Statistics and Probability Letters 96 (2015) 54–60

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Constructions for a bivariate beta distribution

Ingram Olkin [a], Thomas A. Trikalinos [b,*]

[a] Department of Statistics, Stanford University, Stanford, CA 94305, USA

[b] Department of Health Services, Policy & Practice, Brown University, Providence, RI 02912, USA

ARTICLE INFO

Article history:
Received 27 February 2014
Received in revised form 10 September 2014
Accepted 11 September 2014
Available online 21 September 2014

Keywords:
Bivariate beta distribution
Bayesian analysis
Dirichlet distribution
Bivariate families
Hypergeometric functions

ABSTRACT

We provide a new bivariate distribution with beta marginal distributions, positive probability over the unit square, and correlations over the full range. We discuss its extension to three or more dimensions.

1. Introduction

The univariate beta distribution and its bivariate extension are basic distributions that have been used to model data in various fields. For example, in population genetics, Wright (1937) showed that the beta distribution arises from a diffusion equation describing allele frequencies in finite populations, and thus the beta distribution is used to model proportions of alleles at a specific locus. By extension, a bivariate beta distribution may be appropriate to model proportions of alleles in evolutionary-related loci that are in linkage disequilibrium (Gianola et al., 2012). Bivariate beta distributions have also been used to model drought duration and drought intensity in climate science (Nadarajah, 2007), the proportions of diseased second premolars and molars in dentistry (Bibby and Væth, 2011), tree diameter and height in forestry (Hafley and Schreuder, 1977; Li et al., 2002; Wang and Rennolls, 2007), soil strength parameters ('cohesion' and 'coefficient of friction') in civil engineering (A-Grivas and Asaoka, 1982), retinal image recognition measurements in biometry (Adell et al., 2012), decision maker utilities in multi-attribute utility assessment (Libby and Novick, 1982), and joint readership of two monthly magazines (see second example in Danaher and Hardie (2005)).

A second role for the bivariate beta distribution is that of a prior for two correlated proportions. Xie et al. (2013) elicited a bivariate beta distribution from experts to serve as a prior distribution in the analysis of clinical trial data; and Oleson (2010) used a bivariate beta distribution as a prior distribution to correlated proportions when analyzing single-patient trials.

The well known Dirichlet density is a multivariate generalization of the beta distribution, but it is restricted to a lower dimensional simplex. Thus it is not an appropriate model for examples such as the above. Instead, we seek a bivariate distribution with a positive probability on the unit square $(0, 1)^2$, beta marginal distributions, and correlation over the full range.

Balakrishnan et al. (2008) and Nelsen (2006) discuss an array of techniques for constructing continuous bivariate distributions. We selectively outline some, to contextualize relevant literature.

* Corresponding author.
E-mail addresses: olkin@stanford.edu (I. Olkin), thomas_trikalinos@brown.edu (T.A. Trikalinos).

http://dx.doi.org/10.1016/j.spl.2014.09.013
0167-7152/

1.1. General families

First, one can use general families of bivariate distributions that separate the bivariate structure from the marginal distributions. Examples are the Farlie–Gumbel–Morgenstern, Plackett, Mardia, and Sarmanov families. For a fuller discussion of these and other families see Joe (1997), Kotz et al. (2000), or Balakrishnan et al. (2008).

1.2. Variable-in-common and transformation-based constructions

An alternative is explicated by Libby and Novick (1982), who construct a multivariate generalized beta distribution starting from independent gamma variates G_0, G_1, G_2, with parameters α_0 and β_0, α_1 and β_1, and α_2 and β_2 respectively. Then the joint density of

$$X = G_1/(G_1 + G_0), \qquad Y = G_2/(G_2 + G_0) \tag{1.1}$$

is a generalized beta distribution with density

$$f(x, y) = \frac{1}{B(\alpha_0, \alpha_1, \alpha_2)} \frac{\lambda_1^{\alpha_1} x^{\alpha_1 - 1}(1-x)^{-(\alpha_1+1)}\, \lambda_2^{\alpha_2} y^{\alpha_2-1}(1-y)^{-(\alpha_2+1)}}{\left[1 + \lambda_1 x/(1-x) + \lambda_2 y/(1-y)\right]^{\alpha_0+\alpha_1+\alpha_2}}, \tag{1.2}$$

for $0 < x, y < 1$; $\alpha_i, \beta_i, > 0$ for $i = 0, 1, 2$, and $\lambda_i = \beta_i/\beta_0$ for $i = 1, 2$. In (1.2) $B(\alpha_1, \dots, \alpha_k) = \prod \Gamma(\alpha_i)/\Gamma(\sum \alpha_i)$ is the generalized beta function. When $\lambda_i = 1$, the density (1.2) reduces to a bivariate beta distribution with three, rather than five parameters:

$$f(x, y) = \frac{1}{B(\alpha_0, \alpha_1, \alpha_2)} \frac{x^{\alpha_1-1}(1-x)^{\alpha_0+\alpha_2-1} y^{\alpha_2-1}(1-y)^{\alpha_0+\alpha_1-1}}{(1-xy)^{\alpha_0+\alpha_1+\alpha_2}}. \tag{1.3}$$

Jones (2002) obtains the density (1.3) starting from a multivariate F distribution. Olkin and Liu (2003) obtain it independently using a multiplicative or logarithmically additive construction scheme analogous to that in (1.1). It is obvious from the construction (1.1) that X and Y have a positive correlation in [0, 1]. In fact, Olkin and Liu (2003) note that the stronger property of positive quadrant dependence holds for (1.3), so that the probability that the bivariate random variables are simultaneously large (small) is at least as large as if they were independent.

In addition, Nadarajah and Kotz (2005) note that if U, V, W are beta random variates with special relations among the parameters, then $(X = UW,\ Y = U)$ and $(X = UW,\ Y = VW)$ will have a bivariate beta distribution.

Finally, Arnold and Ng (2011) construct a flexible family of bivariate beta distributions starting from five independent gamma variates G_1 through G_5, with positive shape parameters α_1 through α_5, respectively, and common scale parameter 1. They define the pair X, Y

$$X = \frac{G_1 + G_3}{G_1 + G_3 + G_4 + G_5}, \qquad Y = \frac{G_2 + G_4}{G_2 + G_3 + G_4 + G_5}, \tag{1.4}$$

which has a density with beta distribution marginals. The joint density of (X, Y) from construction (1.4) does not have a closed form and must be calculated numerically. It contains the distribution in (1.3) as a special case, namely when $\alpha_3 = \alpha_4 = 0$. Contrary to the other constructions, it allows correlations throughout the full range. However its extension to three or more dimensions is cumbersome.

1.3. Generalization of existing bivariate beta densities

Nadarajah (2007) provides a slight modification to (1.3), in which the denominator $(1 - xy)^{\alpha_0+\alpha_1+\alpha_2}$ becomes $(1 - xy\delta)^{\alpha_0+\alpha_1+\alpha_2}$. In the worked example the estimate of δ was close to 1, suggesting that this parameter had a small effect in fitting the density.

1.4. Constructions based on order statistics

Another construction may be via order statistics. If $X_{(1)} \le \cdots \le X_{(n)}$ are the order statistics from a uniform distribution on [0, 1] then the distribution of a spacing

$$w_{rs} = X_{(s)} - X_{(r)}, \quad s \ge r,$$

has a beta distribution,

$$f(w_{rs}) = \frac{w_{rs}^{d-1}(1 - w_{rs})^{n-d}}{B(d, n-d)}, \quad d = s - r.$$

This suggests that the joint distribution of w_{rs} and w_{ts} is a bivariate beta distribution. We have not followed this line of inquiry but note only that it may lead to some novel results. For further discussion of order statistics see David and Nagaraja (2003).