Bivariate Beta Distributions and Beyond

Jacobs, Rianne
University of Pretoria, Department of Statistics
cnr. Lynnwood Road and Roper Street
Hatfield 0083, South Africa
E-mail: rianne.jacobs@up.ac.za

Bekker, Andriette
University of Pretoria, Department of Statistics
cnr. Lynnwood Road and Roper Street
Hatfield 0083, South Africa
E-mail: andriette.bekker@up.ac.za

Human, Schalk
University of Pretoria, Department of Statistics
cnr. Lynnwood Road and Roper Street
Hatfield 0083, South Africa
E-mail: schalk.human@up.ac.za

Introduction

Bivariate beta distributions make up a major part of statistical distribution theory. They form part of the Dirichlet family of distributions, but have become an important family of distributions in their own right. Many bivariate beta distributions have been derived out of an application or as extensions or generalizations from other well-known bivariate beta distributions. Since bivariate beta distributions are used in a wide variety of applications such as Bayesian statistics and reliability theory, there exist a huge amount of research in the development and derivation of new bivariate beta distributions. In this research, we considered the Kummer type beta distributions and more specifically, in this paper, the *bivariate Kummer-beta type IV* distribution. This distribution is an extension of the *bivariate beta type IV* distribution (or Jones' model) which has its roots in the model proposed by Libby and Novick (1982). It was, however, more explicitly derived by Jones (2001) and Olkin and Liu (2003). Furthermore, it is a special case of the model proposed by Sarabia and Castillo (2006). Although, in this paper we explicitly derive the *bivariate Kummer-beta type IV* distribution, it is also a special case of the *bimatrix variate Kummer-beta type IV distribution* defined by Bekker et al. (2010).

Kummer-type distributions form an integral part of statistical distribution theory and a number of these distributions have been proposed. In the univariate case, for example, the *Kummer-gamma* and the *Kummer-beta are* introduced by Armero and Bayarri (1997) and Ng and Kotz (1995). The latter also proposed and studied the *multivariate Kummer-gamma* and *multivariate Kummer-beta* families of distributions. In the matrix variate case, there is the work by Gupta et al (2001), Nagar and Gupta (2002) and Nagar and Cardeño (2001). These authors proposed and studied matrix variate generalizations of the multivariate Kummer-beta and the multivariate Kummer-gamma families of distributions, which are called the *matrix variate Kummer-Dirichlet* (or the *matrix variate Kummer-beta*) and the *matrix variate Kummer-gamma* distributions, respectively. It should be noted that these Kummer distributions get their name from the fact that their normalizing constants are all defined in terms of one of the two so-called Kummer functions (see e.g. Rainville, 1960, p 124-126).

The rest of this paper is structured as follows. First we derive the joint probability density function (pdf), $g(x_1, x_2)$, of the bivariate Kummer-beta type IV distribution. Then, the pdf's of the marginal distributions, $m(x_i)$ for $i = 1, 2$, the pdf's of the conditional distributions, $h(x_i|x_j)$ for $i, j = 1, 2$ and $i \neq j$, and the product moment, $E(X_1^r X_2^s)$, of the bivariate Kummer-beta type IV are derived. We investigate the influence of the shape parameter, ψ, of the bivariate Kummer-beta type IV distribution. Thus, its effect on the correlation between the components of this distribution as well as its effect on the pdf of (X_1, X_2) and the marginal of X_1. An example of a possible application will be presented.

The Bivariate Kummer-Beta type IV distribution

In this section we derive the pdf of the new bivariate Kummer-beta type IV distribution, $g(x_1, x_2)$, the pdf's of the marginal distributions, $m(x_i)$ for $i = 1, 2$, the pdf's of the conditional distributions, $h(x_i|x_j)$ for $i, j = 1, 2$ and $i \neq j$, and the product moment, $E(X_1^r X_2^s)$.

The pdf, $g(x_1, x_2)$, is derived by extending the kernel of the *Jones' bivariate beta distribution,* which has pdf

$$f(x_1, x_2) = C x_1^{a-1} x_2^{b-1} (1 - x_1)^{b+c-1} (1 - x_2)^{a+c-1} (1 - x_1 x_2)^{-(a+b+c)}$$

for $0 \leq x_1, x_2 \leq 1$, $a, b, c > 0$ and where $C^{-1} = B(a, b, c) = \dfrac{\Gamma(a)\Gamma(b)\Gamma(c)}{\Gamma(a+b+c)}$ denotes the normalizing constant (Jones, 2001). Clearly, X_1 and X_2 each have a standard beta type I distribution, i.e. $X_1 \sim Beta(a, c)$ and $X_2 \sim Beta(b, c)$ over $0 \leq x_1, x_2 \leq 1$.

Theorem 1

The pdf of the bivariate Kummer-beta type IV distribution is given by

$$g(x_1, x_2) = K x_1^{a-1} x_2^{b-1} (1 - x_1)^{b+c-1} (1 - x_2)^{a+c-1} (1 - x_1 x_2)^{-(a+b+c)} e^{-\psi(x_1+x_2)} \tag{1}$$

where $0 \leq x_1, x_2 \leq 1$, $a, b, c > 0$, $-\infty < \psi < \infty$, *and the normalizing constant* K *is given by*

$$K^{-1} = \sum_{k=0}^{\infty} \frac{(a+b+c)_k}{k!} \frac{{}_1F_1(a+k, a+b+c+k; -\psi)\, {}_1F_1(b+k, a+b+c+k; -\psi)}{(B(a+c, b+k)\, B(b+c, a+k))^{-1}}. \tag{2}$$

where $B(.)$ *denotes the beta function and* ${}_1F_1(.)$ *denotes the confluent hypergeometric function (Gradshteyn, 2007, Section 9.2, p 1022). This distribution is denoted as* $(X_1, X_2) \sim BKB^{IV}(a, b, c, \psi)$.

Proof

Define the bivariate Kummer-beta type IV distribution as

$$g(x_1, x_2) = K x_1^{a-1} x_2^{b-1} (1 - x_1)^{b+c-1} (1 - x_2)^{a+c-1} (1 - x_1 x_2)^{-(a+b+c)} e^{-\psi(x_1+x_2)}.$$

The normalizing constant is obtained by integrating over the full support of $g(x_1, x_2)$:

$$\begin{aligned}
& K^{-1} \\
&= \sum_{k=0}^{\infty} \frac{(a+b+c)_k}{k!} \int_0^1 x_2^{b+k-1} (1-x_2)^{a+c-1} e^{-\psi x_2} \int_0^1 x_1^{a+k-1} (1-x_1)^{b+c-1} e^{-\psi x_1} dx_1 dx_2 \\
&= \sum_{k=0}^{\infty} \frac{(a+b+c)_k}{k!} \int_0^1 x_2^{b+k-1} (1-x_2)^{a+c-1} e^{-\psi x_2} \frac{{}_1F_1(a+k; a+b+c+k; -\psi)}{(B(b+c, a+k))^{-1}} dx_2 \\
&= \sum_{k=0}^{\infty} \frac{(a+b+c)_k}{k!} \frac{{}_1F_1(a+k; a+b+c+k; -\psi)\, {}_1F_1(b+k; a+b+c+k; -\psi)}{(B(a+c, b+k)\, B(b+c, a+k))^{-1}}
\end{aligned}$$

The above result is obtained by expanding the term $(1 - x_1 x_2)^{-(a+b+c)}$ as a power series and then using the integral representation of the confluent hypergeometric function, ${}_1F_1(.)$ (Gradshteyn, 2007, Eq 3.383, p 347). ■

Note, the bivariate Kummer-beta type IV distribution may also be obtained by substituting $p = 1$ in the pdf of the bimatrix variate Kummer-beta type IV distribution defined by Bekker et al. (2010) (Section 5.3).

Theorem 2

If $(X_1, X_2) \sim BKB^{IV}(a, b, c, \psi)$, *the marginal pdf of* X_1 *is given by*

$$m(x_1) = Kx_1^{a-1}(1-x_1)^{b+c-1}e^{-\psi x_1}B(b, a+c)\,\Phi_1(b, a+b+c, a+b+c, \psi, x_1) \tag{3}$$

$$= Kx_1^{a-1}(1-x_1)^{b+c-1}e^{-\psi x_1}\sum_{k=0}^{\infty}\frac{(a+b+c)_k\,x_1^k}{k!}\frac{{}_1F_1(b+k; a+b+c+k; -\psi)}{(B(b+k, a+c))^{-1}} \tag{4}$$

where $0 \le x_1, x_2 \le 1$, $a, b, c > 0$, $B(.)$ *denotes the beta function,* $\Phi_1(.)$ *denotes confluent hypergeometric series of two variables (Gradshteyn, 2007, Eq 9.261, p 1031) and* K *is defined in (2).*

Proof

Using (1), the first representation of $m(x_1)$ given in (3), is obtained by using the integral representation of the confluent hypergeometric series of two variables, $\Phi_1(.)$ (Gradshteyn, 2007, Eq 3.385, p 349). The second representation of $m(x_1)$ given in (4), is obtained by expanding the term $(1-x_1x_2)^{-(a+b+c)}$ as a power series and using the integral representation of the confluent hypergeometric function, ${}_1F_1(.)$. ■

Equation (4) is more useful (in the sense that it is easier to implement and/or program) in computer packages such as Mathematica when we want to graph the pdf $m(x_i)$ for $i = 1, 2$.

The expression for the conditional pdf of $X_2|X_1$ can easily be obtained by using the definition $h(x_2|x_1) = \frac{g(x_1,x_2)}{m(x_1)}$. Note that the marginal pdf of X_2 and the conditional pdf of $X_1|X_2$ are obtained by substituting x_2 for x_1 in the respective expressions and interchanging the parameters a and b.

Theorem 3

If $(X_1, X_2) \sim BKB^{IV}(a, b, c, \psi)$, *the product moment,* $E(X_1^r X_2^s)$, *equals*

$$K\sum_{k=0}^{\infty}\frac{(a+b+c)_k}{k!}\frac{{}_1F_1(a+k+r, a+b+c+k+r; -\psi)\,{}_1F_1(b+k+s, a+b+c+k+s; -\psi)}{(B(a+c, b+k+s)\,B(b+c, a+k+r))^{-1}} \tag{5}$$

$$= (A(a, b, c, 0, 0))^{-1} \times A(a, b, c, s, r)$$

where

$$A(a, b, c, s, r) = \sum_{k=0}^{\infty}\frac{(a+b+c)_k}{k!}\frac{{}_1F_1(a+k+r, a+b+c+k+r; -\psi)\,{}_1F_1(b+k+s, a+b+c+k+s; -\psi)}{(B(a+c, b+k+s)\,B(b+c, a+k+r))^{-1}}$$

and $A(a, b, c, 0, 0) = K$ *defined in (2).*

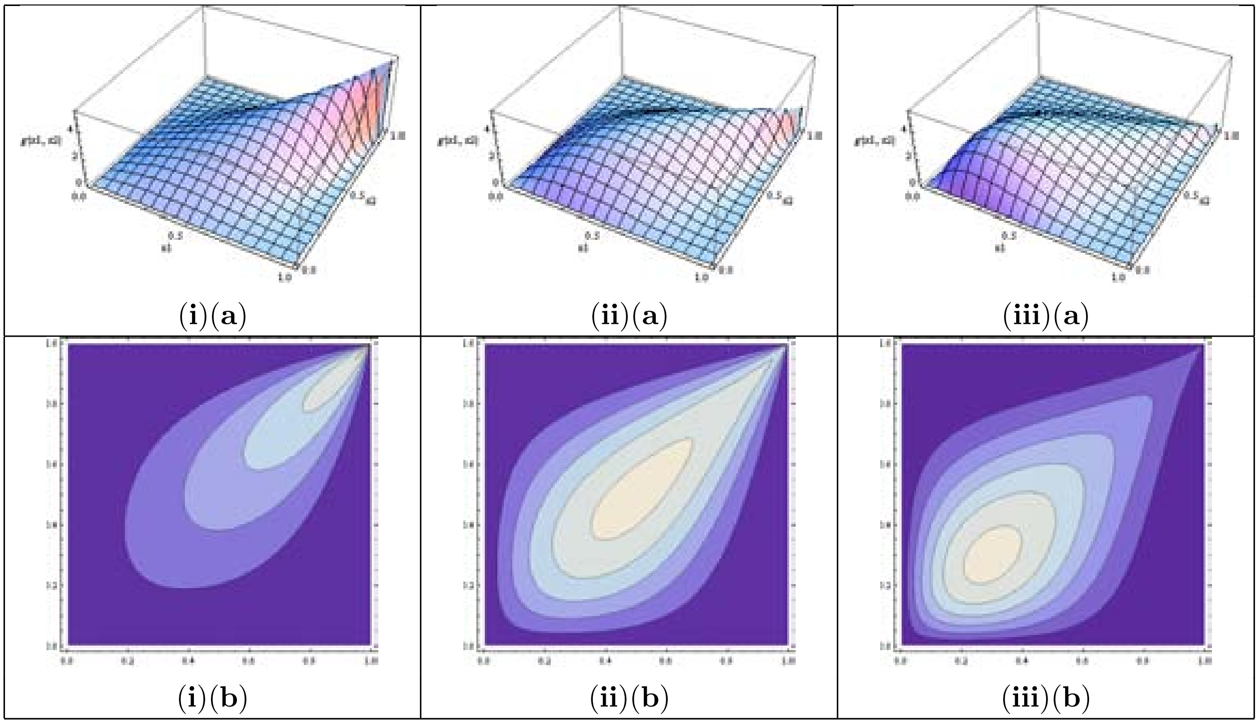
Proof

From (1), expanding the term $(1-x_1x_2)^{-(a+b+c)}$ as a power series and using the integral representation of the confluent hypergeometric function, ${}_1F_1(.)$, we obtain (5). ■

Shape Analysis

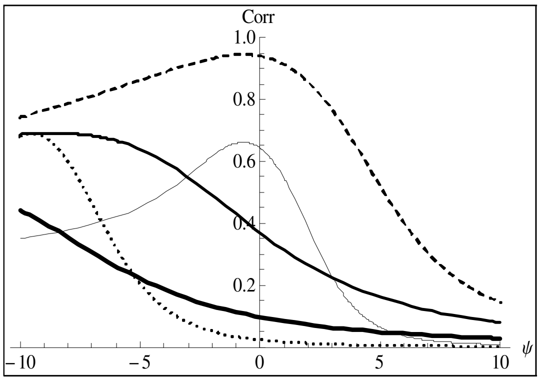
Figure 1 illustrates the effect of the parameter ψ for $\psi = -1.1, 0$ and 1.1 on the bivariate Kummer-beta type IV pdf (see (1)) for $a = b = c = 2$. We also show the contour plots for easy comparison. We note that the domain of the graphs in Figure 1 are all $\mathbb{R}^2 : [0,1] \times [0,1]$. Considering the graphs in Figure 1 we see that the parameter ψ shifts the bell of the density. For negative ψ, the bell of the density moves away from the origin, while for positive ψ, the bell of the density moves towards the origin.

Figure 1: Bivariate Kummer-beta type IV density function (a) and contour plots(b)



(i) ψ=−1.1; (ii) ψ=0; (iii) ψ=1.1. The parameter values are a=b=c=2.

Figure 2: Correlation between dependent components



The five curves are: thick solid line $a = 1$, $b = 1$, $c = 10$; medium solid line $a = 2$, $b = 4$, $c = 5$; thin solid line $a = 0.5$, $b = 0.9$, $c = 0.1$; dashed line $a = 2.5$, $b = 4$, $c = 0.5$; dotted line $a = 0.1$, $b = 0.5$, $c = 5$.

Figure 2 illustrates the effect of the parameter $\psi \in [-10, 10]$ on the correlation between X_1 and X_2 using (5). We see that: (i) the parameter ψ can both increase and decrease the correlations for different values of a, b and c and (ii) we can obtain a wide range of correlations between 0 and 1 - depending on the values of a, b, c and ψ. The correlation cannot be negative because the central bivariate Kummer-beta type IV distribution is totally positive of order 2.

When looking at the marginal density, we found that the parameter ψ introduces skewness to the symmetric cases of the Jones' bivariate beta distribution and changes the kurtosis of the assymetric cases.

An Application

The stress-strength model in the context of reliability is a well-known application of various bivariate beta distributions. This model describes the life of a component with a random strength X_2 subjected to a random stress X_1. The reliability of a component can be expressed as $P(X_1 < X_2)$ or $P(\frac{X_1}{X_2} < 1) = P(R < 1)$. One method that is frequently used in obtaining the reliability is that of obtaining the distribution of the ratio of the two components stress, X_1, and strength, X_2. In this section we derive the exact expression for the pdf of the ratio of the correlated components of the bivariate Kummer-beta type IV distribution i.e. $R = \frac{X_1}{X_2}$, in terms of Meijer's G-function (see Mathai, 1993, Definition 2.1, p 60) using the Mellin transform and the inverse Mellin transform (see Mathai, 1993, Definition 1.8, p 23).

Theorem 4

If $(X_1, X_2) \sim BKB^{IV}(a, b, c, \psi)$ *and we let* $R = \dfrac{X_1}{X_2}$, *then the pdf of* R *is given by*

$$w(r) = K\Gamma(a+c)\Gamma(b+c)\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{j=0}^{\infty}(a+b+c)_k\frac{(-\psi)^{j+l}}{j!k!l!}G_{2,2}^{1,1}\left(r\,\middle|\,\begin{matrix}\alpha_1,\alpha_2\\ \beta_1,\beta_2\end{matrix}\right) \quad \text{for } r \geq 0 \tag{6}$$

where

$$-\alpha_1 = b+k+j \quad \alpha_2 = a+b+c+k+l-1 \quad \beta_1 = a+k+l-1 \quad -\beta_2 = a+b+c+k+j$$

with K *defined in* (2).

Proof

Setting $r = h-1$ and $s = 1-h$ in (5) and using the series representation of the confluent hypergeometric function, ${}_1F_1(.)$, we obtain an expression for the Mellin transform of $g(x_1, x_2)$ as

$$\begin{aligned}
M_g(h) &= E(R^{h-1}) = E\left(\left(\frac{X_1}{X_2}\right)^{h-1}\right) \\
&= (A(a,b,c,0,0))^{-1} \times A(a,b,c,h-1,1-h) \\
&= K\frac{\Gamma(b+c)\Gamma(a+c)}{\Gamma(a+b+c)}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{j=0}^{\infty}\Gamma(a+b+c+k)\frac{\Gamma(1-\alpha_1-h)\Gamma(\beta_1+h)}{\Gamma(1-\beta_2-h)\Gamma(\alpha_2+h)}\frac{(-\psi)^{j+l}}{j!k!l!}
\end{aligned}$$

with

$$-\alpha_1 = b+k+j \quad \alpha_2 = a+b+c+k+l-1 \quad \beta_1 = a+k+l-1 \quad -\beta_2 = a+b+c+k+j$$

Using the inverse Mellin transform, the density of R given in (6) follows. ■

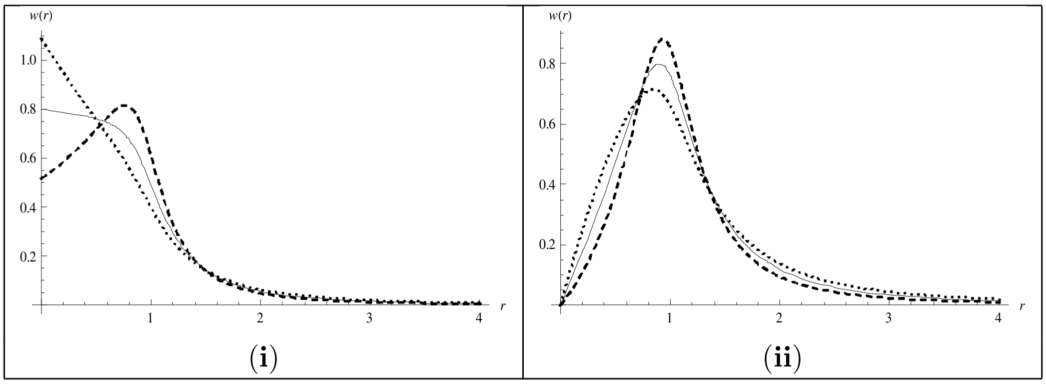
Using (6), we can now calculate the reliability for various combinations of parameter values. Table 1 provides the reliability of the bivariate Kummer type IV distribution for $\psi = -1.1$, 0, 1.1 with parameters $a = 1$, $b = c = 2$ and $a = b = c = 2$. For example, when $a = 1, b = c = 2$ and $\psi = -1.1$, we see that $P(R < 1) = 0.68471$; this implies that the probability that the component will function satisfactorily is 0.68471; or, in other words, the component will fail with probability 0.31529.

Tabel 1: Some reliability values

a	b	c	ψ	$P(R < 1)$
1	2	2	−1.1	0.68471
1	2	2	0	0.74904
1	2	2	1.1	0.75217
2	2	2	−1.1	0.42325
2	2	2	0	0.49782
2	2	2	1.1	0.48213

Figure 3 illustrates the shape of the density of $R = \frac{X_1}{X_2}$ (see Equation 6) for the case $a = 1, b = c = 2$ and $a = b = c = 2$ for different values of ψ. The domain for these graphs is $\mathbb{R} : [0, \infty]$.

Figure 3: The pdf of the ratio



(i) $a = 1, b = c = 2$ and (ii) $a = b = c = 2$. The three curves in each panel are: dashed line $\psi = -1.1$, solid line $\psi = 0$, dotted line $\psi = 1.1$.

Conclusion

In this paper we introduced, derived and studied the new bivariate Kummer-beta type IV distribution. It was shown that the densities can take different shapes and, therefore, the bivariate Kummer-beta type IV distribution can be used to analyse skewed bivariate data sets. The expressions derived in this paper are a valuable contribution to the existing literature on *Continuous Bivariate Distributions* as the comprehensive work by Balakrishnan and Lai (2009). We also obtained an exact expression for the density function of the ratio of the components of the bivariate distribution which is useful in reliability. For future research, another possible application can be explored in the Beta-Binomial context.

REFERENCES

1. Armero, C. and Bayarri, M.J. (1997). A Bayesian analysis of queueing system with unlimited service, *Journal of Statistical Planning and Inference*, **58**, 241-261.

2. Balakrishnan, N. and Lai, C.D. (2009). *Continuous Bivariate Distributions*, 2nd Edition, Springer, New York.

3. Bekker, A., Roux, J.J.J., Ehlers, R. and Arashi, M. (2010). Bimatrix variate beta type IV distribution:relation to Wilks's statistic and bimatrix variate Kummer-beta type IV distribution, *Communications and Statistics-Theory and methods* (in press).

4. Gradshteyn, I.S. and Ryzhik, I.M. (2007). *Table of Integrals, Series, and Products*, Academic Press, Amsterdam.

5. Gupta, A.K., Cardeño, L. and Nagar, D.K. (2001). Matrix-variate Kummer-Dirichlet distributions, *J. Appl. Math.*, **1(**3), 117-139.

6. Johnson, N.L., Kemp, A.W. and Kotz, S. (2005). *Univariate Discrete Distributions, 3rd Edition*, Wiley, Hoboken.

7. Jones, M.C. (2001). Multivariate t and the beta distributions associated with the multivariate F distribution, *Metrika*, **54**, 215-231.

8. Libby, D.L. and Novick, R.E. (1982). Multivariate generalized beta distributions with applications to utility assessment, *Journal of Educational Statistics*, **7**(4), 271-294.

9. Mathai, A.M. (1993). *A Handbook of Generalized Special Functions for Statistical and Physical Sciences*, Clarendon Press, Oxford.

10. Nagar, D.K. and Cardeño, L. (2001). Matrix variate Kummer-Gamma distributions, *Random Oper. and Stoch. Equ.*, **9**(3), 207-218.

11. Nagar, D.K. and Gupta, A.K. (2002). Matrix-variate Kummer-Beta distribution, *J. Austral. Math. Soc.*, **73**, 11-25.

12. Ng, K.W. and Kotz, S. (1995). Kummer-gamma and Kummer-beta univariate and multivariate distributions, *Research report*, **84**, The University of Hong Kong.

13. Olkin, I. and Liu, R. (2003). A Bivariate beta distribution, *Statistics and Probability Letters*, **62**, 407-412.

14. Rainville, E.D. (1960). *Special Functions*, Macmillan, New York.

15. Sarabia, J.M. and Castillo, E. (2006). Bivariate Distributions Based on the Generalized Three-Parameter Beta Distribution, In *Advances in Distribution Theory, Order Statistics, and Inference*, (Ed., N. Balakrishnan, E. Castillo, J.M. Sarabia), pp. 85-110, Birkhäuser, Boston.