A TRUNCATED BIVARIATE INVERTED DIRICHLET DISTRIBUTION

Saralees Nadarajah

1. INTRODUCTION

The need for truncated bivariate distributions with heavy tails is frequently encountered in many applied areas, see, for example, Lien (1985, 1986). A popular bivariate distribution with heavy tails is the bivariate inverted dirichlet distribution. Because of the heavy tails, this distribution does not possess finite moments of all orders. In this note, we overcome this weakness by introducing a truncated version. The bivariate inverted dirichlet distribution is given by the joint probability density function (pdf):

$$g(x, y) = \frac{\Gamma(a+b+c) x^{a-1} y^{b-1}}{\Gamma(a)\Gamma(b)\Gamma(c)(1+x+y)^{a+b+c}} \tag{1}$$

for $x > 0, y > 0$, $a > 0$, $b > 0$ and $c > 0$. The corresponding joint cumulative distribution function (cdf) is:

$$G(x, y) = \frac{\Gamma(a+b+c) x^{a} y^{b}}{\Gamma(a+1)\Gamma(b+1)\Gamma(c)(1+x+y)^{a+b+c}} \times F_2\left(a+b+c; 1, 1; a+1, b+1; \frac{x}{1+x+y}, \frac{y}{1+x+y}\right), \tag{2}$$

where F_2 denotes the Appell hypergeometric function of the second kind defined by

$$F_2(a, b, b'; c, c'; z, \eta) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{(a)_{k+l} (b)_k (b')_l z^k \eta^l}{(c)_k (c')_l k! l!}$$

for $|z| + |\eta| < 1$, where $(e)_k = e(e+1)\cdots(e+k-1)$ denotes the ascending factorial. The truncated version is given by the pdf:

$$f(x,y)=\frac{\Gamma(a+b+c)x^{a-1}y^{b-1}}{\Omega\Gamma(a)\Gamma(b)\Gamma(c)(1+x+y)^{a+b+c}} \tag{3}$$

for $0<B\leq x\leq A<\infty$ and $0<D\leq y\leq C<\infty$, where $\Omega(A,B,C,D,a,b,c)=G(A,C)-G(A,D)-G(B,C)+G(B,D)$. The cdf associated with (3) is:

$$F(x,y)=\frac{1}{\Omega}\{G(x,y)-G(x,D)-G(B,y)+G(B,D)\}. \tag{4}$$

The corresponding marginal cdfs and marginal pdfs are

$$F_X(x)=\frac{1}{\Omega}\{G(x,C)-G(x,D)-G(B,C)+G(B,D)\},$$

$$F_Y(y)=\frac{1}{\Omega}\{G(A,y)-G(A,D)-G(B,y)+G(B,D)\},$$

$$f_X(x)=\frac{\Gamma(a+b+c)x^{a-1}}{\Omega\Gamma(a)\Gamma(b+1)\Gamma(c)(1+x)^{a+b+c}}\left\{C^b\ {}_2F_1\left(b,a+b+c;b+1;-\frac{C}{1+x}\right)\right.$$
$$\left.-D^b\ {}_2F_1\left(b,a+b+c;b+1;-\frac{D}{1+x}\right)\right\}$$

and

$$f_Y(y)=\frac{\Gamma(a+b+c)y^{b-1}}{\Omega\Gamma(a+1)\Gamma(b)\Gamma(c)(1+y)^{a+b+c}}\left\{A^a\ {}_2F_1\left(a,a+b+c;a+1;-\frac{A}{1+y}\right)\right.$$
$$\left.-B^a\ {}_2F_1\left(a,a+b+c;a+1;-\frac{B}{1+y}\right)\right\},$$

where ${}_2F_1$ denotes the Gauss hypergeometric function defined by

$${}_2F_1(a,b;c;x)=\sum_{k=0}^{\infty}\frac{(a)_k(b)_k}{(c)_k}\frac{x^k}{k!}.$$

We refer to (3) as the truncated bivariate inverted dirichlet distribution. Because (3) is defined over a finite interval, the truncated bivariate inverted dirichlet distribution has all it moments. Thus, (3) may prove to be a better model for certain practical situations than one based on just the bivariate inverted dirichlet distribution. Below, we discuss one such situation.

The bivariate inverted dirichlet distribution has received applications in many areas, including biological analyses, clinical trials, stochastic modelling of decreasing failure rate life components, study of labor turnover, queueing theory, and reliability (see, for example, Nayak (1987) and Lee and Gross (1991)). For data from these areas, there is no reason to believe that empirical moments of any order should be infinite. Thus, the choice of the bivariate inverted dirichlet distribution as a model is unrealistic since its product moments $E(X^m Y^n)$ are not finite for all m and n. The alternative given by (3) will be a more appropriate model for the kind of data mentioned. The choice of the limits, A, B, C and D could be easily based on historical records. For example, to model the dependence between the exchange rate data of the United Kingdom Pound and the Canadian Dollar (as compared to the United States Dollar) for last 100 years, one could choose $D=0$, $C=10$, $B=0$ and $A=10$ (see http://www.gobalfindata.com).

The rest of this note is organized as follows. Explicit expressions for the moments of (3) are derived in Sections 2 and 3. Maximum likelihood estimators and the associated Fisher information matrix are derived in Section 4. The calculations use the special functions defined above as well as the hypergeometric function defined by

$$ {}_3F_2(a,b,c;d,e;x) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k (c)_k}{(d)_k (e)_k} \frac{x^k}{k!}. $$

The properties of the above special functions can be found in Prudnikov *et al.* (1986) and Gradshteyn and Ryzhik (2000).

2. MOMENTS

We derive two representations for the product moment $E(X^m Y^n)$. Theorem 1 expresses it as an infinite sum of incomplete beta functions while the expression given by Theorem 2 is in terms of the hypergeometric functions.

Theorem 1 If (X,Y) has the joint pdf (3) then

$$ E\left(X^m Y^n\right) = \frac{\Gamma(a+b+c)}{\Gamma(a)\Gamma(b)\Gamma(c)} \{H(A,C) - G(A,D) - G(B,C) + G(B,D)\} \tag{5} $$

for $m \geq 1$ and $n \geq 1$, where

$$ H(P,Q) = \frac{Q^{n+b}}{n+b} \sum_{k=0}^{\infty} \frac{(n+b)_k (a+b+c)_k (-Q)^k}{(n+b+1)_k k!} B_{P/(1+P)}(m+a, k+b+c-m). $$

Proof: Using (3), one can write

$$
\begin{aligned}
E(X^m Y^n) &= \frac{\Gamma(a+b+c)}{\Gamma(a)\Gamma(b)\Gamma(c)} \int_B^A \int_D^C \frac{x^{m+a-1} y^{n+b-1}}{(1+x+y)^{a+b+c}} dydx \\
&= \frac{\Gamma(a+b+c)}{\Gamma(a)\Gamma(b)\Gamma(c)} \left\{ \int_0^A \int_0^C \frac{x^{m+a-1} y^{n+b-1}}{(1+x+y)^{a+b+c}} dydx - \int_0^A \int_0^D \frac{x^{m+a-1} y^{n+b-1}}{(1+x+y)^{a+b+c}} dydx \right. \\
&\quad \left. - \int_0^B \int_0^C \frac{x^{m+a-1} y^{n+b-1}}{(1+x+y)^{a+b+c}} dydx + \int_0^B \int_0^D \frac{x^{m+a-1} y^{n+b-1}}{(1+x+y)^{a+b+c}} dydx \right\} \\
&= \frac{\Gamma(a+b+c)}{\Gamma(a)\Gamma(b)\Gamma(c)} \{ H(A,C) - H(A,D) - H(B,C) + H(B,D) \},
\end{aligned}
$$

where

$$
H(P,Q) = \int_0^P \int_0^Q \frac{x^{m+a-1} y^{n+b-1}}{(1+x+y)^{a+b+c}} dydx. \tag{6}
$$

Application of equation (2.2.6.1) in Prudnikov *et al.* (1986, volume 1) shows that the double integral $H(P,Q)$ can be expressed as

$$
\begin{aligned}
H(P,Q) &= \int_0^P x^{m+a-1} \int_0^Q \frac{y^{b+n-1}}{(1+x+y)^{a+b+c}} dydx \\
&= \frac{Q^{n+b}}{n+b} \int_0^P x^{m+a-1} (1+x)^{-a-b-c} {}_2F_1\left(n+b, a+b+c; n+b+1; -\frac{Q}{1+x} \right) dx \\
&= \frac{Q^{n+b}}{n+b} \int_0^P x^{m+a-1} (1+x)^{-a-b-c} \sum_{k=0}^{\infty} \frac{(n+b)_k (a+b+c)_k}{(n+b+1)_k k!} \left(-\frac{Q}{1+x} \right)^k dx \\
&= \frac{Q^{n+b}}{n+b} \sum_{k=0}^{\infty} \frac{(n+b)_k (a+b+c)_k (-Q)^k}{(n+b+1)_k k!} \int_0^P x^{m+a-1} (1+\theta x)^{-a-b-c-k} dx \\
&= \frac{Q^{n+b}}{n+b} \sum_{k=0}^{\infty} \frac{(n+b)_k (a+b+c)_k (-Q)^k}{(n+b+1)_k k!} \int_0^{P/(1+P)} y^{m+a-1} (1-y)^{k+b+c-m-1} dy \\
&= \frac{Q^{n+b}}{n+b} \sum_{k=0}^{\infty} \frac{(n+b)_k (a+b+c)_k (-Q)^k}{(n+b+1)_k k!} B_{P/(1+P)}(m+a, k+b+c-m),
\end{aligned}
$$

where we have set $y = x/(1+x)$ and used the definitions of the incomplete beta and the Gauss hypergeometric functions.

Theorem 2 If (X,Y) has the joint pdf (3) and if $a \geq 1$ and $b \geq 1$ are integers then $E(X^m Y^n)$ is given by (5) for $m \geq 1$ and $n \geq 1$, where

$$H(P,Q)=\sum_{k=0}^{b+n-1}\binom{b+n-1}{k}\frac{(-1)^{b+n-1-k}}{k-a-b-c+1}\sum_{l=0}^{b+n-k-1}\binom{b+n-k-1}{l}\{\alpha(k,l)-\beta(k,l)\},$$

$$\alpha(k,l)=\begin{cases}\dfrac{P^{a+m+l}(Q+1)^{k-a-b-c+1}}{a+m+l}\\ \quad\times\ {}_2F_1\left(a+m+l,a+b+c-k-1;a+m+l+1;-\dfrac{P}{Q+1}\right), & \text{if } k\neq a+b+c-1,\\ \dfrac{P^{a+m+l}\log(Q+1)}{a+m+l}+\dfrac{P^{a+m+l+1}/(Q+1)}{a+m+l+1}\\ \quad\times\ {}_3F_2\left(a+m+l+1,1,1;2,a+m+l+2;-\dfrac{P}{Q+1}\right), & \text{if } k=a+b+c-1\end{cases}$$

and

$$\beta(k,l)=\begin{cases}\dfrac{P^{a+m+l}}{a+m+l}\ {}_2F_1(a+m+l,a+b+c-k-1;a+m+l+1;-P), & \text{if } k\neq a+b+c-1,\\ \dfrac{P^{a+m+l+1}}{a+m+l+1}\ {}_3F_2(a+m+l+1,1,1;2,a+m+l+2;-P), & \text{if } k=a+b+c-1.\end{cases}$$

Proof: Setting $z=1+x+y$ in (6), the double integral $H(P,Q)$ can be expressed as

$$\begin{aligned}H(P,Q)&=\int_0^P x^{a+m-1}\int_{1+x}^{1+x+Q} z^{-a-b-c}(z-x-1)^{b+n-1}dzdx\\ &=\int_0^P x^{a+m-1}\sum_{k=0}^{b+n-1}\binom{b+n-1}{k}(-1-x)^{b+n-1-k}\int_{1+x}^{1+x+Q} z^{k-a-b-c}dzdx\\ &=\int_0^P x^{a+m-1}\sum_{k=0}^{b+n-1}\binom{b+n-1}{k}(-1-x)^{b+n-1-k}\\ &\qquad\times\frac{(1+x+Q)^{k-a-b-c+1}-(1+x)^{k-a-b-c+1}}{k-a-b-c+1}dx\\ &=\sum_{k=0}^{b+n-1}\binom{b+n-1}{k}\frac{(-1)^{b+n-k-1}}{k-a-b-c+1}\sum_{l=0}^{b+n-1-k}\binom{b+n-1-k}{l}\\ &\qquad\times\int_0^P x^{a+m+l-1}\left\{(1+x+Q)^{k-a-b-c+1}-(1+x)^{k-a-b-c+1}\right\}dx\\ &=\sum_{k=0}^{b+n-1}\binom{b+n-1}{k}\frac{(-1)^{b+n-k-1}}{k-a-b-c+1}\sum_{l=0}^{b+n-1-k}\binom{b+n-1-k}{l}\{\alpha(k,l)-\beta(k,l)\},\end{aligned}$$

where

$$\alpha(k,l)=\int_0^P x^{a+m+l-1}(1+x+Q)^{k-a-b-c+1}dx$$

and

$$\beta(k,l)=\int_0^P x^{a+m+l-1}(1+x)^{k-a-b-c+1}dx.$$

If $k\neq a+b+c-1$ then a direct application of equation (2.2.6.1) in Prudnikov *et al.* (1986, volume 1) shows that $\alpha(k,l)$ and $\beta(k,l)$ reduce to the forms given in the theorem. If $k=a+b+c-1$ then a direct application of equation (2.6.10.31) in Prudnikov *et al.* (1986, volume 1) shows that $\alpha(k,l)$ and $\beta(k,l)$ reduce to the forms given in the theorem.

3. PARTICULAR CASES

Using special properties of the hypergeometric functions, one can obtain elementary forms for $E(X^mY^n)$ for given integer values of a, b, m and n. This is illustrated in the corollaries below.

Corollary 1 If (X,Y) has the joint pdf (3) with $a=b=1$ then $E(X)$ is given by (5) with

$$\begin{aligned}H(A,B)=\{&1-(1+\theta A)^{-a}-(1+\theta A)^{-a}a\theta A-(1+\phi B)^{-a}+(1+\phi B+\theta A)^{-a}a\theta A\\&+(1+\phi B+\theta A)^{-a}\phi B-(1+\phi B)^{-a}\phi B+(1+\phi B+\theta A)^{-a}\}\\&/\{\phi\theta^2a(a^2-1)\}.\end{aligned}$$

Corollary 2 If (X,Y) has the joint pdf (3) with $a=b=1$ then $E(X^2)$ is given by (5) with

$$\begin{aligned}H(A,B)=\{&2+(1+\theta A)^{-a}\theta^2A^2a-(1+\theta A)^{-a}\theta^2A^2a^2-2(1+\theta A)^{-a}a\theta A-2(1+\theta A)^{-a}\\&-2(1+\phi B)^{-a}+(1+\phi B+\theta A)^{-a}\theta^2A^2a^2+2(1+\phi B+\theta A)^{-a}\\&-4(1+\phi B)^{-a}B\phi+4(1+\phi B+\theta A)^{-a}B\phi+2(1+\phi B+\theta A)^{-a}a\theta A\\&-2(1+\phi B)^{-a}\phi^2B^2+2(1+\phi B+\theta A)^{-a}a\theta A\phi B\\&-(1+\phi B+\theta A)^{-a}\theta^2A^2a+2(1+\phi B+\theta A)^{-a}\phi^2B^2\}\\&/\{a\theta^3\phi(-2a^2-a+2+a^3)\}.\end{aligned}$$

Corollary 3 If (X,Y) has the joint pdf (3) with $a=b=1$ then $E(XY)$ is given by (5) with

$$\begin{aligned}H(A,B)=\{&1-(1+\theta A)^{-a}a\theta A-(1+\theta A)^{-a}\theta^2A^2a+(1+\theta A)^{-a}\theta^2A^2-(1+\theta A)^{-a}\\&-(1+\phi B+\theta A)^{-a}\phi^2B^2-(1+\phi B)^{-a}a\phi B+(1+\phi B+\theta A)^{-a}a^2\theta A\phi B\\&+(1+\phi B+\theta A)^{-a}a\theta A+(1+\phi B+\theta A)^{-a}-(1+\phi B+\theta A)^{-a}a\theta A\phi B\\&-(1+\phi B)^{-a}a\phi^2B^2+(1+\phi B+\theta A)^{-a}\theta^2A^2a-(1+\phi B)^{-a}\\&-(1+\phi B+\theta A)^{-a}\theta^2A^2+(1+\phi B+\theta A)^{-a}a\phi B+(1+\phi B)^{-a}\phi^2B^2\\&+(1+\phi B+\theta A)^{-a}a\phi^2B^2\}/\{\phi^2\theta^2a(-2a^2-a+2+a^3)\}.\end{aligned}$$

4. ESTIMATION

Here, we consider estimation of the seven parameters of (3) by the method of maximum likelihood and provide the associated Fisher information matrix. The log-likelihood for a random sample $(x_1,y_1),\ldots,(x_n,y_n)$ from (3) is:

$$\begin{aligned}\log L(A,B,C,D,a,b,c)=\;&n\log\Gamma(a+b+c)-n\log\Omega+n\log\Gamma(a)-n\log\Gamma(b)-n\log\Gamma(c)\\&+(a-1)\sum_{j=1}^{n}\log x_j+(b-1)\sum_{j=1}^{n}\log y_j\\&-(a+b+c)\sum_{j=1}^{n}\log(1+x_j+y_j).\end{aligned}$$

It follows that maximum likelihood estimators of the seven parameters are the solutions of the simultaneous equations

$$n\Psi(a+b+c)-n\Psi(a)-\frac{n}{\Omega}\frac{\partial\Omega}{\partial a}=\sum_{j=1}^{n}\log\left(1+x_j+y_j\right)-\sum_{j=1}^{n}\log x_j,$$

$$n\Psi(a+b+c)-n\Psi(b)-\frac{n}{\Omega}\frac{\partial\Omega}{\partial b}=\sum_{j=1}^{n}\log\left(1+x_j+y_j\right)-\sum_{j=1}^{n}\log y_j,$$

$$n\Psi(a+b+c)-n\Psi(c)-\frac{n}{\Omega}\frac{\partial\Omega}{\partial c}=\sum_{j=1}^{n}\log\left(1+x_j+y_j\right),$$

$$\frac{n}{\Omega}\frac{\partial\Omega}{\partial A}=0,$$

$$\frac{n}{\Omega}\frac{\partial\Omega}{\partial B}=0,$$

$$\frac{n}{\Omega}\frac{\partial\Omega}{\partial C}=0,$$

and

$$\frac{n}{\Omega}\frac{\partial\Omega}{\partial D}=0,$$

where $\Psi(x)=d\log\Gamma(x)/dx$ is the digamma function. The second order derivatives of $\log L$ are all constants, so the elements of the associated Fisher information matrix are

$$E\left(-\frac{\partial^2\log L}{\partial a^2}\right)=\frac{n}{\Omega}\frac{\partial^2\Omega}{\partial a^2}-\frac{n}{\Omega^2}\left(\frac{\partial\Omega}{\partial a}\right)^2-n\Psi'(a+b+c)+n\Psi'(a),$$

$$E\left(-\frac{\partial^2\log L}{\partial a\partial b}\right)=\frac{n}{\Omega}\frac{\partial^2\Omega}{\partial a\partial b}-\frac{n}{\Omega^2}\frac{\partial\Omega}{\partial a}\frac{\partial\Omega}{\partial b}-n\Psi'(a+b+c),$$

$$E\left(-\frac{\partial^2\log L}{\partial a\partial c}\right)=\frac{n}{\Omega}\frac{\partial^2\Omega}{\partial a\partial c}-\frac{n}{\Omega^2}\frac{\partial\Omega}{\partial a}\frac{\partial\Omega}{\partial c}-n\Psi'(a+b+c),$$

$$E\left(-\frac{\partial^2\log L}{\partial b^2}\right)=\frac{n}{\Omega}\frac{\partial^2\Omega}{\partial b^2}-\frac{n}{\Omega^2}\left(\frac{\partial\Omega}{\partial b}\right)^2-n\Psi'(a+b+c)+n\Psi'(b),$$

$$E\left(-\frac{\partial^2\log L}{\partial b\partial c}\right)=\frac{n}{\Omega}\frac{\partial^2\Omega}{\partial b\partial c}-\frac{n}{\Omega^2}\frac{\partial\Omega}{\partial b}\frac{\partial\Omega}{\partial c}-n\Psi'(a+b+c),$$

and

$$E\left(-\frac{\partial^2\log L}{\partial c^2}\right)=\frac{n}{\Omega}\frac{\partial^2\Omega}{\partial c^2}-\frac{n}{\Omega^2}\left(\frac{\partial\Omega}{\partial c}\right)^2-n\Psi'(a+b+c)+n\Psi'(c),$$

where $\Psi'(x)$ is the first derivative of $\Psi(x)$. The remaining elements of the Fisher information matrix all take the form

$$E\left(-\frac{\partial^2\log L}{\partial\theta_1\partial\theta_2}\right)=\frac{n}{\Omega}\frac{\partial^2\Omega}{\partial\theta_1\partial\theta_2}-\frac{n}{\Omega^2}\frac{\partial\Omega}{\partial\theta_1}\frac{\partial\Omega}{\partial\theta_2}.$$

School of Mathematics
University of Manchester

SARALEES NADARAJAH

REFERENCES

I.S. GRADSHTEYN, I.M. RYZHIK, (2000), *Table of Integrals, Series, and Products (sixth edition),* Academic Press, San Diego.

M.L.T. LEE, A.J. GROSS, (1991), *Lifetime distributions under unknown environment*, "Journal of Statistical Planning and Inference", 29, pp. 137-143.

D.D. LIEN, (1985), *Moments of truncated bivariate log–normal distributions*, "Economics Letters", 19, pp. 243-247.

T.K. NAYAK, (1987), *Multivariate lomax distribution - properties and usefulness in reliability theory*, "Journal of Applied Probability", 24, 170-177.

A.P. PRUDNIKOV, Y.A. BRYCHKOV, O.I. MARICHEV, (1986), *Integrals and Series* (volumes 1, 2 and 3), Gordon and Breach Science Publishers, Amsterdam.

SUMMARY

A truncated bivariate inverted Dirichlet distribution

A truncated version of the bivariate inverted dirichlet distribution is introduced. Unlike the inverted dirichlet distribution, this possesses finite moments of all orders and could therefore be a better model for certain practical situations. One such situation is discussed. The moments, maximum likelihood estimators and the Fisher information matrix for the truncated distribution are derived.