A Family of J-Shaped Bivariate Distributions

Saralees Nadarajah, *Department of Statistics and Applied Probability, University of California, Santa Barbara, California 93106-3110, United States of America, E-mail: snadar@pstat.ucsb.edu*

Samuel Kotz, *Department of Engineering Management and Systems Engineering, George Washington University, Washington, D.C. 20052, United States of America, E-mail: kotz@seas.gwu.edu*

1. Introduction

Topp and Leone (1955) proposed a family univariate distributions by formulating the cumulative distribution function (cdf), F, as

$$F(x) = \begin{cases} \left(\frac{x}{b}\right)^{\nu} \left(2 - \frac{x}{b}\right)^{\nu} & \text{if } 0 \leq x \leq b < \infty, \\ 0 & \text{if } x < 0, \\ 1 & \text{if } x > b, \end{cases} \tag{1}$$

where $0 < \nu < 1$. It is easily seen that $f(x) > 0$, $f^{'}(x) < 0$ and $f^{''}(x) > 0$ for all $0 < x < b$, where $f(x)$ is the probability density function (pdf) obtained by differentiating (1), $f^{'}$ is the first derivative of f and $f^{''}$ is the second derivative of f. Thus the pdfs of (1) have graphs which are positive between 0 and b, have negative slopes between 0 and b, and for each of them the slope is always increasing between 0 and b. Hence, the curves are J-shaped and the distribution (1) is referred to as a J-shaped distribution.

2. Bivariate Generalization

We propose the following bivariate generalization of (1):

$$\begin{aligned} f(x_1, x_2) &= \frac{2\nu_1\nu_2(1+\nu_2)}{b_1 b_2} \left(\frac{x_1}{b_1}\right)^{\nu_1 - 1} \left(1 - \frac{x_1}{b_1}\right) \left(\frac{x_2}{b_2}\right)^{\nu_2 - 1} \left(1 - \frac{x_2}{b_2}\right) \\ &\quad \times {}_2F_1\left(1 - \nu_1, 2 + \nu_2; 2; -\left(1 - \frac{x_1}{b_1}\right)\left(1 - \frac{x_2}{b_2}\right)\right) \end{aligned} \tag{2}$$

for $0 \leq x_1 \leq b_1 < \infty$, $0 \leq x_2 \leq b_2 < \infty$ and $0 < \nu_2 \leq \nu_1 < 1$, where

$${}_pF_q\left(\alpha_1, \alpha_2, \ldots, \alpha_p; \beta_1, \beta_2, \ldots, \beta_q; x\right) = \sum_{j=0}^{\infty} \frac{(\alpha_1)_j (\alpha_2)_j \cdots (\alpha_p)_j}{(\beta_1)_j (\beta_2)_j \cdots (\beta_q)_j} \frac{x^j}{j!} \tag{3}$$

is the generalized hypergeometric function – see Section 9.1 of Gradshteyn and Ryzhik (1995) for details. Here f denotes the joint pdf. Let us verify whether f is indeed a pdf. First we show that it integrates to 1: under the transformations $u_k = x_k / b_k$, $k = 1, 2$,

$$\begin{aligned} & \int_0^{b_2} \int_0^{b_1} f(x_1, x_2)\, dx_1 dx_2 \\ = \;& 2\nu_1\nu_2 (1+\nu_2) \int_0^1 u_2 (1-u_2)^{\nu_2 - 1} \int_0^1 u_1 (1-u_1)^{\nu_1 - 1}\; {}_2F_1(1-\nu_1, 2+\nu_2; 2; -u_1u_2)\, du_1 du_2 \\ = \;& \frac{2\nu_2(1+\nu_2)}{1+\nu_1} \int_0^1 u_2 (1-u_2)^{\nu_2 - 1}\; {}_2F_1(1-\nu_1, 2+\nu_2; 2+\nu_1; -u_2)\, du_2, \end{aligned} \tag{4}$$

where we have used equation (7.512.11) in Gradshteyn and Ryzhik (1995) to reduce the integration with respect to u_1. Similarly, applying equation (7.512.12) for the integral with respect to u_2, we can reduce (4) to:

$$\frac{2}{1+\nu_1}\; {}_3F_2(2, 1-\nu_1, 2+\nu_2; 2+\nu_2, 2+\nu_1; -1),$$

which is equal to

$$\frac{2}{1+\nu_1}\ {}_2F_1\left(2, 1-\nu_1; 2+\nu_1; -1\right)$$

by definition (3). But,

$${}_2F_1\left(2, 1-\nu_1; 2+\nu_1; -1\right) = \sum_{j=0}^{\infty} \frac{(2)_j (1-\nu_1)_j}{(2+\nu_1)_j} \frac{(-1)^j}{j!} = \sum_{j=0}^{\infty} (j+1) \frac{(1-\nu_1)_j}{(2+\nu_1)_j} (-1)^j = \frac{1+\nu_1}{2},$$

where the final step follows by equation (6.6.24) in Hansen (1975). Thus, we have established that f integrates to 1. To show non-negativity, rewrite f, using definition (3), as

$$f\left(x_1, x_2\right) = \frac{2\nu_1\nu_2(1+\nu_2)}{b_1 b_2} \left(\frac{x_1}{b_1}\right)^{\nu_1-1} \left(1-\frac{x_1}{b_1}\right) \left(\frac{x_2}{b_2}\right)^{\nu_2-1} \left(1-\frac{x_2}{b_2}\right) \sum_{j=0}^{\infty} (-1)^j a_j, \tag{5}$$

where

$$a_j = \frac{(1-\nu_1)_j (2+\nu_2)_j}{(1)_j (2)_j} \left(1-\frac{x_1}{b_1}\right)^j \left(1-\frac{x_2}{b_2}\right)^j, \qquad j \geq 0.$$

Since $0 < \nu_2 \leq \nu_1 < 1$ it is not difficult to note that a_j, $j \geq 0$ is a decreasing sequence of positive numbers. Thus, the infinite sum must be bounded between $a_0 = 1$ and $a_0 - a_1 = 1 - a_1/a_0 \geq 0$. Hence, from (5), f must be a non-negative function.

Now let us see how (2) generalizes (1). Consider the case $\nu_1 = \nu_2 = \nu$, say. Then the first marginal density

$$\begin{aligned}
f_1(x_1) &= \int_0^{b_2} f\left(x_1, x_2\right) dx_2 \\
&= \frac{2\nu^2(1+\nu)}{b_1} \left(\frac{x_1}{b_1}\right)^{\nu-1} \left(1-\frac{x_1}{b_1}\right) \\
&\quad \times \int_0^1 u_2 \left(1-u_2\right)^{\nu-1}\ {}_2F_1\left(1-\nu, 2+\nu; 2; -\left(1-\frac{x_1}{b_1}\right) u_2\right) du_2 \\
&= \frac{2\nu}{b_1} \left(\frac{x_1}{b_1}\right)^{\nu-1} \left(1-\frac{x_1}{b_1}\right)\ {}_2F_1\left(1-\nu, 2+\nu; 2+\nu; -\left(1-\frac{x_1}{b_1}\right)\right),
\end{aligned} \tag{6}$$

where again we have used the transformation $u_2 = x_2/b_2$ and applied equation (7.512.11) in Gradshteyn and Ryzhik (1995) to reduce the integration. By definition (3),

$$\begin{aligned}
{}_2F_1\left(1-\nu, 2+\nu; 2+\nu; -\left(1-\frac{x_1}{b_1}\right)\right) &= {}_1F_0\left(1-\nu; -\left(1-\frac{x_1}{b_1}\right)\right) \\
&= \sum_{j=0}^{\infty} \binom{\nu-1}{j} \left(1-\frac{x_1}{b_1}\right)^j \\
&= \left(2-\frac{x_1}{b_1}\right)^{\nu-1}.
\end{aligned}$$

Substituting this into (6), we get the expression

$$\frac{2\nu}{b_1} \left(\frac{x_1}{b_1}\right)^{\nu-1} \left(1-\frac{x_1}{b_1}\right) \left(2-\frac{x_1}{b_1}\right)^{\nu-1}$$

for the marginal density, which is identical to the derivative of (1). Hence, (2) is indeed a generalization of (1).

Actually, following similar calculations, one can show that for general ν_1, ν_2

$$f_1\left(x_1\right) = \frac{2\nu_1}{b_1} \left(\frac{x_1}{b_1}\right)^{\nu_1-1} \left(1-\frac{x_1}{b_1}\right) \left(2-\frac{x_1}{b_1}\right)^{\nu_1-1} \tag{7}$$

and

$$f_2\left(x_2\right) = \frac{2\nu_2(1+\nu_2)}{b_2(1+\nu_1)} \left(\frac{x_2}{b_2}\right)^{\nu_2-1} \left(1-\frac{x_2}{b_2}\right)\ {}_2F_1\left(1-\nu_1, 2+\nu_2; 2+\nu_1; -\left(1-\frac{x_2}{b_2}\right)\right).$$

3. Joint and Marginal CDFs

The joint cdf, F, associated with (2) can be obtained by integrating each term of (5) as an incomplete beta integral. Calculations show that

$$\begin{aligned} F(x_1, x_2) &= \frac{2}{1+\nu_1} \sum_{j=0}^{\infty} (-1)^j \frac{(j+1)(1-\nu_1)_j}{(2+\nu_1)_j} \left\{ 1 - B\left(\nu_1, j+2; 1 - \frac{x_1}{b_1}\right) \right\} \\ &\quad \times \left\{ 1 - B\left(\nu_2, j+2; 1 - \frac{x_2}{b_2}\right) \right\}, \end{aligned}$$

where

$$B(\alpha, \beta; u) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^u x^{\alpha-1}(1-x)^{\beta-1} dx$$

is a normalized incomplete beta integral. The marginal cdfs, F_1 and F_2, follow easily from the above. We have

$$F_1(x_1) = 1 - \frac{2}{1+\nu_1} \sum_{j=0}^{\infty} (-1)^j \frac{(j+1)(1-\nu_1)_j}{(2+\nu_1)_j} B\left(\nu_1, j+2; 1 - \frac{x_1}{b_1}\right)$$

and

$$F_2(x_2) = 1 - \frac{2}{1+\nu_1} \sum_{j=0}^{\infty} (-1)^j \frac{(j+1)(1-\nu_1)_j}{(2+\nu_1)_j} B\left(\nu_2, j+2; 1 - \frac{x_2}{b_2}\right).$$

4. Characteristic Function and Moments

Set $U_k = X_k/b_k$, $k = 1, 2$. Then, using the representation (5), the characteristic function (cf)

$$\begin{aligned} E\left(\exp\left\{i\left(t_1 X_1 + t_2 X_2\right)\right\}\right) &= 2\nu_1\nu_2\left(1+\nu_2\right) \exp\left\{i\left(t_1 b_1 + t_2 b_2\right)\right\} \sum_{j=0}^{\infty} (-1)^j \frac{(1-\nu_1)_j (2+\nu_2)_j}{(1)_j (2)_j} \\ &\quad \times \int_0^1 \exp\left(-i t_1 b_1 u_1\right) u_1^{j+1} \left(1-u_1\right)^{\nu_1 - 1} du_1 \\ &\quad \times \int_0^1 \exp\left(-i t_2 b_2 u_2\right) u_2^{j+1} \left(1-u_2\right)^{\nu_2 - 1} du_2. \end{aligned} \tag{8}$$

By equations (3.38.3) and (9.212.1) in Gradshteyn and Ryzhik (1995),

$$\begin{aligned} \int_0^1 \exp\left(-i t_k b_k u_k\right) u_k^{j+1} \left(1-u_k\right)^{\nu_k - 1} du_k &= \frac{(2)_j}{(\nu_k)_{j+2}}\ {}_1F_1\left(j+2; j+2+\nu_k; -i t_k b_k\right) \\ &= \frac{(2)_j}{(\nu_k)_{j+2}} \exp\left(-i t_k b_k\right)\ {}_1F_1\left(\nu_k; j+2+\nu_k; i t_k b_k\right). \end{aligned}$$

Substituting this into (8) for $k = 1, 2$, we obtain the following expression for the cf:

$$\frac{2}{1+\nu_1} \sum_{j=0}^{\infty} (-1)^j \frac{(j+1)(1-\nu_1)_j}{(2+\nu_1)_j}\ {}_1F_1\left(\nu_1; j+2+\nu_1; i t_1 b_1\right)\ {}_1F_1\left(\nu_2; j+2+\nu_2; i t_2 b_2\right).$$

It follows that the product moment

$$E\left(X_1^m X_2^n\right) = \frac{2 b_1^m b_2^n}{1+\nu_1} \sum_{j=0}^{\infty} (-1)^j \frac{(j+1)(1-\nu_1)_j}{(2+\nu_1)_j} \frac{(\nu_1)_m}{(j+2+\nu_1)_m} \frac{(\nu_2)_n}{(j+2+\nu_2)_n},$$

which simplifies to

$$\frac{2 b_1^m b_2^n}{1+\nu_1} \sum_{j=0}^{\infty} (-1)^j (j+1)(1-\nu_1)_j (2+\nu_2)_j \frac{(\nu_1)_m}{(2+\nu_1)_{m+j}} \frac{(\nu_2)_n}{(2+\nu_2)_{n+j}}$$

$$
\begin{aligned}
&= \frac{2b_1^m b_2^n}{1+\nu_1}\sum_{j=0}^{\infty}(-1)^j \frac{(2)_j(1-\nu_1)_j(2+\nu_2)_j}{(1)_j(m+2+\nu_1)_j(n+2+\nu_2)_j}\frac{(\nu_1)_m}{(2+\nu_1)_m}\frac{(\nu_2)_n}{(2+\nu_2)_n} \\
&= \frac{2b_1^m b_2^n \nu_1\nu_2(1+\nu_2)}{(m+\nu_1)(m+1+\nu_1)(n+\nu_2)(n+1+\nu_2)} \\
&\quad \times {}_3F_2\left(2,1-\nu_1,2+\nu_2;m+2+\nu_1,n+2+\nu_2;-1\right),
\end{aligned}
$$

where the final equality follows from definition (3).

5. Conditional Densities and Moments

¿From (2) and (7), the conditional pdf of X_2 given $X_1 = x_1$ is

$$
\frac{\nu_2(1+\nu_2)}{b_2}\left(\frac{x_2}{b_2}\right)^{\nu_2-1}\left(1-\frac{x_2}{b_2}\right)\left(2-\frac{x_1}{b_1}\right)^{1-\nu_1}{}_2F_1\left(1-\nu_1,2+\nu_2;2;-\left(1-\frac{x_1}{b_1}\right)\left(1-\frac{x_2}{b_2}\right)\right).
$$

Under the transformation $u_2 = x_2/b_2$, the conditional moment

$$
\begin{aligned}
E\left(X_2^n \mid X_1 = x_1\right) &= \nu_2\left(1+\nu_2\right)b_2^n\left(2-\frac{x_1}{b_1}\right)^{1-\nu_1}\int_0^1 u_2\left(1-u_2\right)^{n-1+\nu_2} \\
&\quad \times {}_2F_1\left(1-\nu_1,2+\nu_2;2;-u_2\left(1-\frac{x_1}{b_1}\right)\right)du_2 \\
&= \frac{\nu_2(1+\nu_2)b_2^n}{(n+\nu_2)(n+1+\nu_2)}\left(2-\frac{x_1}{b_1}\right)^{1-\nu_1} \\
&\quad \times {}_2F_1\left(1-\nu_1,2+\nu_2;n+2+\nu_2;-\left(1-\frac{x_1}{b_1}\right)\right),
\end{aligned}
$$

where again we have used equation (7.512.11) in Gradshteyn and Ryzhik (1995). Similar calculations show that the conditional pdf of X_1 given $X_2 = x_2$ is

$$
\frac{\nu_1(1+\nu_1)}{b_1}\left(\frac{x_1}{b_1}\right)^{\nu_1-1}\left(1-\frac{x_1}{b_1}\right)\frac{{}_2F_1\left(1-\nu_1,2+\nu_2;2;-\left(1-\frac{x_1}{b_1}\right)\left(1-\frac{x_2}{b_2}\right)\right)}{{}_2F_1\left(1-\nu_1,2+\nu_2;2+\nu_1;-\left(1-\frac{x_2}{b_2}\right)\right)}
$$

and that the corresponding conditional moment

$$
E\left(X_1^m \mid X_2 = x_2\right) = \frac{\nu_1(1+\nu_1)b_1^m}{(m+\nu_1)(m+1+\nu_1)}\frac{{}_2F_1\left(1-\nu_1,2+\nu_2;m+2+\nu_1;-\left(1-\frac{x_2}{b_2}\right)\right)}{{}_2F_1\left(1-\nu_1,2+\nu_2;2+\nu_1;-\left(1-\frac{x_2}{b_2}\right)\right)}.
$$

REFERENCES

Gradshteyn, I.S. and Ryzhik, I.M. (1995). *Tables of Integrals, Series, and Products.* New York: Academic Press.

Hansen, E.R. (1975). *A Table of Series and Products.* New Jersey: Prentice-Hall.

Topp, C.W. and Leone, F.C. (1955). A family of J-shaped frequency functions. *Journal of the American Statistical Association*, **50**, 209-219.

RESUME

This paper concerns a family of univariate distributions suggested by Topp and Leone in 1955. To the best of our knowledge this family of distributions does not appear to have been studied further since it appeared. Here we propose a bivariate generalization of the family. We derive its characteristic function, product moments, conditional densities and conditional moments.