Journal of Statistical Planning and Inference 139 (2009) 2559 – 2570

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

On the Dirichlet distributions on symmetric matrices

Mohamed Ben Farah, Abdelhamid Hassairi*

Laboratory of Probability and Statistics, Sfax University, B.P.802, Sfax, Tunisia

ARTICLE INFO

Article history:
Received 8 September 2008
Received in revised form
5 November 2008
Accepted 25 November 2008
Available online 7 December 2008

Keywords:
Wishart distribution
Beta distribution
Dirichlet distribution
Spherical Fourier transform

ABSTRACT

In this paper, we introduce a generalization of the Dirichlet distribution on symmetric matrices which represents the multivariate version of the Connor and Mosimann generalized real Dirichlet distribution. We establish some properties concerning this generalized distribution. We also extend to the matrix Dirichlet distribution a remarkable characterization established in the real case by Darroch and Ratcliff.

1. Dirichlet distribution on symmetric matrices

The Dirichlet distribution is one of the most important distributions in statistics because it offers high flexibility for modeling data and because it is conjugate prior of the parameters of the multinomial distribution in Bayesian statistics. It has also many nice probabilistic characterization properties in different directions, such as independence properties, zero regression or conditional distributions. The Dirichlet distribution is usually considered as the multivariate generalization of the beta distribution. However, with regard to the way in which the real beta and Dirichlet distributions are constructed with the gamma distribution, it seems more natural to consider that the multivariate generalizations of these distributions are the ones constructed with the multivariate version of the gamma distribution, that is with the Wishart distribution (see Casalis and Letac, 1996) or more generally with the Riesz distribution on the cone of positive definite symmetric matrices (see Hassairi and Lajmi, 2001). We also mention that Connor and Mosimann (1969) have introduced the so-called generalized Dirichlet distribution which is a more general version of the ordinary Dirichlet distribution with real margins. The extension of the properties of the real beta and real Dirichlet distributions to their matrix versions involves in general some conceptual and technical difficulties, however, many interesting results in this direction have been established. The most famous is the Olkin and Rubin (1962) characterization of the Wishart distribution which is a generalization of the Lukacs (1955) characterization of the gamma distribution. Also Ben Farah and Hassairi (2007) have extended to the matrix variate Dirichlet distribution a remarkable characterization of the real Dirichlet distribution proved in Chamayou and Letac (1994). Another remarkable characterization of the real Dirichlet distribution is due to Darroch and Ratcliff (1971).

It says that if $X_1, \dots, X_n$ are positive random variables with continuous probability density functions (pdfs) such that $\sum_{i=1}^n X_i < 1$, then the vector $(X_1, \dots, X_n)$ has a Dirichlet distribution if and only if, for every $i \in \{1, \dots, n\}$, $X_i/1 - \sum_{j \neq i} X_j$ is independent of the set $\{X_j; j \neq i\}$.

In the present paper, we define a matrix version of the Connor and Mosimann's generalized Dirichlet distribution and we establish some results concerning this distribution. This involves the matrix beta distributions of the first and the second kind.

* Corresponding author.
E-mail address: Abdelhamid.Hassairi@fss.rnu.tn (A. Hassairi).

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2008.11.014

We also extend the Darroch and Ratcliff characterization to the Dirichlet distribution on symmetric matrices. For the proof, we establish some analytic properties of functions defined on symmetric matrices and related to determinants, and we use the fact that a real function, defined on the cone of positive definite matrices, which is invariant by the orthogonal group is characterized by its spherical Fourier transform.

We will denote by V the space of real symmetric $r \times r$ matrices, and by Ω the cone of positive definite elements of V. The identity matrix is denoted by e, the determinant of an element x of V by $\Delta(x)$ and its trace by $\mathrm{tr}(x)$. For the definition of the Dirichlet distribution on symmetric matrices, we need to define a notion of "quotient" based on a division algorithm on matrices. In fact, there is not a single way to define a quotient of two matrices, and for each definition of a quotient, corresponds a definition of a beta distribution and of a Dirichlet distribution (see Olkin and Rubin, 1964). For instance if y is a positive definite matrix, we can for example, write $y = y^{1/2}y^{1/2}$ and define the ratio of an element x of V by y as $y^{-1/2}xy^{-1/2}$. We can also use the Cholesky decomposition, that is the fact that an element y of Ω can be written in a unique manner as $y = tt'$, where t is a lower triangular matrix with strictly positive diagonal and t' is its transpose. In this case, for an element x in V, we set $y(x) = txt'$, which may be seen as the product of x by y, and we define the "quotient" of x by y as $y^{-1}(x) = t^{-1}xt'^{-1}$. This algorithm is the one which suits us in the present work, because it has a nice property which allows the calculation of the generalized power of the product $y(x) = txt'$ and of the quotient $y^{-1}(x)$, when y and x are both in Ω.

Consider the absolutely continuous Wishart distribution on Ω, with shape parameter $p > (r-1)/2$ and scale parameter e,

$$W_p(\mathrm{d}x) = \frac{1}{\Gamma_\Omega(p)} \exp(-\mathrm{tr}(x))\Delta(x)^{p-(r+1)/2}\mathbf{1}_\Omega(x)\,\mathrm{d}x,$$

where $\Gamma_\Omega(.)$ is the multivariate gamma function defined by

$$\Gamma_\Omega(p) = (2\pi)^{r(r-1)/4}\prod_{k=1}^{r}\Gamma\left(p - \frac{k-1}{2}\right).$$

Let $p_1, \ldots, p_{n+1}$ be in $((r-1)/2, +\infty)$ and let $Y_1, \ldots, Y_{n+1}$ be independent random matrices in V with Wishart distributions $W_{p_1}, \ldots, W_{p_{n+1}}$, respectively. Define

$$S = Y_1 + \cdots + Y_{n+1} \quad \text{and} \quad X = (X_1, \ldots, X_n) = (S^{-1}(Y_1), \ldots, S^{-1}(Y_n)).$$

Then the distribution of X is called the Dirichlet distribution on V with parameters $(p_1, \ldots, p_{n+1})$ and is denoted by $D_{(p_1,\ldots,p_{n+1})}$. Its pdf is given by

$$f(x_1, \ldots, x_n) = \frac{\Gamma_\Omega(\sum_{i=1}^{n+1} p_i)}{\prod_{i=1}^{n+1}\Gamma_\Omega(p_i)}\prod_{i=1}^{n}\Delta(x_i)^{p_i-(r+1)/2}\Delta\left(e - \sum_{i=1}^{n} x_i\right)^{p_{n+1}-(r+1)/2}\mathbf{1}_{T_n}(x_1, \ldots, x_n),$$

where $T_n = \{(x_1, \ldots, x_n) \in \Omega^n; \sum_{i=1}^{n} x_i \in \Omega \cap (e - \Omega)\}$.

When $n = 1$, that is when we have two independent Wishart random matrices X and Y with the same scale parameter e and respective shape parameters p_1 and p_2, the random matrix $Z = (X+Y)^{-1}(X)$ has the distribution

$$(B_\Omega(p_1, p_2))^{-1}\Delta(z)^{p_1-(r+1)/2}\Delta(e-z)^{p_2-(r+1)/2}\mathbf{1}_{\Omega\cap(e-\Omega)}(z)\,\mathrm{d}z, \tag{1.1}$$

where $B_\Omega(.)$ is the bivariate beta function defined by

$$B_\Omega(p_1, p_2) = \frac{\Gamma_\Omega(p_1)\Gamma_\Omega(p_2)}{\Gamma_\Omega(p_1 + p_2)}.$$

This distribution is called the beta distribution of the first kind on V with parameters (p_1, p_2) and is denoted by $\beta^{(1)}_{p_1,p_2}$. It is easy to see that, for $\sigma \in \Omega$, $(\sigma^{-1}(X_1), \ldots, \sigma^{-1}(X_n))$ has the Dirichlet distribution $D_{(p_1,\ldots,p_{n+1})}$ if and only if $(X_1, \ldots, X_n)$ has the pdf

$$f(x_1, \ldots, x_n) = \frac{\Gamma_\Omega(\sum_{i=1}^{n+1} p_i)}{\prod_{i=1}^{n+1}\Gamma_\Omega(p_i)}\prod_{i=1}^{n}\frac{\Delta(x_i)^{p_i-(r+1)/2}\Delta(\sigma - \sum_{i=1}^{n} x_i)^{p_{n+1}-(r+1)/2}}{\Delta(\sigma)^{\sum_{i=1}^{n+1} p_i-(r+1)/2}}\mathbf{1}_{T_n(\sigma)}(x_1, \ldots, x_n),$$

where $T_n(\sigma) = \{(x_1, \ldots, x_n) \in \Omega^n, \sum_{i=1}^{n} x_i \in \Omega \cap (\sigma - \Omega)\}$.

To close this section, we introduce the orthogonal group K of V, the notion of generalized power, and the spherical Fourier transform of a function defined on Ω and invariant by K. Specifically, for an orthogonal real $r \times r$ matrix a, corresponds the automorphism of V defined by $x \mapsto axa^{-1}$. The set of such automorphisms is a group called the orthogonal group of V and denoted by K. Now for $x = ((x_{ij})_{1 \leqslant i,j \leqslant n})$ in Ω and $1 \leqslant k \leqslant r$, let $\Delta_k(x)$ denote the principal minor of order k of x, that is the determinant of the sub-matrix $((x_{ij})_{1 \leqslant i,j \leqslant k})$. The generalized power of x is defined for $s = (s_1, \ldots, s_r)$ in $\mathbb{R}^r$, by

$$\Delta_s(x) = \Delta_1(x)^{s_1-s_2}\Delta_2(x)^{s_2-s_3}\ldots\Delta_r(x)^{s_r}.$$