A bivariate *F* distribution with marginals on arbitrary numerator and denominator degrees of freedom, and related bivariate beta and *t* distributions

Journal Item

A bivariate F distribution with marginals on arbitrary numerator and denominator degrees of freedom, and related bivariate beta and t distributions

A.H. El-Bassiouny[a] & M.C. Jones[b]

[a] *King Saud University, Saudi Arabia*

[b] *The Open University, UK*

Abstract

The classical bivariate F distribution arises from ratios of chi-squared random variables with common denominators. A consequent disadvantage is that its univariate F marginal distributions have one degree of freedom parameter in common. In this paper, we add a further independent chi-squared random variable to the denominator of one of the ratios and explore the extended bivariate F distribution, with marginals on arbitrary degrees of freedom, that results. Transformations linking F, beta and skew t distributions are then applied componentwise to produce bivariate beta and skew t distributions which also afford marginal (beta and skew t) distributions with arbitrary parameter values. We explore a variety of properties of these distributions and give an example of a potential application of the bivariate beta distribution in Bayesian analysis.

Keywords: Chi-squared distribution; Positive dependence; Transformation

1. Introduction

Let X_i, $i = 1, 2, 3$, be independent chi-squared random variables with degrees of freedom (d.f.) $n_i > 0$, $i = 1, 2, 3$. The usual bivariate F distribution (e.g. Johnson & Kotz, 1972, Chapter 40, Sections 7 and 8, Hutchinson & Lai, 1990, Section 6.3) is the joint distribution of

$$F_1' = \frac{n_3 X_1}{n_1 X_3} \quad \text{and} \quad F_2' = \frac{n_3 X_2}{n_2 X_3}. \tag{1}$$

The univariate marginal distributions are, of course, F distributions and the common denominator random variable serves to introduce positive correlation into the distribution. However, this common denominator also gives rise to a particular disadvantage of the distribution as a model for data: its univariate F marginals have the same denominator d.f., the marginal d.f. being $\{n_1, n_3\}$ and $\{n_2, n_3\}$.

In this paper, we provide an alternative distribution with F marginals, each with its own arbitrary numerator and denominator d.f. parameters. To this end, let X_4 be a further chi-squared random variable, independent of X_1, X_2, X_3, with d.f. $n_4 \geq 0$ and write $\upsilon_1 = n_3, \upsilon_2 = n_3 + n_4$. The proposed bivariate F distribution with marginal F distributions on arbitrary numerator and denominator d.f., $\{n_1, \upsilon_1\}$ and $\{n_2, \upsilon_2\}$, is defined as the joint distribution of

$$F_1 = \frac{\upsilon_1 X_1}{n_1 X_3} \quad \text{and} \quad F_2 = \frac{\upsilon_2 X_2}{n_2 (X_3 + X_4)}. \tag{2}$$

Note that $\upsilon_2 \geq \upsilon_1$ and the way that F_1 and F_2 relate to the data variables has to be arranged accordingly. Positive correlation still arises through the continued presence of X_3 in the denominators of both F_1 and F_2. The standard bivariate F distribution, of course, corresponds to $n_4 = 0$.

This new bivariate F distribution is the first major focus of the paper and its properties are investigated in Section 2. There, we derive its density function, conjecture and provide evidence for its unimodality, illustrate its shapes graphically, present its moments, investigate its dependence properties (including correlation and proving positive quadrant dependence), and look at related distributions (including its conditionals).

Now, in one dimension, there is a simple set of transformations linking $Y \sim F$, $B \sim$ beta and $T \sim$ skew t distributions (Jones & Faddy, 2003) given by

$$T = \frac{\sqrt{\omega}}{2}\left(\sqrt{\frac{n}{\upsilon}}\sqrt{Y} - \sqrt{\frac{\upsilon}{n}}\frac{1}{\sqrt{Y}}\right); \quad Y = \frac{\upsilon}{\omega n}\left(T + \sqrt{\omega + T^2}\right)^2, \qquad (3.a)$$

$$B = \frac{1}{2}\left(1 + \frac{T}{\sqrt{\omega + T^2}}\right); \quad T = \frac{\sqrt{\omega}}{2}\frac{(2B-1)}{\sqrt{B(1-B)}}, \qquad (3.b)$$

$$Y = \frac{\upsilon}{n}\frac{B}{1-B}; \quad B = \frac{nY}{\upsilon + nY} \qquad (3.c)$$

where $\omega = (n+\upsilon)/2$. In fact, while $\{n, \upsilon\}$ are the parameters of the F distribution of Y, in the standard parametrisation of the beta distribution of B the parameters are $\{n/2, \upsilon/2\}$ and likewise for the skew t distribution of T. When $n = \upsilon$, T follows a symmetric Student t distribution on υ d.f. Jones (2001) applies transformations (3) to both marginal variates F_1' and F_2' of the standard F distribution of (1) (in place of Y in (3)) to obtain bivariate beta and skew t distributions with parameters $\{n_1/2, n_3/2\}$ and $\{n_2/2, n_3/2\}$, respectively. The bivariate beta distribution was, unfortunately, not acknowledged to be a special case of the multivariate generalized beta distribution proposed by Libby & Novick (1982) and has since also been reinvented by Olkin & Liu (2003). Note, again, the restriction that the second parameter of each marginal distribution be the same.

In Sections 3 and 4, we obtain bivariate beta and skew t distributions with arbitrary marginal parameters $\{n_1/2, \upsilon_1/2\}$ and $\{n_2/2, \upsilon_2/2\}$ by applying transformations (3) to both marginal variates F_1 and F_2 of the F distribution of (2) (as Y). The bivariate beta and skew t distributions that appear in Jones (2001) correspond to $n_4 = 0$ i.e. $\upsilon_1 = \upsilon_2$. We concentrate particularly on the bivariate beta distribution because of its greater tractability; Section 3 contains, for this distribution, a study of a subset of the properties studied for the bivariate F distribution in Section 2. Section 4 provides only the briefest of introductions to the related bivariate skew t distribution.

All of these distributions have practical potential as empirical distributions for data with marginals of the prescribed sort, or as families of prior distributions. The brief example of use of these distributions that we present in the closing Section 5 arises in the latter context and is specific to the case of the bivariate beta distribution.

2. The bivariate F distribution

2.1. Density

Theorem 2.1. *The joint density function of F_1 and F_2 defined in (2) has the following closed form expression:*

$$\begin{aligned} f_{F_1,F_2}(f_1,f_2) &= C_{12} f_1^{n_1/2-1} f_2^{n_2/2-1} \left(1+\frac{n_1}{\upsilon_1}f_1+\frac{n_2}{\upsilon_2}f_2\right)^{-N/2} \\ &\times \; F\left(\frac{N}{2},\frac{\upsilon_2-\upsilon_1}{2};\frac{n_1+\upsilon_2}{2};\frac{\frac{n_1}{\upsilon_1}f_1}{1+\frac{n_1}{\upsilon_1}f_1+\frac{n_2}{\upsilon_2}f_2}\right) \end{aligned} \quad (4)$$

on $f_1, f_2 \geq 0$ where

$$C_{12} = \frac{\left(\frac{n_1}{\upsilon_1}\right)^{n_1/2}\left(\frac{n_2}{\upsilon_2}\right)^{n_2/2}}{B\left(\frac{n_1}{2},\frac{\upsilon_1}{2}\right)B\left(\frac{n_2}{2},\frac{n_1+\upsilon_2}{2}\right)} = \left(\frac{n_1}{\upsilon_1}\right)^{n_1/2}\left(\frac{n_2}{\upsilon_2}\right)^{n_2/2} C'_{12}, \text{ say,}$$

$N = n_1+n_2+\upsilon_2$, $B(\cdot,\cdot)$ is the beta function and $F(\cdot,\cdot;\cdot;\cdot)$ is the Gauss hypergeometric function.

Proof for $\upsilon_2 > \upsilon_1$. Write $H_1 = (n_1/\upsilon_1)F_1$, $H_2 = (n_2/\upsilon_2)F_2$; the joint density of H_1 and H_2 will be derived, it being a simple rescaling step short of the density of F_1 and F_2. The joint density of $X_1, ..., X_4$ is

$$K_1 \prod_{i=1}^{4} x_i^{n_i/2-1}\exp(-\tfrac{1}{2}x_i) \quad \text{where} \quad 1/K_1 = \prod_{i=1}^{4} 2^{n_i/2}\Gamma\left(\tfrac{1}{2}n_i\right).$$

Utilise the transformation $H_1 = X_1/X_3$, $H_2 = X_2/(X_3+X_4)$, $Y_1 = X_3$, $Y_2 = X_3+X_4$ whose Jacobian is Y_1Y_2. The required density comes to be $h_1^{(n_1/2)-1}h_2^{(n_2/2)-1}J(h_1,h_2)$ where $J = J(h_1,h_2)$ is given by

$$J \equiv K_1 \int\int_{0<y_1<y_2<\infty} y_1^{(n_1+n_3)/2-1}e^{-y_1h_1/2}y_2^{n_2/2}(y_2-y_1)^{n_4/2-1}e^{-y_2(1+h_2)/2}dy_2dy_1.$$

Write I for the integral over y_2 in J:

$$I = \frac{y_1^{(n_2+n_4-2)/4}2^{(n_2+n_4+2)/4}\Gamma\left(\frac{1}{2}(n_4)\right)}{(1+h_2)^{(n_2+n_4+2)/4}e^{y_1(1+h_2)/4}}W_{(n_2-n_4+2)/4,(n_2+n_4)/4}\left(\tfrac{1}{2}y_1(1+h_2)\right)$$

where W is the Whittaker function (Gradshteyn & Rhyzik, 1994, 3.384.4, Sections 9.22–9.23). Writing $K(h_2) = 1/\{2^{(2n_1+n_2+\upsilon_1+\upsilon_2-2)/4}\Gamma\left(\frac{1}{2}n_1\right)\Gamma\left(\frac{1}{2}n_2\right)$ $\Gamma\left(\frac{1}{2}\upsilon_1\right)(1+h_2)^{(n_2+n_4+2)/4}\}$,

$$\begin{aligned} J &= K(h_2)\int_0^\infty x^{(2n_1+n_2+\upsilon_1+\upsilon_2-6)/4}e^{-x(1+2h_1+h_2)/4} \\ &\qquad\qquad \times W_{(n_2-n_4+2)/4,(n_2+n_4)/4}\left(\tfrac{1}{2}x(1+h_2)\right)dx \\ &= K(h_2)\frac{\Gamma\left(\frac{1}{2}N\right)\Gamma(\frac{1}{2}(n_1+\upsilon_1))}{\Gamma(\frac{1}{2}(n_1+\upsilon_2))}\frac{2^{(2n_1+n_2+\upsilon_1+\upsilon_2-2)/4}(1+h_2)^{(n_2+n_4+2)/4}}{(1+h_1+h_2)^{N/2}} \\ &\qquad\qquad \times F\left(\tfrac{1}{2}N, \tfrac{1}{2}(\upsilon_2-\upsilon_1); \tfrac{1}{2}(n_1+\upsilon_2); h_1/(1+h_1+h_2)\right) \end{aligned}$$

(Gradshteyn & Rhyzik, 1994, 7.621.3). All that remains is some simple further manipulation. ◇

Remark. Although the proof is given only for the case $\upsilon_2 > \upsilon_1$, formula (4) holds for the case $\upsilon_2 = \upsilon_1$ too, when it reduces to the standard bivariate F density (48) of Johnson & Kotz (1972, Chapter 40). The key observation is that the hypergeometric function is unity when its second argument is zero.

2.2. Unimodality and graphs of density

We conjecture that density f_{F_1,F_2} is either monotonically decreasing or unimodal, depending on the values of its parameters. Differentiating log $f_{F_1,F_2}(f_1, f_2)$ with respect to f_1 and f_2 and setting each derivative equal to zero results in a pair of equations which are satisfied by the single value

$$y_0 = \frac{(n_2-2)\upsilon_2}{n_2(\upsilon_2+4)}.$$

When $n_2 > 2$, any 'internal' modes, (x_0, y_0) say, of f_{F_1,F_2} must be associated with this value y_0. For the classical bivariate F distribution with $\upsilon_1 = \upsilon_2$, it is easy to show that $x_0 = (n_1-2)\upsilon_2/\{n_1(\upsilon_2+4)\}$. Unfortunately, the corresponding value(s) of x_0 cannot be evaluated explicitly when $\upsilon_1 < \upsilon_2$ since they solve an equation, (*) say, involving two Gauss hypergeometric functions. However, non-unimodality could arise only if there are two or more values of $x_0 \geq 0$ satisfying (*). In extensive numerical investigations, we have never observed such behaviour. Hence, we conjecture unimodality when $n_2 > 2$.

When $n_2 < 2$, $f_{F_1,F_2}(0, f_2) = f_{F_1,F_2}(f_1, 0) = \infty$ and the density is monotonically decreasing on $f_1, f_2 > 0$. When $n_1 < n_2 = 2$, $f_{F_1,F_2}(0, f_2) = \infty$ but

$$f_{F_1,F_2}(f_1, 0) = C_{12} f_1^{n_1/2-1} \left(1 + \frac{n_1}{\upsilon_1} f_1\right)^{-N/2} F\left(\frac{N}{2}, \frac{\upsilon_2 - \upsilon_1}{2}; \frac{n_1 + \upsilon_2}{2}; \frac{\frac{n_1}{\upsilon_1} f_1}{1 + \frac{n_1}{\upsilon_1} f_1}\right),$$

$f_1 > 0$. When $n_1 = n_2 = 2$, $f_{F_1,F_2}(f_1, 0)$ simplifies a little, while $f_{F_1,F_2}(0, f_2) = C_{12}(1 + 2f_2/\upsilon_2)^{-N/2}$ which is monotonically decreasing from the value C_{12} at $(0, 0)$. Monotonicity of the density is conjectured from numerical evidence in both $n_1 < 2$ and $n_1 = 2$ cases, with a finite mode at the origin apparent when $n_1 = n_2 = 2$.

* * * Figs 1 and 2 about here * * *

Figs 1 and 2 display typical examples of the unimodal and finite decreasing kind, respectively. (For these graphs and Fig. 3 to follow, we used the FORTRAN routine for computing the hypergeometric function associated with Zhang & Jin (1996), available from `http://jin.ece.uiuc.edu`.) Their shapes tie in with the discussion above. The first has parameter values $n_1 = 10, n_2 = 20, \upsilon_1 = 2, \upsilon_2 = 3$ and hence marginal F distributions on $\{10, 2\}$ and $\{20, 3\}$ degrees of freedom, respectively; the second has parameter values $n_1 = n_2 = 2, \upsilon_1 = 1, \upsilon_2 = 30$ and hence marginal F distributions on $\{2, 1\}$ and $\{2, 30\}$ degrees of freedom.

2.3. Product moments and correlation

Theorem 2.2. *Let r_1 and r_2 be nonnegative such that $r_1 < \nu_1/2$ and $r_1 + r_2 < \nu_2/2$, then*

$$E(F_1^{r_1} F_2^{r_2}) = \left(\frac{\upsilon_1}{n_1}\right)^{r_1} \left(\frac{\upsilon_2}{n_2}\right)^{r_2} \frac{\Gamma\left(\frac{n_1}{2} + r_1\right) \Gamma\left(\frac{n_2}{2} + r_2\right) \Gamma\left(\frac{\upsilon_1}{2} - r_1\right) \Gamma\left(\frac{\upsilon_2}{2} - r_1 - r_2\right)}{\Gamma\left(\frac{n_1}{2}\right) \Gamma\left(\frac{n_2}{2}\right) \Gamma\left(\frac{\upsilon_1}{2}\right) \Gamma\left(\frac{\upsilon_2}{2} - r_1\right)}. \tag{5}$$

Proof. From the definition of F_1 and F_2 in (2), we can write

$$E(F_1^{r_1} F_2^{r_2}) = \left(\frac{\upsilon_1}{n_1}\right)^{r_1} \left(\frac{\upsilon_2}{n_2}\right)^{r_2} E X_1^{r_1} E X_2^{r_2} E\{X_3^{-r_1}(X_3 + X_4)^{-r_2}\}.$$

Formula (5) follows from the facts that $E(Y^r) = 2^r\Gamma(\frac{1}{2}n + r)/\Gamma(\frac{1}{2}n)$ when $Y \sim \chi_n^2$ (Johnson, Kotz & Balakrishnan, 1994a, p.420) and

$$E\left\{\frac{1}{X_3^{r_1}(X_3+X_4)^{r_2}}\right\} = \frac{\Gamma(\frac{1}{2}\upsilon_1 - r_1)\Gamma(\frac{1}{2}\upsilon_2 - r_1 - r_2)}{2^{r_1+r_2}\Gamma(\frac{1}{2}\upsilon_1)\Gamma(\frac{1}{2}\upsilon_2 - r_1)}$$

provided $r_1 < \upsilon_1/2$ and $r_1 + r_2 < \upsilon_2/2$ as was shown by Jones (2002, p.165). $\diamond$

Remark. Expression (5) reduces to the correct formula for product moments of the bivariate F distribution when $\upsilon_2 = \upsilon_1$ (expression (44) of Johnson & Kotz, 1972, Chapter 40).

Now, provided that $\upsilon_1 > 4$, the variance of F_1 takes its usual form (Johnson, Kotz & Balakrishnan, 1994b, p.326) $2\upsilon_1^2(\upsilon_1 + n_1 - 2)/\{(\upsilon_1 - 2)^2(\upsilon_1 - 4)n_1\}$. Similarly for $\mathrm{Var}(F_2)$ when $\upsilon_2 > 4$. The covariance of F_1 and F_2 is $2\upsilon_1\upsilon_2/\{(\upsilon_1 - 2)(\upsilon_2 - 2)(\upsilon_2 - 4)\}$ provided $\upsilon_1 > 2$ and $\upsilon_2 > 4$ in which case it is positive. Hence,

$$\rho_F^2 = \mathrm{Corr}^2(F_1, F_2) = \frac{n_1 n_2(\upsilon_1 - 4)}{(\upsilon_1 + n_1 - 2)(\upsilon_2 + n_2 - 2)(\upsilon_2 - 4)}, \tag{6}$$

$\upsilon_1, \upsilon_2 > 4$. Note that ρ_F is symmetric in n_1 and n_2 if $\upsilon_2 = \upsilon_1$ but not otherwise. As a function of n_1 or n_2 for fixed values of the other d.f., ρ_F increases monotonically from 0 as n_1 or $n_2 \to 0$ to an upper limit depending on the other parameter values. Now fix n_1, n_2. Recalling that $\upsilon_2 \geq \upsilon_1$, $\rho_F^2 \to n_1 n_2/(n_1 + 2)(n_2 + 2)$ as $\upsilon_2 \to 4$. (Unit correlation is therefore approached for large n_1, n_2 and small υ_1, υ_2.) For fixed υ_1, ρ_F decreases monotonically to 0 as $\upsilon_2 \to \infty$; for fixed υ_2, ρ_F increases monotonically as a function of υ_1 to the correlation associated with the classical bivariate F distribution. However, ρ_F continues to approach 0 if both υ_1 and υ_2 tend to infinity.

2.4. Positive quadrant dependence

The positive nature of the dependence between F_1' and F_2' is further reflected in the fact that they are positively quadrat dependent (PQD), i.e. $P(F_1' \leq f_1, F_1' \leq f_1) \geq P(F_1' \leq f_1)P(F_2' \leq f_2)$; this was proved by Kimball (1951).

Theorem 2.3. *F_1 and F_2 are PQD.*

Proof. Kimball's (1951) proof of the PQD nature of F_1' and F_2' can readily be adapted to the case of F_1 and F_2. All that is required is to note that Kimball's function "$f_2(q_3)$" is replaced by $P(\upsilon_2 X_2 - n_2 f_2 X_4 \leq n_2 f_2 q_3)$ which remains a strictly monotonically increasing function of q_3, and the rest of the proof goes through. $\diamond$

Theorem 2.3 implies that further scalar dependence measures such as Kendall's tau and Spearman's rho are necessarily positive (e.g. Joe, 1997, Section 2.2).

When $\upsilon_2 = \upsilon_1$ it is easy to show that the much stronger property of TP_2 (totally positive of order 2) dependence holds, but the argument does not seem extendable to the general case.

2.5. A related distribution

We have already emphasised that when $\upsilon_2 = \upsilon_1$, the new bivariate F distribution reduces to the classical bivariate F distribution. Another link with the existing literature is the joint distribution of $T_1 = \pm\sqrt{F_1}, T_2 \pm \sqrt{F_2}$ when $n_1 = n_2 = 1$. This can readily be checked to be the bivariate t distribution of Jones (2002) which has density

$$f_{T_1,T_2}(t_1,t_2) = K_{12}\left(1+\frac{t_1^2}{\upsilon_1}+\frac{t_2^2}{\upsilon_2}\right)^{-(\upsilon/2+1)} F\left(\frac{\upsilon_2}{2}+1, \frac{\upsilon_2-\upsilon_1}{2}; \frac{\upsilon_2+1}{2}; \frac{\frac{t_1^2}{\upsilon_1}}{1+\frac{t_1^2}{\upsilon_1}+\frac{t_2^2}{\upsilon_2}}\right)$$

on $t_1, t_2 \in \mathcal{R}$ where

$$K_{12} = \frac{1}{\pi}\frac{\Gamma\left(\frac{\upsilon_1+1}{2}\right)\Gamma\left(\frac{\upsilon_2}{2}+1\right)}{\sqrt{\upsilon_1\upsilon_2}\Gamma\left(\frac{\upsilon_1}{2}\right)\Gamma\left(\frac{\upsilon_2+1}{2}\right)}.$$

This distribution has univariate t marginals on υ_1 and υ_2 d.f.; it is straightforward to show that it is a unimodal distribution with mode at the origin and its contours are what Jones describes as "squashed ellipses". The usual symmetric bivariate t distribution with d.f. υ_1 (Kotz & Nadarajah, 2004) is the special case of this distribution when $\upsilon_2 = \upsilon_1$.

2.6. Conditional distributions

Write $u_i = 1 + n_i f_i/\upsilon_i$, $i = 1, 2$. The conditional density function of $F_2|F_1 = f_1$ is equal to the distribution of $Y_2 \equiv F_2/u_1$ where

$$\begin{aligned} f_{Y_2}(y_2) &= C_{2|1}(u_1)\, y_2^{\frac{n_2}{2}-1}\left(1+\frac{n_2}{\upsilon_2}y_2\right)^{-\frac{n_1+n_2+\upsilon_2}{2}} \\ &\qquad \times F\left(\frac{N}{2}, \frac{\upsilon_2-\upsilon_1}{2}; \frac{n_1+\upsilon_2}{2}; \frac{u_1-1}{u_1\left(1+\frac{n_2}{\upsilon_2}y_2\right)}\right) \end{aligned} \tag{7}$$

where

$$C_{2|1}(u_1) = \left(\frac{n_2}{\upsilon_2}\right)^{\frac{n_2}{2}} \Big/ \left\{u_1^{(\upsilon_2-\upsilon_1)/2} B\left(\frac{n_2}{2}, \frac{n_1+\upsilon_2}{2}\right)\right\}.$$

When $\upsilon_2 = \upsilon_1$, this is a scaled version of the F distribution on $\{n_2, n_1+\upsilon_2\}$ d.f. When $\upsilon_2 > \upsilon_1$, this is a scaled version of a generalized F distribution. Let $\alpha = n_2/2$ and $\beta = (n_1+\upsilon_2)/2$ and introduce $0 < c < \beta$ and $0 < p < 1$; this generalized F distribution, which we believe must be monotone or unimodal, is a scaled version of the distribution with density $y^{\alpha-1}(1+y)^{-(\alpha+\beta)}F(\alpha+\beta, \beta-c; \beta; p/(1+y)), y > 0$.

Similarly, the conditional density function of $F_1|F_2 = f_2$ is equal to the distribution of $Y_1 \equiv F_1/u_2$ where

$$\begin{aligned} f_{Y_1}(y_1) &= C_{1|2}\, y_1^{\frac{n_1}{2}-1}\left(1+\frac{n_1}{\upsilon_1}y_1\right)^{-\frac{n_1+n_2+\upsilon_2}{2}} \\ &\qquad \times F\left(\frac{N}{2}, \frac{\upsilon_2-\upsilon_1}{2}; \frac{n_1+\upsilon_2}{2}; 1-\frac{1}{1+\frac{n_1}{\upsilon_1}y_1}\right) \end{aligned} \tag{8}$$

and

$$C_{1|2} = \left(\frac{n_1}{\nu_1}\right)^{\frac{n_1}{2}} \frac{\Gamma\left(\frac{\upsilon_2}{2}\right)\Gamma\left(\frac{n_1+\upsilon_1}{2}\right)\Gamma\left(\frac{N}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{\upsilon_1}{2}\right)\Gamma\left(\frac{n_1+\upsilon_2}{2}\right)\Gamma\left(\frac{n_2+\upsilon_2}{2}\right)}.$$

Note the possibly surprising fact that density (8) has no further dependence on u_2 (in contrast to (7) and u_1). Again, when $\upsilon_2 = \upsilon_1$, (8) is a scaled version of the appropriate F distribution, while when $\upsilon_2 > \upsilon_1$, it is a scaled version of a slightly different generalized F distribution: it has (scaled) density of the form $y^{\alpha-1}(1+y)^{-(\alpha+\beta)}F(\alpha+\beta, \beta-c; \alpha+\beta-c+d; y/(1+y)), y > 0, \alpha, \beta, c, d > 0, c < \beta$.

Conditional rth moments of $F_2|F_1 = f_1$ exist provided $r < (n_1 + \upsilon_2)/2$, but are omitted because they depend on u_1 in an unedifying way in terms of the hypergeometric function. Conditional rth moments of $F_1|F_2 = f_2$ are, however, given because, from (8) being independent of u_2, the rth conditional moment of $F_1|F_2$ is proportional to the rth power of u_2.

Theorem 2.4. *Provided $r < \upsilon_1/2$, then*

$$E\left(F_1^r|F_2\right) = \left(\frac{\upsilon_1}{n_1}\right)^r u_2^r \frac{\Gamma\left(\frac{n_1}{2}+r\right)\Gamma\left(\frac{\upsilon_1}{2}-r\right)\Gamma\left(\frac{\upsilon_2}{2}\right)\Gamma\left(\frac{n_2+\upsilon_2}{2}-r\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{\upsilon_1}{2}\right)\Gamma\left(\frac{\upsilon_2}{2}-r\right)\Gamma\left(\frac{n_2+\upsilon_2}{2}\right)}. \quad (9)$$

Proof. We provide the proof in terms of Y_1 and Y_2.

$$\begin{aligned} E\left(Y_1^r|y_2\right) &= C_{1|2}\int_0^\infty y^{\frac{n_1}{2}+r-1}(1+\frac{n_1}{\upsilon_1})^{-\frac{n_1+n_2+\upsilon_2}{2}} \\ &\times F\left(\frac{n_1+n_2+\upsilon_2}{2},\frac{\upsilon_2-\upsilon_1}{2};\frac{n_1+\upsilon_2}{2};1-\frac{1}{1+\frac{n_1}{\upsilon_1}y}\right)dy_1 \\ &= C_{1|2}\left(\frac{\upsilon_1}{n_1}\right)^{\frac{n_1}{2}+r}\int_0^1 z^{\frac{n_1}{2}+r-1}(1-z)^{\frac{n_2+\upsilon_2}{2}-r-1} \\ &\qquad \times F\left(\frac{n_1+n_2+\upsilon_2}{2},\frac{\upsilon_2-\upsilon_1}{2};\frac{n_1+\upsilon_2}{2};z\right)dz \\ &= C_{1|2}\left(\frac{\upsilon_1}{n_1}\right)^{\frac{n_1}{2}+r}\frac{\Gamma\left(\frac{n_1+\upsilon_2}{2}\right)\Gamma\left(\frac{n_1}{2}+r\right)\Gamma\left(\frac{n_2+\upsilon_2}{2}-r\right)\Gamma\left(\frac{\upsilon_1}{2}-r\right)}{\Gamma\left(\frac{n_1+n_2+\upsilon_2}{2}\right)\Gamma\left(\frac{n_1+\upsilon_1}{2}\right)\Gamma\left(\frac{\upsilon_2}{2}-r\right)}. \end{aligned}$$

We used Gradshteyn & Ryzhik (1994, 7.512.3) here. Minor further manipulation completes the proof. ◇

Remark. In particular, the regression mean of F_1 is linear in f_2, being given by

$$E(F_1|F_2=f_2) = \frac{\upsilon_1(\upsilon_2-2)}{(\upsilon_1-2)(n_2+\upsilon_2-2)}\left(1+\frac{n_2}{\upsilon_2}f_2\right).$$

3. The bivariate beta distribution

3.1. Density

Let (B_1, B_2) be defined by transformation (3.c) applied to the pair F_1, F_2. The Jacobian associated with this transformation is $(n_1 n_2)^{-1} \upsilon_1 \upsilon_2 (1 - b_1)^{-2}$ $(1 - b_2)^{-2}$ and the following result is immediate.

Theorem 3.1. *The joint density function of B_1 and B_2 is given by*

$$\begin{aligned} f_{B_1,B_2}(b_1, b_2) &= C'_{12} \frac{b_1^{n_1/2-1} (1-b_1)^{\frac{n_2+\upsilon_2}{2}-1} b_2^{n_2/2-1} (1-b_2)^{\frac{n_1+\upsilon_2}{2}-1}}{(1-b_1 b_2)^{N/2}} \\ &\quad \times F\left(\frac{N}{2}, \frac{\upsilon_2-\upsilon_1}{2}; \frac{n_1+\upsilon_2}{2}; \frac{b_1(1-b_2)}{1-b_1 b_2}\right), \end{aligned} \tag{10}$$

$0 < b_1, b_2 < 1$, *where C'_{12} is given beneath (4).*

Remarks. The univariate marginals of distribution (10) are, of course, each beta distributions with arbitrary parameters $\{n_1/2, \upsilon_1/2\}$ and $\{n_2/2, \upsilon_2/2\}$, respectively. When $\upsilon_2 = \upsilon_1$, (10) reduces to the bivariate beta density in Libby & Novick (1982), Jones (2001) and Olkin & Liu (2003).

* * * Fig. 3 about here * * *

A graph of density f_{B_1,B_2} is shown in Fig. 3. The beta marginal distributions in this case have parameter values $\{4, 4\}$ (a symmetric beta marginal) and $\{4, 10\}$, respectively. The correlation (from Table 1 to follow) is 0.233. Graphs of f_{B_1,B_2} when $\upsilon_1 = \upsilon_2$ are given in Jones (2001, Fig. 4, though the two frames of that figure have had their labelling swopped) and Olkin & Liu (2003, Fig. 1).

3.2. Product moments and correlation

Theorem 3.2. *For any $r_1, r_2 > 0$,*

$$\begin{aligned} E\left(B_1^{r_1} B_2^{r_2}\right) &= \frac{\Gamma\left(\frac{n_1+\upsilon_1}{2}\right)\Gamma\left(\frac{n_2+\upsilon_2}{2}\right)\Gamma\left(\frac{N}{2}\right)\Gamma\left(\frac{n_1}{2}+r_1\right)\Gamma\left(\frac{n_2}{2}+r_2\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)\Gamma\left(\frac{\upsilon_1}{2}\right)\Gamma\left(\frac{N}{2}+r_1\right)\Gamma\left(\frac{N}{2}+r_2\right)} \\ &\quad \times \ {}_3F_2\left(\frac{N}{2}, \frac{n_2+\upsilon_2-\upsilon_1}{2}+r_2, \frac{n_1}{2}+r_1; \frac{N}{2}+r_2, \frac{N}{2}+r_1; 1\right). \end{aligned} \tag{11}$$

${}_3F_2(\cdot,\cdot,\cdot;\cdot,\cdot;\cdot)$ *is a generalized hypergeometric function.*

Proof. From (10),

$$E\left(B_1^{r_1}B_2^{r_2}\right) = C'_{12}\int_0^1 b_1^{n_1/2+r_1-1}\left(1-b_1\right)^{(n_2+\upsilon_2)/2-1} I\, db_1$$

where

$$I = \int_0^1 \frac{b_2^{n_2/2+r_2-1}\left(1-b_2\right)^{(n_1+\upsilon_2)/2-1}}{\left(1-b_1b_2\right)^{N/2}} F\left(\frac{N}{2}, \frac{\upsilon_2-\upsilon_1}{2}; \frac{n_1+\upsilon_2}{2}; \frac{b_1\left(1-b_2\right)}{1-b_1b_2}\right) db_2.$$

Using Gradshteyn & Ryzhik (1994, 7.512.8) with $\alpha = 0$,

$$I = \frac{\Gamma\left(\frac{n_2}{2}+r_2\right)\Gamma\left(\frac{n_1+\upsilon_2}{2}\right)}{\Gamma\left(\frac{N}{2}+r_2\right)} F\left(\frac{N}{2}, \frac{n_2+\upsilon_2-\upsilon_1}{2}+r_2; \frac{N}{2}+r_2; b_1\right).$$

(11) then arises from application of Gradshteyn & Ryzhik (1994, 7.512.5). ⋄

Remark. It is interesting that the generalization of the current paper has made the *form* of the moments no more complicated for this bivariate beta distribution than they are when $\upsilon_1 = \upsilon_2$.

The variance of B_1 is $2n_1\upsilon_1/(n_1+\upsilon_1)^2(n_1+\upsilon_1+2)$ ((11) does reduce to this which, of course, is the standard formula according with e.g. Johnson et al., 1994a, p.217); similarly for the variance of B_2. From (11) and a little further manipulation, one can get

$$\begin{aligned}\mathrm{Cov}(B_1,B_2) \;=\; n_1n_2\Bigg\{ {}_3F_2\left(\frac{N}{2}, \frac{n_2+\upsilon_2-\upsilon_1}{2}+1, \frac{n_1}{2}+1; \frac{N}{2}+1, \frac{N}{2}+1; 1\right) \\ \times \frac{\Gamma\left(\frac{n_1+\upsilon_1}{2}\right)\Gamma\left(\frac{n_2+\upsilon_2}{2}\right)}{2N\Gamma\left(\frac{N}{2}+1\right)\Gamma\left(\frac{\upsilon_1}{2}\right)} - \frac{1}{\left(n_1+\upsilon_1\right)\left(n_2+\upsilon_2\right)}\Bigg\}. \quad (12)\end{aligned}$$

From (12) and the expressions for the variances, $\rho_B \equiv \mathrm{Corr}(B_1,B_2)$ is symmetric in n_1 and n_2 if $\upsilon_2 = \upsilon_1$ but not otherwise. For further investigation, we resort to computational evaluation of ρ_B using Maple (Maplesoft, 2005). Table 1 contains many such values. (The very few numerical values of ρ_B in common with those given when $\upsilon_2 = \upsilon_1$ by Jones (2001) and Olkin & Liu (2003) are confirmed in these calculations.)

* * * Table 1 about here * * *

Table 1 indicates a pattern of dependence of ρ_B on n_1, n_2, υ_1 and υ_2 which reflects precisely the dependence of ρ_F on n_1, n_2, υ_1 and υ_2, obtained analytically and described at the end of Section 2.3. Particularly obvious because of the layout of Table 1 is the way that ρ_B decreases monotonically as υ_2 increases for any fixed n_1, n_2 and υ_1. It is also the case, in common with ρ_F, that ρ_B appears to increase as either n_1 or n_2 increases or, indeed, as υ_1 increases. Unit correlation is again approached for large n_1, n_2 and small υ_1, υ_2; small ρ_B is particularly associated with large υ_2.

That the PQD property holds for B_1 and B_2 follows immediately from Theorem 2.3 and the strictly monotone nature of transformations (1.c) (Joe, 1997, Theorem 2.2). In fact, densities (4) and (10) also share the same copula (e.g. Nelsen, 2006). As for the F distribution, when $\upsilon_2 = \upsilon_1$ the bivariate beta distribution is TP_2 dependent (Olkin & Liu, 2003).

3.3. Conditional distributions

The conditional density functions of $B_2|B_1$ and $B_1|B_2$ are given by

$$f(b_2|b_1) = C_{2|1}(b_1)\frac{b_2^{\frac{n_2}{2}-1}(1-b_2)^{\frac{n_1+\upsilon_2}{2}-1}}{(1-b_1b_2)^{\frac{N}{2}}}F\left(\frac{N}{2},\frac{\upsilon_2-\upsilon_1}{2};\frac{n_1+\upsilon_2}{2};\frac{b_1(1-b_2)}{1-b_1b_2}\right)$$

and

$$f(b_1|b_2) = C_{1|2}(b_2)\frac{b_1^{\frac{n_1}{2}-1}(1-b_1)^{\frac{n_2+\upsilon_2}{2}-1}}{(1-b_1b_2)^{\frac{N}{2}}}F\left(\frac{N}{2},\frac{\upsilon_2-\upsilon_1}{2};\frac{n_1+\upsilon_2}{2};\frac{b_1(1-b_2)}{1-b_1b_2}\right),$$

where

$$C_{2|1}(b_1) = \frac{(1-b_1)^{\frac{n_2+\upsilon_2-\upsilon_1}{2}}}{B\left(\frac{n_2}{2},\frac{n_1+\upsilon_2}{2}\right)}, \quad \text{and} \quad C_{1|2}(b_2) = \frac{(1-b_2)^{\frac{n_1}{2}}\,\Gamma\left(\frac{\upsilon_2}{2}\right)\Gamma\left(\frac{N}{2}\right)}{B\left(\frac{n_1}{2},\frac{\upsilon_1}{2}\right)\Gamma\left(\frac{n_2+\upsilon_2}{2}\right)\Gamma\left(\frac{n_1+\upsilon_2}{2}\right)}.$$

When $\upsilon_2 = \upsilon_1$, the conditional distributions are the three-parameter generalized beta (G3B) distributions of Libby & Novick (1982); see also Pham-Gia & Duong (1990). Else, the conditional distributions are unexplored extensions thereof. Interestingly, while the conditional moments of $B_2|B_1 = b_1$ are tractable at the same level, and with similar results, as the conditional moments of $F_2|F_1 = f_1$ (Section 2.6), the integral that comprises the expression for the conditional moments of $B_1|B_2 = b_2$ does not reduce to closed form, even in terms of hypergeometric functions.

3. Bivariate t/skew t distribution

From (3.b), the Jacobian in transforming from (B_1, B_2) to (T_1, T_2) is

$$\frac{1}{4}\frac{\omega_1\omega_2}{(\omega_1 + t_1^2)^{3/2}\,(\omega_2 + t_2^2)^{3/2}}$$

where $\omega_i = (n_i + \upsilon_i)/2$. Therefore, from (10), and using the notation $s_i = t_i/\sqrt{\omega_i + t_i^2}$,

$$\begin{aligned} f_{T_1,T_2}(t_1,t_2) &= 4C'_{12}\omega_1\omega_2 \\ &\times \frac{(1+s_1)^{n_1/2-1}(1-s_1)^{\frac{n_2+\upsilon_2}{2}-1}(1+s_2)^{n_2/2-1}(1-s_2)^{\frac{n_1+\upsilon_2}{2}-1}}{(w_1+t_1^2)^{3/2}(w_2+t_2^2)^{3/2}\{4-(1+s_1)(1+s_2)\}^{N/2}} \\ &\times F\left(\frac{N}{2}, \frac{\upsilon_2-\upsilon_1}{2}; \frac{n_1+\upsilon_2}{2}; \frac{(1+s_1)(1-s_2)}{4-(1+s_1)(1+s_2)}\right), \end{aligned} \tag{13}$$

$t_1, t_2 \in \mathcal{R}$. The marginal densities associated with (15) are, by construction, Jones & Faddy (2003) skew t distributions with parameters $\{n_1, \upsilon_1\}$ and $\{n_2, \upsilon_2\}$, respectively. Symmetric Student t marginals arise for $n_1 = \upsilon_1$, in which case the d.f. are n_1, and for $n_2 = \upsilon_2$. A little effort shows that this reduces to formula (6) of Jones (2001) when $\upsilon_1 = \upsilon_2$. See Fig. 1 of Jones (2001) for plots of three bivariate t distributions when $\upsilon_1 = \upsilon_2$ and his Fig. 2 for two further plots when, in addition, $n_1 = n_2 = \upsilon_1$. We will not pursue the general case further here partly because it is not very tractable and partly because of flagging excitement, we imagine, on behalf of the reader!

4. A bivariate beta prior

Cole, Lee, Whitmore & Zaslavsky (1995) considered an empirical Bayes model for Markov-dependent binary sequences with randomly missing observations. A family of prior distributions was required for a pair of Markov chain transition probabilities. Such quantities clearly take values on $(0,1) \times (0,1)$ and it is natural to think in terms of marginal beta distributions for the probabilities individually and to allow dependence between them. Cole et al. employed a particular Sarmanov–Lee bivariate beta distribution (Sarmanov, 1966, Lee, 1996) and, utilising an empirical Bayes approach, decided on marginal beta distributions with parameters $\{0.30, 3.68\}$ and $\{2.36, 5.61\}$ (for their u_0 and $1 - u_1$, respectively) along with no fewer than *five* further parameters. The resulting log prior density is plotted in Fig. 4(a).

* * * Fig. 4 (a) and (b) about here * * *

The log of the bivariate beta density (10) with the same marginal distributions — which has no further parameters to be specified either empirically or subjectively — is shown in Fig. 4(b). In such a prior specification context, we feel that both the fewer parameters and the much smoother and more regular (yet, at the largest scale, similar) shape of the new bivariate beta distribution offers greater potential for practical application and defensibility.

Acknowledgement

The first author is happy to acknowledge Professor N.N. Leonenko for his useful remarks on part of this work.

References

Cole, B.F., Lee, M.L.T., Whitmore, G.A. and Zaslavsky, A.M. (1995) An Empirical Bayes Model for Markov-Dependent Binary Sequences with Randomly Missing Observations. Journal of the American Statistical Association, 90, 1364-1372.

Gradshteyn, I.S. and Ryzhik, I.M. (1994) *Table of Integrals, Series, and Products*, Jeffrey, A. (Ed.), 5th Edition. Academic Press, San Diego.

Hutchinson, T.P. and Lai, C.D. (1990) *Continuous Bivariate Distributions, Emphasising Applications.* Rumsby, Adelaide.

Joe, H. (1997) *Multivariate Models and Dependence Concepts.* Chapman and Hall, London.

Johnson, N.L. and Kotz, S. (1972) *Distributions in Statistics: Continuous Multivariate Distributions.* Wiley, New York.

Johnson, N.L., Kotz, S. and Balakrishnan, N. (1994a) *Continuous Univariate Distributions*, Vol. 1, 2nd Edition. Wiley, New York.

Johnson, N.L., Kotz, S. and Balakrishnan, N. (1994b) *Continuous Univariate Distributions*, Vol. 2, 2nd Edition. Wiley, New York.

Jones, M.C. (2001) Multivariate t and Beta Distributions Associated with the Multivariate F Distribution. Metrika, 54, 215-231.

Jones, M.C. (2002) A Dependent Bivariate t Distribution with Marginals on Different Degrees of Freedom. Statistics & Probability Letters, 56, 163-170.

Jones, M.C. and Faddy, M.J. (2003) A Skew Extension of the t Distribution, with Applications. Journal of the Royal Statistical Society, Series B, 65, 159-174.

Kimball, A.W. (1951) On Dependent Tests of Significance in the Analysis of Variance. Annals of Mathematical Statistics, 22, 600-602.

Kotz, S. and Nadarajah, S. (2004) *Multivariate t Distributions and Their Applications*. Cambridge University Press, Cambridge.

Lee, M.L.T. (1996) Properties and Applications of the Sarmanov Family of Distributions. Communications in Statistics – Theory and Methods, 25, 1207-1222.

Libby D.L. and Novick, M.R. (1982) Multivariate Generalized Beta Distributions With Applications to Utility Assessment. Journal of Educational Statistics, 7, 271-294.

Maplesoft (2005) Maple 10 software, `www.maplesoft.com`.

Nelsen, R.B. (2005) *An Introduction to Copulas*, 2nd Edition. Springer, London.

Olkin, I. and Liu, R. (2003) A Bivariate Beta Distribution. Statistics and Probability Letters, 62, 407-412.

Pham-Gia, T. and Duong, Q.P. (2003) The Generalized Beta- and F-Distributions in Statistical Modelling. Mathematical and Computer Modelling, 12, 1613-1625.

Sarmanov, O.V. (1966) Generalized Normal Correlation and Two Dimensional Fréchet Classes. Doklady AN SSSR, 168, 32-35.

Zhang, S. and Jin, J. (1996) *Computation of Special Functions*. Wiley, New York.

Table 1.

Values of ρ_B for selected values of n_1, n_2, υ_1 and υ_2.

			υ_2					
n_1	n_2	υ_1	1	2	3	8	20	30
1	1	1	0.455	0.251	0.172	0.065	0.026	0.017
1	3	1	0.515	0.318	0.233	0.099	0.042	0.028
1	3	2		0.423	0.307	0.129	0.054	0.036
2	2	1	0.578	0.440	0.253	0.106	0.044	0.029
2	2	2		0.478	0.343	0.140	0.057	0.038
2	2	3			0.404	0.163	0.066	0.044
2	2	10					0.088	0.058
2	3	20					0.115	0.076
3	1	5				0.063	0.042	0.041
3	1	20					0.084	0.055
3	2	3			0.442	0.182	0.074	0.049
3	2	20					0.115	0.076
3	40	8				0.454	0.255	0.192
8	8	8				0.497	0.233	0.162
10	20	2		0.804	0.642	0.357	0.189	0.138
20	10	2		0.804	0.632	0.332	0.164	0.117
20	20	20					0.499	0.363
20	30	1	0.902	0.628	0.505	0.289	0.160	0.119
20	30	2		0.873	0.702	0.401	0.222	0.038
20	30	3			0.845	0.483	0.267	0.199
20	30	30						0.448
30	20	2		0.873	0.698	0.389	0.207	0.152
30	20	3			0.845	0.471	0.250	0.183
30	30	30						0.500
40	3	8				0.454	0.202	0.138
40	100	1	0.953	0.670	0.545	0.326	0.196	0.154
40	100	2		0.938	0.762	0.546	0.274	0.215
50	100	1	0.959	0.675	0.548	0.328	0.197	0.155
50	100	2		0.946	0.769	0.460	0.276	0.217

Fig. 1. Contour plot of density (4) when $n_1 = 10, n_2 = 20, \upsilon_1 = 2, \upsilon_2 = 3$.

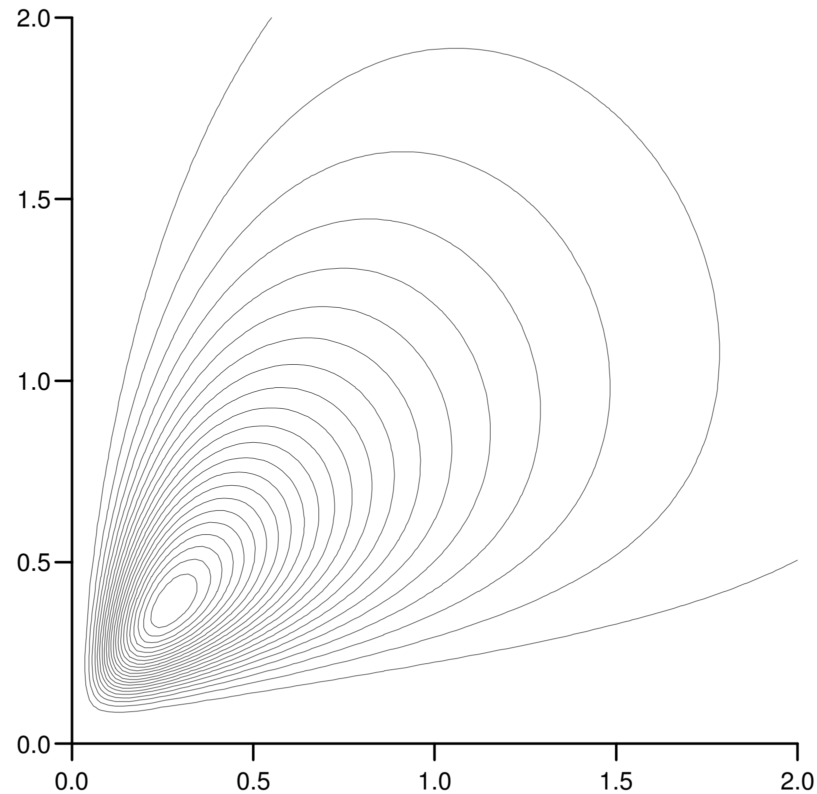

Fig. 2. Contour plot of density (4) when $n_1 = n_2 = 2, \upsilon_1 = 1, \upsilon_2 = 30$.

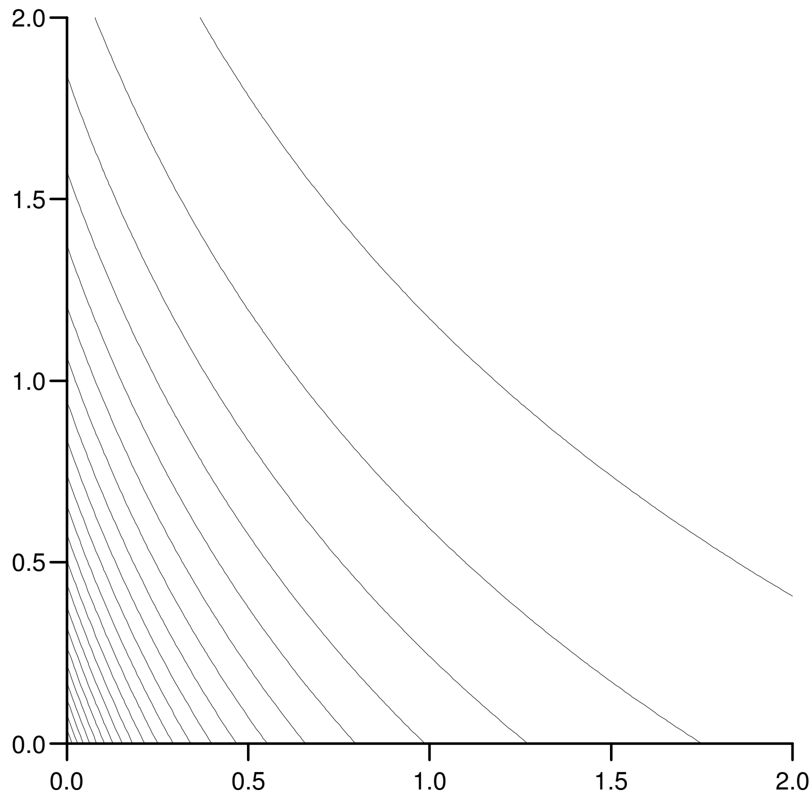

Fig. 3. Contour plot of density (10) when $n_1 = n_2 = \upsilon_1 = 8, \upsilon_2 = 20$.

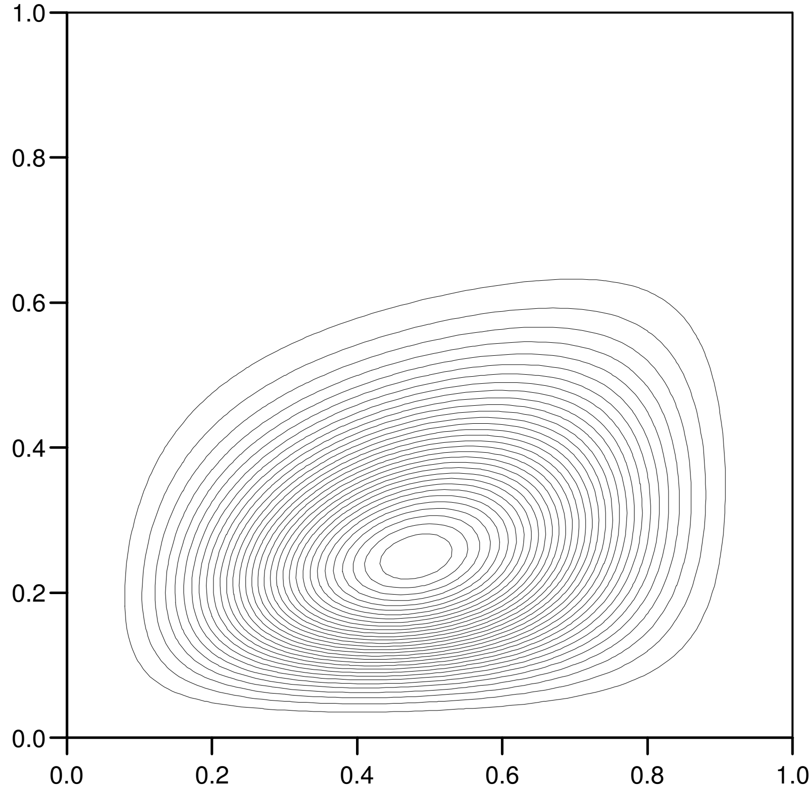

Fig. 4. Contour plot of log of bivariate beta density specified in Cole et al. (1995, pp.1365,1370).

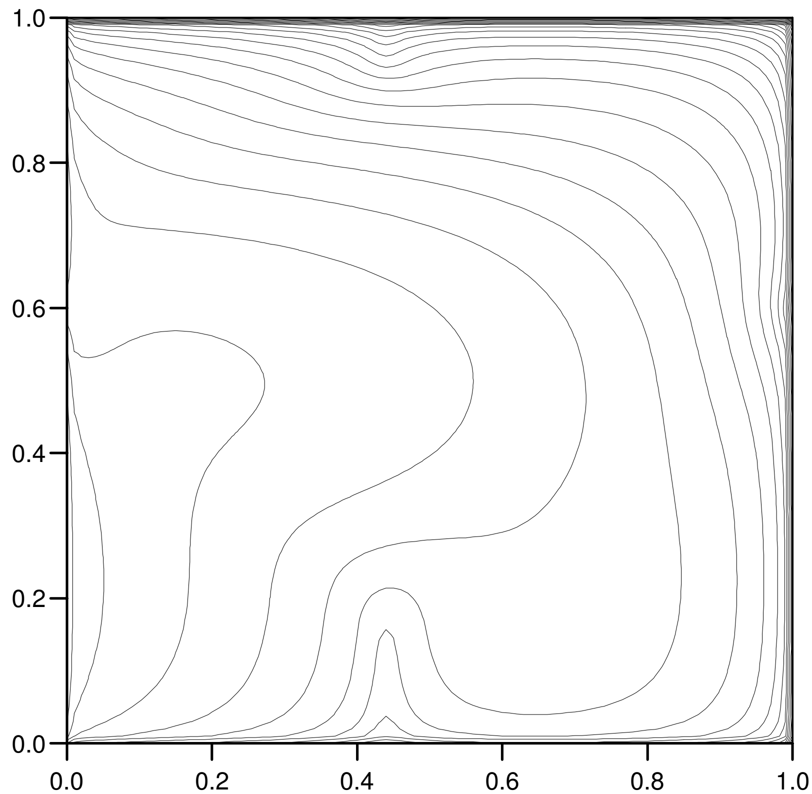

Fig. 5. Contour plot of log of density (10) when $n_1 = 0.60$, $n_2 = 4.72$, $\upsilon_1 = 7.36$, $\upsilon_2 = 11.22$. These parameter values ensure that the marginals of this density match those of the distribution in Fig. 4.