfying $A(1)=0=B(1)$, $A(0)=1=B(0)$ and $A(x)B^{'}(x)-A^{'}(x)B(x)\geq 0$ or $\leq 0$ for all $x$ in [0,1], Then

$$P(X>Y)E(X|X>Y)=P(Y>X)E(Y|X>Y)$$

if and only if $\alpha_1$=$\alpha_2$ and $A(x)=B(x)$ for all $x$.

*Proof:* The result follows from the fact that for (2.11)

$$\begin{aligned}P(X>Y)E(X|X>Y) &= \frac{1}{3}-\alpha_1\int_0^1 x^2A(x)dx-\frac{\alpha_3}{2}\int_0^1 x^2A^2(x)dx\\ &= P(Y>X)E(Y|X>Y)\end{aligned}$$

and Theorem 2.3.

The problem of characterizing bivariate distributions through their regression functions have received considerable attention, see for example, Rao and Sinha (1988), Bryc (2012). A traditional approach in statistical modelling is to select a flexible family of distributions and then to find a member of the family that is appropriate for the given data. One characteristic of the family amenable to easy verification is the regression function. The forms of the regression functions $b_1(x)=E(Y|X=x)$ and $b_2(y)=E(X|Y=y)$ can be detected from the observations and the model that conforms with it is a reasonable choice for the data.

Notes

We show that $(b_1(x), b_2(y))$ determines the model (2.2) uniquely and provide example of members that have simple functional forms for them.

*Theorem 2.5*

Let *(X, Y)* be continuous random vector with distribution specified by (2.2). Then the regression functions $b_1(x)$ and $b_2(y)$ uniquely determine the distribution of *(X, Y)*.

Notes

*Proof:* For the distribution (2.2), the conditional distribution of $Y$ given $X{=}x$ is

$$f(y|x) = \frac{1 + \alpha_1 \frac{d}{dx} xA(x) + \alpha_2 \frac{d}{dy} yB(y) + \alpha_3 (\frac{d}{dx} xA(x))(\frac{d}{dy} yB(y))}{1 + \alpha_1 \frac{d}{dx} xA(x)}.$$

Some direct calculations give

$$\begin{aligned} b_1(x) &= \int_0^1 y f(y|x) dy \\ &= \frac{1}{2} - \frac{[\alpha_2 + \alpha_3 \frac{d}{dx} xA(x)] \int_0^1 yB(y)dy}{1 + \alpha_1 \frac{d}{dx} xA(x)} \end{aligned}$$

Solving

$$\frac{d}{dx} xA(x) = \frac{d_1(x) - \alpha_2}{\alpha_3 - \alpha_1 d_1(x)}, \quad d_1(x) = \frac{\frac{1}{2} - b_1(x)}{\int_0^1 yB(y)dy}$$

and hence

$$A(x) = \frac{1}{x} \int_0^x \frac{d_1(t) - \alpha_2}{\alpha_3 - \alpha_1 d_1(t)} dt. \tag{2.12}$$

Similarly

$$B(y) = \frac{1}{y} \int_0^y \frac{d_2(t) - \alpha_1}{\alpha_3 - \alpha_2 d_2(t)} dt, \quad d_2(y) = \frac{\frac{1}{2} - b_2(y)}{\int_0^1 xA(x)dx}. \tag{2.13}$$

Equations (2.12) and (2.13) determine the distribution (2.2).

*Example 2.1*

$$\text{Let} \qquad d_1(x) = \frac{\alpha_2 + \alpha_3(1 - 2x)}{1 + \alpha_1(1 - 2x)} \qquad \text{and} \quad d_2(y) = \frac{\alpha_1 + \alpha_3(1 - 2y)}{1 + \alpha_2(1 - 2y)}$$

Then from (2.12),

$$A(x) = \frac{1}{x} \int_0^x (1 - 2t)dt = 1 - x,$$

and similarly $A(y) = 1 - y$. Thus the distribution is given by

$$F(x, y) = xy[1 + \alpha_1(1 - x) + \alpha_2(1 - y) + \alpha_3(1 - x)(1 - y)], \quad 0 \le x, y \le 1.$$

*Example 2.2*

Let $d_1(x) = \dfrac{\alpha_2 + \alpha_3(1-x)(1-3x)}{1+\alpha_1(1-x)(1-3x)}$ and $d_2(y) = \dfrac{\alpha_1 + \alpha_3(1-y)(1-3y)}{1+\alpha_2(1-y)(1-3y)}$

Then

$$A(x) = \frac{1}{x}\int_0^x (1-t)(1-3t)dt = (1-x)^2,$$

and similarly $A(y) = (1-y)^2$, giving

$$F(x,y) = xy[1+\alpha_1(1-x)^2 + \alpha_2(1-y)^2 + \alpha_3(1-x)^2(1-y)^2], \quad 0 \le x, y \le 1.$$

We conclude this work by noting that Theorems 2.1 through 2.4 extends the work of Bairamov and Kotz (2002) to a more general family of bivariate distributions and Theorem 2.5 provides a new result that helps in identifying a distribution belonging to the general family we have presented.

## References Références Referencias


1. Amblard, C., and Girard, S., (2009). A new extension of FGM copulas, *Metrika,* 70, 1-17.
2. Bairamov, I., and Kotz, S., (2002). Dependence structure and symmetry of Huang-Kotz FGM distributions and their extensions, *Metrika,* 56, 55-72.
3. Bairamov, I., Kotz, S., and Bekci, M., (2001). New generalized Farlie-Gumbel-Morgenstern distributions and concomitants of order statistics, *Journal of Applied Statistics,* 28, 5, 521-536.
4. Bryc, W., (2012). *The Normal distributions; Characterizations and Applications,* Springer-Verlag.
5. Cambanis, S., (1977). Some properties and generalizations of Multivariate Eyraud-Gumbel-Morgenstern Distributions, *Journal of Multivariate Analysis,* 7, 551-559.
6. Carles, M., Cudras, C. M., and Walter Daz, (2012). Another generaization of the bivariate FGM distribution with two- dimensional extensions, *Acta Et Commentationes Universitatis Tartuinsis De Mathematica,* 16, 1.
7. Domma, F., and Giordano, S., (2013). A copula-based approach to account for dependence in stress-strength models, *Statistical Papers,* 54, 3, 807826.
8. Huang, J. S., and Kotz, S., (1984). Correlation structure in iterated Farlie-Gumbel-Morgenstern distributions, *Biometrika* 71, 3, 633-636.
9. Huang, J. S., and Kotz, S., (1999). Modifications of the Farlie-Gumbel-Morgenstern distributions. A tough hill to climb. *Metrika,* 49, 135-145.
10. Kotz, S., Lumelskii, Y., Pensky, M. (2003). The Stress-strength Model and its Generalizations, World Scientific Publising, Singapore.
11. Nair, N. U., Scaria, J., and Mohan, S. (2016). The Cambanis family of bivariate distributions: Properties and applications, *Statistica,* LXXVI, 2, 169-184.
12. Rao, B. V, and Sinha, B. K. (1988). A characterization of Dirichlet distributions, *Journal of Multivariate Analysis,* 25, 25-30.


Notes