Journal of Multivariate Analysis 102 (2011) 948–957

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

On Pearson–Kotz Dirichlet distributions

N. Balakrishnan [a], E. Hashorva [b,*]

[a] Department of Mathematics and Statistics, McMaster University, 1280 Main Street West, Hamilton, Ontario, Canada L8S 4K1

[b] Department of Actuarial Science, Faculty of Business and Economics, University of Lausanne, Bâtiment Extranef, UNIL-Dorigny, 1015 Lausanne, Switzerland

ARTICLE INFO

Article history:
Received 23 November 2009
Available online 22 February 2011

AMS subject classifications:
primary 60F05
secondary 60G70

Keywords:
Pearson–Kotz Dirichlet distribution
Dirichlet distribution
Kotz type distribution
Elliptically symmetric distribution
t-distribution
Beta distribution
Beta distribution of second
Conditional limiting theorem
Conditional excess distribution

ABSTRACT

In this paper, we discuss some basic distributional and asymptotic properties of the Pearson–Kotz Dirichlet multivariate distributions. These distributions, which appear as the limit of conditional Dirichlet random vectors, possess many appealing properties and are interesting from theoretical as well as applied points of view. We illustrate an application concerning the approximation of the joint conditional excess distribution of elliptically symmetric random vectors.

1. Introduction

Let $\boldsymbol{Z} = (Z_1, \ldots, Z_k)^\top, k \geq 2$, be a mean zero Gaussian random vector with independent components with $\mathbf{Var}\{Z_i\} = 1, 1 \leq i \leq k$. If λ/V, where $\lambda \in (0, \infty)$, is a gamma distributed random variable with parameters $(\theta/2, 1/2)$, with $\theta \in (0, \infty)$ being independent of $\boldsymbol{Z}$, then the random vector $\boldsymbol{X} = (X_1, \ldots, X_k)^\top$ defined by

$$X_i \stackrel{d}{=} \sqrt{V} Z_i, \quad i = 1, \ldots, k$$

has a generalized (Student) t-distribution with parameters λ and θ, as introduced by Arellano-Valle and Bolfarine [2]. Here, $^\top$ and $\stackrel{d}{=}$ denote the transpose sign and equality of the distribution functions, respectively. It is well-known [see [29]] that the Gaussian random vector $\boldsymbol{Z}$ is a canonical example of a spherically symmetric random vector with radial representation

$$\boldsymbol{Z} \stackrel{d}{=} R_* \boldsymbol{\mathcal{U}}, \tag{1}$$

where $\boldsymbol{\mathcal{U}}$ is uniformly distributed on the unit sphere of $\mathbb{R}^k$ and is independent of the positive random radius R_* such that R_*^2 is chi-square distributed with k degrees of freedom. Consequently, a random vector $\boldsymbol{X}$ with the generalized t-distribution has the radial representation

$$\boldsymbol{X} \stackrel{d}{=} R \boldsymbol{\mathcal{U}}, \quad \text{with } R = \sqrt{\lambda W_1 / W_2}, \tag{2}$$

* Corresponding author.
E-mail addresses: bala@mcmaster.ca (N. Balakrishnan), enkelejd.hashorva@unil.ch, enkelejd@stat.unibe.ch (E. Hashorva).

0047-259X/$ – see front matter
doi:10.1016/j.jmva.2011.01.012

where W_1 and W_2 are two independent Gamma random variables with parameters $(k/2, 1/2)$ and $(\theta/2, 1/2)$, respectively, being further independent of $\boldsymbol{\mathcal{U}}$.

The generalized t-distributions, which include as a special case the well-known (Student) t-distribution, possess many interesting distributional properties and flexibility, and are therefore useful for statistical modeling. An elaborate discussion on various t-distributions has been provided by Kotz and Nadarajah [30]. One may also refer to [1,8,35,3,9,28,6,31–33,36] for related discussions.

The radial representation in (2) is with respect to L_2-norm. Gupta and Song [14] introduced Person Type VII distributions with respect to L_p-norm; see also [15,39]. This class of distributions includes the generalized t-distributions for some special choice of the parameters.

A further generalization, motivated by the aforementioned model, is the Pearson Type VII Dirichlet distribution discussed by Hashorva et al. [21]. One important feature of these distributions is that the marginal distributions need not be identical, which is the case for the generalized t-distributions. Indeed, the Pearson Type VII Dirichlet distribution is a natural generalization of the t-distribution motivated as follows: Taking $\boldsymbol{\mathcal{U}}$ in (2) as a symmetrized Dirichlet random vector with parameters $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_k)^\top \in (0, \infty)^k$ and $p \in (0, \infty)$ (see Definition 1 below), $\boldsymbol{X}$ in the stochastic representation in (2) becomes a L_p Dirichlet random vector. For instance, a natural extension of the generalized t-distribution is achieved if in (2) is considered a symmetrized Dirichlet distribution for $\boldsymbol{\mathcal{U}}$, and W_1 independent of W_2 are such that W_1 is gamma distributed with parameters $(\sum_{i=1}^k \alpha_i, 1/p)$ and W_2 is gamma distributed with parameters $(\theta, 1/p)$, with $\theta \in (0, \infty)$. This choice of W_1 and W_2 is motivated by the beta-independent splitting property of Dirichlet distributions [see (27) in Appendix]. For such $\boldsymbol{X}$, the stochastic representation in (1) holds as well, with $\boldsymbol{Z}$ being a standard Kotz Type I Dirichlet random vector, provided that $\boldsymbol{\mathcal{U}}, W_1, W_2$ are mutually independent. We, therefore, propose to refer to the random vector $\boldsymbol{X}$ with stochastic representation in (2) as a Pearson–Kotz Dirichlet random vector.

The class of Pearson–Kotz Dirichlet random vectors inherits the main distributional properties of the t-random vectors, and at the same time provides for more flexible models since the components of Pearson–Kotz Dirichlet random vectors need not have the same distribution function.

The principal goal of this paper is the investigation of the main asymptotic properties of Pearson–Kotz Dirichlet random vectors. We start with the derivation of the marginal and conditional distributions. These are useful for our main result in which we show that the Pearson–Kotz Dirichlet random vectors possess a crucial asymptotic property. Specifically, in Theorem 2, we provide a conditional limit result for regularly varying Dirichlet random vectors. The importance of our result is that even though the distribution function of Dirichlet random vectors can be unknown, in a conditional framework, the approximating random vector has a known distribution function, viz., the Pearson–Kotz Dirichlet distribution. This fact is of interest for statistical modeling. We finally illustrate an application concerning the approximation of joint conditional excess distribution of elliptically symmetric random vectors.

The rest of this paper is organized as follows. In Section 2, we introduce some notation and preliminary results. In Section 3, we present some basic distributional properties of the Pearson–Kotz Dirichlet random vectors, which are then used in the derivation of our main results in Section 4 concerning the asymptotic approximation of conditioned Dirichlet random vectors. For elliptically symmetric random vectors we show further an asymptotic approximation of the joint conditional excess distribution, while some related results and the proofs omitted in the next sections are presented in the Appendix.

2. Preliminaries

In this section, we first present some notation and then review some key facts about Dirichlet random vectors. Let I be a non-empty subset of $\{1, \ldots, k\}$, $k \geq 2$. The number of elements in I is denoted by $|I|$. In the following A denotes a $k \times k$ real matrix. For any k-dimensional vector $\boldsymbol{x} = (x_1, \ldots, x_k)^\top \in \mathbb{R}^k$, let $\boldsymbol{x}_I := (x_i, i \in I)^\top$ with $^\top$ the transpose sign. Similarly, if $J := \{1 \cdots k\} \setminus I$ is non-empty we define submatrices $A_{II}, A_{IJ}, A_{JI}, A_{JJ}$ of A, and write A_{II}^{-1} for the inverse of A_{II} if it possesses an inverse. For k-dimensional vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, let us define

$$\boldsymbol{x} + \boldsymbol{y} := (x_1 + y_1, \ldots, x_k + y_k),$$
$$\boldsymbol{x} > \boldsymbol{y} \quad \text{if } x_i > y_i \quad \forall\, i = 1, \ldots, k,$$
$$\boldsymbol{x}\boldsymbol{y} := (x_1 y_1, \ldots, x_k y_k)^\top,$$
$$\|\boldsymbol{x}_I\|_p := \Big(\sum_{i \in I} |x_i|^p\Big)^{1/p}, \qquad \|\boldsymbol{x}_I\|_{A,p} := \|A_{II}^{-1}\boldsymbol{x}_I\|_p \quad \text{for } p \in (0, \infty).$$

Further, we shall denote by $beta(a, b)$ and $gamma(a, b)$ the distribution functions of a beta and a gamma random variable with positive parameters a, b with density functions

$$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}, \quad x \in (0, 1), \quad \text{and} \quad \frac{b^a}{\Gamma(a)} x^{a-1} \exp(-bx), \quad x \in (0, \infty),$$

respectively, where $\Gamma(\cdot)$ is the Euler gamma function.

If a random vector $\boldsymbol{Z}/c, c \in (0, \infty)$ has the distribution function Q, we will denote it by $\boldsymbol{Z} \sim cQ$. Note that if $X \sim beta(b, a)$, then $Y := (1/X) - 1$ has the so-called beta distribution of second kind with parameters (a, b), which is denoted by $beta_2(a, b)$, see [26].