SOME PROPERTIES OF A GENERALIZED TYPE-1 DIRICHLET DISTRIBUTION

E.V. Mayamol

1. INTRODUCTION

In this paper we study properties of some new generalizations of Dirichlet densities. Johann Peter Gustav Lejeune Dirichlet in (1839) evaluated an integral which later gave rise to the well-known probability density which now bears his name. Wilks (1962) was the first to use the terminology “Dirichlet Distributions” for random variables which have the density function in (1). Dirichlet distribution is the generalization of beta distribution. Standard real type-1 and type-2 beta distributions are extended to standard type-1 and type-2 Dirichlet distributions. These Dirichlet distributions are further extended in various directions. The standard Dirichlet distribution can be found in textbooks on mathematical statistics. The standard real type-1 Dirichlet density with parameters \((\alpha_1, \ldots, \alpha_k; \alpha_{k+1})\) is given by

\[
f_{\alpha}(x_1, \ldots, x_k) = c_k \begin{cases} x_1^{\alpha_1-1} \cdots x_k^{\alpha_k-1} (1 - x_1 - \cdots - x_k)^{-\alpha_{k+1}-1} & \text{for } x_i \geq 0, i = 1, \ldots, k, \\
0 & \text{elsewhere}
\end{cases}
\]

In statistical problems the parameters are real and hence the parameters are assumed to be real. But the integrals and corresponding results will hold for complex parameters. In that case, for example \(\alpha_j > 0\) is to be replaced by \(\Re(\alpha_j) > 0\) where \(\Re(.)\) denotes the real part of \((.)\). Complex parameters are needed if inverse Mellin transform is used to establish the uniqueness of the corresponding densities.

The standard real type-2 Dirichlet density with parameters \((\alpha_1, \ldots, \alpha_k; \alpha_{k+1})\) is given by
This distribution is also known as inverted Dirichlet distribution. In (1) and (2) the normalizing constant \( c_k \) is the same as the one given in (3), and it is evaluated by integrating out variables one at a time.

\[
f_2(x_1, \ldots, x_k) = c_k \begin{cases} 
\frac{x_1^{\alpha_1-1} \cdots x_k^{\alpha_k-1} (1 + x_1 + \cdots + x_k)^{-(\alpha_1 + \cdots + \alpha_k + 1)}}{i = 1, \ldots, k, \quad \alpha_j > 0, j = 1, \ldots, k + 1} \\
0 \quad \text{elsewhere}
\end{cases}
\]

\[
f_2(x_1, \ldots, x_k) = c_k \begin{cases} 
\frac{x_1^{\alpha_1-1} \cdots x_k^{\alpha_k-1} (1 + x_1 + \cdots + x_k)^{-(\alpha_1 + \cdots + \alpha_k + 1)}}{i = 1, \ldots, k, \quad \alpha_j > 0, j = 1, \ldots, k + 1} \\
0 \quad \text{elsewhere}
\end{cases}
\]

The Dirichlet density is used in a variety of contexts. Here we discuss mostly the type-1 Dirichlet density and its generalizations and hence when we state Dirichlet density it will mean a type-1 Dirichlet density. It has found applications in order statistics, reliability and survival analysis, Bayesian analysis etc. Use of Dirichlet distribution for approximating multinomial probabilities may be seen from Johnson (1960). Mosimann (1962, 1963) obtained several characterizations of Dirichlet density and utilized this distribution as the prior for the parameters of the multinomial and negative multinomial distributions. Spiegelhalter et al. (1994) used Dirichlet density as a prior model to study the frequencies of congenital heart disease. Applications of Dirichlet distribution in modeling the buying behavior was discussed by Goodhardt, Ehrenberg and Chatfield (1984). Its application in the distribution of sparse and crowded cells in occupancy models were considered by Sobel and Uppuluri (1974). Lange (1995) used Dirichlet distribution to model the contributions from different ancestral populations in computing forensic match probabilities. Applications of Dirichlet models in random division and other geometrical probability problems may be seen from Mathai (1999). Several applications involving linear combinations of the components of a Dirichlet random vector are pointed out by Provost and Cheong (2000). Generalized Dirichlet in Bayesian analysis may be seen from Wong (1998).

In this paper, Section 2 briefly introduces different generalized models of the Dirichlet density. In Section 3, we discuss different structural representations of \( x_1 \) in one of the generalized Dirichlet models and its applications to geometrical probability problems are pointed out. In Section 4, multiple regression and Bayesian estimates are given.

2. GENERALIZATIONS OF THE DIRICHLET MODEL

There are different generalizations of the Dirichlet distributions in the literature, some of them are reviewed here. Connor and Mosimann (1969) introduced a generalization of the Dirichlet density based on the neutrality principle of pro-
portions. Let \((x_1, \ldots, x_k)\) be the vector of proportions and \(x_i, x_2/(1-x_1), \ldots, (x_k/(1-x_1-\cdots-x_{k-1}))\) be independently beta distributed with parameters \((\alpha_i, \beta_i)\), \(i = 1, \ldots, k\). Then \((x_1, \ldots, x_k)\) has the joint density function,

\[
g_1(x_1, \ldots, x_k) = \left\{ \prod_{i=1}^{k} [B(\alpha_i, \beta_i)]^{-1} \left[ x_j^{\alpha_i - 1} \left( 1 - \sum_{j=1}^{i-1} x_j \right)^{\beta_i - (\alpha_i + \beta_i)} \right] \left[ 1 - \sum_{j=1}^{k} x_j \right]^{\beta_i - 1} \right\} (4)
\]

where

\[
B(\alpha_i, \beta_i) = \frac{\Gamma(\alpha_i)\Gamma(\beta_i)}{\Gamma(\alpha_i + \beta_i)}, \quad \alpha_i > 0, \beta_i > 0.
\]

Wong (1998) studied this generalized Dirichlet distribution (4) and showed that it has a more general covariance function than the Dirichlet distribution. As well as the Dirichlet distribution, Wong has also shown that the generalized Dirichlet distribution (4) is conjugate to multinomial sampling. It is of interest to note that the construction (4) also has interesting applications in Bayesian non-parametric inference; see e.g. Ishwaran and James (2001).

In some problems in reliability and survival analysis, the need for considering sums of Dirichlet variables arises. Hence, Mathai (2003) introduced a general multivariate density of the following form:

\[
g_2(x_1, \ldots, x_k) = C_k x_1^{\alpha_1 - 1} \cdots x_k^{\alpha_k - 1} (1 - x_1)^{\beta_1} (1 - x_1 - x_2)^{\beta_2} \cdots \left( 1 - x_1 - \cdots - x_{k-1} \right)^{\beta_k} (1 - x_1 - \cdots - x_k)^{\beta_k - 1} (5)
\]

for \(0 \leq x_j \leq 1, 0 \leq x_j - x_{j-1} \leq \cdots \leq x_{j-1}, j = 2, \ldots, k-1, k, \alpha_j > 0, j = 1, \ldots, k + 1, \alpha_{j+1} + \cdots + \alpha_{k+1} + \beta_j + \cdots + \beta_k > 0, j = 1, \ldots, k, \) and \(g_2(x_1, \ldots, x_k) = 0 \) elsewhere. Here

\[
C_k^{-1} = \prod_{j=1}^{k} \frac{\Gamma(\alpha_j)\Gamma(\alpha_{j+1} + \cdots + \alpha_{k+1} + \beta_j + \cdots + \beta_k)}{\Gamma(\alpha_j + \cdots + \alpha_{k+1} + \beta_j + \cdots + \beta_k)}
\]

The densities in (4) and (5) can be shown to be identical.

Here we consider a generalization of the Dirichlet density to the following form:

\[
f_3(x_1, \ldots, x_k) = C_k^* x_1^{\alpha_1 - 1} \cdots x_k^{\alpha_k - 1} (x_1 + x_2)^{\beta_1} \cdots (x_1 + \cdots + x_k)^{\beta_k} \left( 1 - x_1 - \cdots - x_k \right)^{\alpha_{k+1} - 1} (6)
\]

for \(0 < x_i < 1, i = 1, \ldots, k, 0 < x_1 + \cdots + x_k < 1\). The normalizing constant \(C_k^*\) can
be obtained by changing the variables to \( u_1 = x_1, u_2 = x_1 + x_2, \ldots, u_k = x_1 + \ldots + x_k \)
and integrating variables successively. We can show that

\[
\frac{1}{C_k^*} = \frac{\Gamma(\alpha_1) \Gamma(\alpha_{k+1})}{\Gamma(\alpha_1 + \alpha_2)} \frac{\Gamma(\alpha_1 + \alpha_2 + \beta_2)}{\Gamma(\alpha_1 + \alpha_2 + \alpha_3 + \beta_2)} \ldots \times \frac{\Gamma(\alpha_1 + \ldots + \alpha_{k-1} + \beta_2 + \ldots + \beta_{k-1})}{\Gamma(\alpha_1 + \ldots + \alpha_k + \beta_2 + \ldots + \beta_{k-1})} \frac{\Gamma(\alpha_1 + \ldots + \alpha_k + \beta_2)}{\Gamma(\alpha_1 + \ldots + \alpha_{k+1} + \beta_2 + \ldots + \beta_k)}
\]

(7)

for \( \alpha_j > 0, j = 1, \ldots, k+1, \alpha_1 + \ldots + \alpha_{j+1} + \beta_2 + \ldots + \beta_j > 0, j = 1, \ldots, k \). When
\( \beta_2 = \ldots = \beta_k = 0 \) we have the Dirichlet density. A sample of the surface for
\( k = 2 \) is given in figure 1.

![Figure 1](image)

*Figure 1 – Generalized Dirichlet density with \( k = 2, \alpha_1 = 5, \alpha_2 = 3, \alpha_3 = 2, \beta_2 = 2. \)*

Here the proposed work is based on the generalized model given in (6).

3. STRUCTURAL REPRESENTATIONS OF \( x_1 \)

This section contains structural representations of \( x_1 \) when \((x_1, \ldots, x_k)\) has
the generalized Dirichlet density in (6). Let us consider the joint product moment
for some arbitrary \((t_1, \ldots, t_k)\) when \((x_1, \ldots, x_k)\) has the joint density in (6). This
can be easily seen to be the following, which can be written down by observing
the normalizing constant in (7).
Some properties of a generalized type-1 Dirichlet distribution

\[ E(x_1^{b_1}x_2^{b_2} \ldots x_k^{b_k}) = C_k^* \left\{ \frac{\Gamma(a_1 + t_1) \ldots \Gamma(a_2 + t_2) \Gamma(a_{k+1})}{\Gamma(a_1 + a_2 + t_1 + t_2)} \times \frac{\Gamma(a_1 + \alpha_2 + \beta_2 + t_1 + t_2)}{\Gamma(a_1 + \alpha_2 + \beta_2 + t_1 + t_2 + t_3)} \ldots \right. \\
\left. \frac{\Gamma(a_1 + \ldots + \alpha_{k-1} + \beta_2 + \ldots + \beta_{k-1} + t_1 + \ldots + t_{k-1})}{\Gamma(a_1 + \ldots + \alpha_k + \beta_2 + \ldots + \beta_{k-1} + t_1 + \ldots + t_k)} \times \frac{\Gamma(a_1 + \ldots + \alpha_k + \beta_2 + \ldots + \beta_k + t_1 + \ldots + t_k)}{\Gamma(a_1 + \ldots + \alpha_{k+1} + \beta_2 + \ldots + \beta_k + t_1 + \ldots + t_k)} \right\} \\
\] 

(8)

for \( a_j + t_j > 0, \ \alpha_1 + \ldots + \alpha_{j+1} + \beta_2 + \ldots + \beta_j + t_1 + \ldots + t_j > 0, j = 1, \ldots, k, \) where \( C_k^* \) is the same quantity appearing in (7).

In (8) put \( t_1 = b \) and \( t_2 = t_3 = \ldots = t_k = 0 \) then we get,

\[ E(x_1^{b_1}) = \prod_{j=1}^{k} \left\{ \frac{\Gamma(a_1 + \alpha_2 + \ldots + \alpha_j + \beta_2 + \ldots + \beta_j + b)}{\Gamma(a_1 + \alpha_2 + \ldots + \alpha_j + \beta_2 + \ldots + \beta_j)} \times \frac{\Gamma(a_1 + \ldots + \alpha_{j+1} + \beta_2 + \ldots + \beta_j)}{\Gamma(a_1 + \ldots + \alpha_{j+1} + \beta_2 + \ldots + \beta_j + b)} \right\} \\
\] 

(9)

This is the \( b^{th} \) moment of \( x_1 \). The random variable \( x_1 \) has some interesting structural properties, that are of interest in many situations. Note that (9) is nothing but the \( b^{th} \) moment of a product of independent type-1 beta random variables. That is \( E(x_1^{b_1}) = E(v_1^{b_1})E(v_2^{b_1}) \ldots E(v_k^{b_1}), \) where \( v_j \) is a type-1 beta variable with parameters \( (a_1 + \alpha_2 + \ldots + \alpha_j + \beta_2 + \ldots + \beta_j, \alpha_{j+1}), \) \( j = 1, \ldots, k. \)

Hence it is worth studying \( x_1 \) further. Theorems 1 and 2 show the transformations needed for connecting a set of type-2 beta random variables to \( x_1 \). A type-2 beta density is the following:

\[ g_3(x) = \begin{cases} 
\frac{\Gamma(a + \beta)}{\Gamma(a)\Gamma(\beta)} x^{a-1}(1+x)^{-(a+\beta)} , & 0 \leq x < \infty, \ \alpha > 0, \ \beta > 0 \\
0 & \text{elsewhere.}
\end{cases} 
\]

Theorem 1. Let \( (x_1, \ldots, x_k) \) have a generalized Dirichlet distribution (6) and \( z_1, \ldots, z_k \) be \( k \) independently distributed type-2 beta random variables with parameters \( (a_1 + \ldots + \alpha_j + \beta_2 + \ldots + \beta_j, \alpha_{j+1}), \) \( j = 1, \ldots, k. \) Consider
then we can write \( \chi_1 \) as the product of \( y_1, \ldots, y_k \) in terms of the type-2 beta variables \( \zeta_1, \ldots, \zeta_k \) and further, \( y_1, \ldots, y_k \) are independently distributed type-1 beta random variables with parameters \( (\alpha_1 + \cdots + \alpha_j + \beta_2 + \cdots + \beta_j, \alpha_{j+1}) \), \( j = 1, \ldots, k \).

**Proof.** Let us consider the joint moments \( E((y_1, \ldots, y_k)^b) \) for an arbitrary \( b \).

Since \( \zeta_1, \ldots, \zeta_k \) are independently distributed we can write

\[
E((y_1 \cdots y_k)^b) = E \left[ \prod_{j=1}^{k} \left( \frac{\zeta_j}{1 + \zeta_j} \right)^b \right].
\]

Now, integrating out over the joint density of \( \zeta_1, \ldots, \zeta_k \) we have

\[
E \left[ \prod_{j=1}^{k} \left( \frac{\zeta_j}{1 + \zeta_j} \right)^b \right] = \frac{\prod_{j=1}^{k} \Gamma(\alpha_1 + \cdots + \alpha_j + \beta_2 + \cdots + \beta_j)}{\Gamma(\alpha_{j+1}) \Gamma(\alpha_1 + \cdots + \alpha_j + \beta_2 + \cdots + \beta_j)}
\]

\[
\times \int_0^{\infty} \left( \frac{\zeta_j}{1 + \zeta_j} \right)^{\alpha_1 + \cdots + \alpha_j + \beta_2 + \cdots + \beta_j - 1} \zeta_j^{\alpha_1 + \cdots + \alpha_j + \beta_2 + \cdots + \beta_j - 1} d\zeta_j, \quad \text{for} \quad \alpha_{j+1} > 0,
\]

\( j = 1, \ldots, k, \quad \alpha_1 + \cdots + \alpha_j + \beta_2 + \cdots + \beta_j > 0, \quad j = 2, \ldots, k \)

\[
= \frac{\prod_{j=1}^{k} \Gamma(\alpha_1 + \cdots + \alpha_j + \beta_2 + \cdots + \beta_j + b)}{\Gamma(\alpha_1 + \cdots + \alpha_j + \beta_2 + \cdots + \beta_j)}
\]

\[
\times \frac{\Gamma(\alpha_1 + \cdots + \alpha_j + \beta_2 + \cdots + \beta_j)}{\Gamma(\alpha_1 + \cdots + \alpha_j + \beta_2 + \cdots + \beta_j + b)}
\]

(10)

= gamma product in (9).

Observe from the gamma product in (10) that it is of the form

\[
E(y_1^b)E(y_2^b)\cdots E(y_k^b)
\]

where \( y_1, \ldots, y_k \) are type-1 beta random variables with
parameters \((\alpha_1 + \ldots + \alpha_j + \beta_2 + \ldots + \beta_j, \alpha_{j+1})\), \(j = 1, \ldots, k\). Hence \(x_1\) has a structural representation of the form \(x_1 = y_1 \ldots y_k\) where \(y_1, \ldots, y_k\) are type-1 beta random variables.

**Theorem 2.** Let \((x_1, \ldots, x_k)\) have a generalized Dirichlet distribution (6) and \((z_1, \ldots, z_k)\) be \(k\) independently distributed type-2 beta random variables with parameters \((\alpha_{j+1}, \alpha_1 + \ldots + \alpha_j + \beta_2 + \ldots + \beta_j), j = 1, \ldots, k\). Consider

\[
\begin{align*}
y_1 &= \frac{1}{1 + z_1} \\
y_2 &= \frac{1}{1 + z_2} \\
&\vdots \\
y_k &= \frac{1}{1 + z_k}
\end{align*}
\]

then we can write \(x_1\) as the product of \(y_1, \ldots, y_k\) in terms of the type-2 beta variables \((z_1, \ldots, z_k)\) and further, \(y_1, \ldots, y_k\) are independently distributed type-1 beta random variables with parameters \((\alpha_1 + \ldots + \alpha_j + \beta_2 + \ldots + \beta_j, \alpha_{j+1})\), \(j = 1, \ldots, k\).

**Proof.** The proof is similar to that of theorem 1.

Thomas and George (2004) considered a “short memory property”. Let \(x_1, \ldots, x_k\) be such that \(0 < x_i < 1\), \(i = 1, \ldots, k\), \(0 < x_1 + \ldots + x_k < 1\). and let

\[
\begin{align*}
y_1 &= \frac{x_1}{x_1 + x_2}, y_2 = \frac{x_1 + x_2}{(x_1 + x_2 + x_3)}, \ldots, y_{k-1} = \frac{x_1 + \ldots + x_{k-1}}{(x_1 + \ldots + x_k)}, y_k = \frac{x_1 + \ldots + x_k}{1 + x_k}
\end{align*}
\]

be independently distributed. This will be called short memory property. It is shown in Thomas and George (2004) that if \(y_1, \ldots, y_k\) are independently distributed beta variables with parameters \((\alpha_1, \alpha_2), (\alpha_1 + \alpha_2 + \beta_2, \alpha_3), \ldots, (\alpha_1 + \ldots + \alpha_k + \beta_2 + \ldots + \beta_k, \alpha_{k+1})\) respectively, then the joint density of \((x_1, \ldots, x_k)\) is that given in (6). Thomas and Thannippara (2008) established a connection of the \(\Lambda\) criterion for sphericity test to this generalized Dirichlet model.
3.1. Application to Geometrical Probability

We see that $x_1$ can be represented as the product of type-1 beta random variables. Now we try to find a geometrical interpretation for $x_1$ and its applications. Let $X_j, j = 1 \ldots p$, be an ordered set of random points in the Euclidean $n$-space $\mathbb{R}^n$, $n \geq p$. Let $O$ denote the origin of a rectangular co-ordinate system. Now the $1 \times n$ vector $X_j$ can be considered as a point in $\mathbb{R}^n$. If $X_1, \ldots, X_p$ are linearly independent then the convex hull generated by these $p$-points almost surely determines a $p$-parallelotope in $\mathbb{R}^n$ with the sides $\overline{OX_1}, \ldots, \overline{OX_p}$. The random volume or $p$-content $\mathbf{v}_{p,n}$ of this random $p$-parallelotope is given by

$$\mathbf{v}_{p,n} = \left| \mathbf{X} \mathbf{X}' \right|^{1/2}$$

where $\mathbf{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix}$ is a matrix of order $p \times n$, $\mathbf{X}'$ is the transpose of $\mathbf{X}$ and $\left| (\cdot) \right|$ denotes the determinant of $(\cdot)$.

Let the joint distribution of the elements in the real $p \times n, n \geq p$, matrix $\mathbf{X}$ have an absolutely continuous distribution with the density function $f(\mathbf{X})$ which can be expressed as the function of $\mathbf{XX}'$. Let

$$f(\mathbf{X}) = g(\mathbf{XX}')$$

where $g(\mathbf{XX}') > 0$ with probability 1 on the support of $\mathbf{XX}'$. Let the rows of $\mathbf{X}$ be linearly independent so that $\mathbf{X}$ is of full rank $p$. If the density of $\mathbf{X}$ can be expressed as in (12) then $\mathbf{X}$ has a spherically symmetric distribution. The density of a spherically symmetric distribution remains invariant under orthogonal transformations, (see Mathai, 1999, 1997). Then, writing $\mathbf{S} = \mathbf{XX}'$ we have the following; where $f(\mathbf{X})$ is a density and its total integral is 1. We can convert $\mathbf{X}$ into $\mathbf{S}$ and a semi orthogonal matrix, (see Mathai, 1997) then the last step in the following line will follow.

$$1 = \int_{\mathbf{X}} f(\mathbf{X}) \, d\mathbf{X} = \int_{\mathbf{S}} g(\mathbf{S}) \, d\mathbf{X} = \frac{\pi^{p/2}}{\Gamma_p \left( \frac{n}{2} \right)} \int_{S = S' > 0} |S|^{-\frac{p+1}{2}} g(S) \, dS .$$

Now let us consider the following results that are given in Mathai (1999). Let the $p \times n, n \geq p$ real random matrix $\mathbf{X}$ of full rank $p$ have the density
Some properties of a generalized type-1 Dirichlet distribution

\[ f_1(X) = C_1 |XX'|^\alpha e^{-tr(BXX')} \text{ for } XX' > 0, \ B = B' > 0, \ \Re(\alpha) > -1, \]  

(13)

which is the density of a real rectangular matrix-variate gamma, where \( C_1 \) is the normalizing constant and \( B \) is a constant positive definite parametric matrix of order \( p \times p \).

Then

\[ E_s[|S_X|^\alpha] = |B|^\frac{\alpha}{2} \frac{\Gamma_p \left( \frac{\alpha + n + b}{2} \right)}{\Gamma_p \left( \frac{\alpha}{2} \right)} , \ b > -\frac{n}{2} + \frac{p-1}{2} - \alpha \quad \text{with } \ S_X = XX' \]

where the real matrix-variate gamma function is given by

\[ \Gamma_p(\alpha) = \pi^{\frac{p(p-1)}{4}} \Gamma(\alpha) \prod_{i=1}^{\frac{p-1}{2}} \Gamma \left( \alpha - \frac{p-1}{2} \right) , \ \Re(\alpha) > \frac{p-1}{2} . \]

Here the variable \( S_X \) has a real matrix-variate density. A general matrix-variate gamma density has the following form

\[ f(S_X) = \frac{|B|^\alpha}{\Gamma_p(\alpha)} |S_X|^\frac{\alpha - p + 1}{2} e^{-tr(BS_X)} , \]  

(14)

\( S_X = S'_X > 0, \ B = B' > 0, \ \Re(\alpha) > \frac{p-1}{2} , \) where the matrix-variate gamma variable is the \( p \times p \) real symmetric positive definite matrix \( S_X \) and the real \( p \times p \) positive definite matrix \( B \) is a constant parameter matrix, \( \alpha \) is a scalar parameter and \( \Gamma_p(\alpha) \) is the real matrix-variate gamma function. When we put \( \alpha = \frac{p + 1}{2} \)

and \( B = \frac{1}{2} \Sigma^{-1} \) in (14), then the variable \( S_X \) has Wishart density with \( n \) degrees of freedom.

\[ f(S_X) = \frac{1}{2^\frac{n}{2} \Gamma_p \left( \frac{n}{2} \right) } |S_X|^\frac{\alpha - p + 1}{2} e^{-\frac{1}{2} tr(\Sigma^{-1}S_X)} , \ S_X > 0, \ \Sigma = \Sigma' > 0, \]

where \( \Sigma \) is usually a nonsingular covariance matrix.

Let the \( p \times n, n \geq p, \) real random matrix \( Y \) of full rank \( n \) have the density

\[ f_2(Y) = C_2 |YY'|^{\alpha + \beta} e^{-tr(BYY')} \text{ for } YY' > 0, \ B = B' > 0, \ \Re(\alpha + \beta) > -1. \]  

(15)
Here $E_x[|S_Y|^p] = |\beta|^p \frac{\Gamma_p \left( \alpha + \beta + \frac{n}{2} + b \right)}{\Gamma_p \left( \alpha + \beta + \frac{n}{2} \right)}$, $b > -\frac{n}{2} + \frac{p-1}{2} - \alpha - \beta$ with $S_Y = YY'$. 

**Theorem 3.** Let the real random matrices $X$ and $Y$ have densities as in (13) and (15) respectively. If $|S_Y|$ and $\frac{|S_X|}{|S_Y|}$ are independently distributed, then $\chi_1$ obtained from (6) with the specified set of parameters $(\alpha_1 = \alpha + \frac{n}{2}, \alpha_2 = \alpha_3 = \ldots \alpha_p = \beta, \beta_2 = \beta_3 = \ldots = \beta_p = -(\beta + \frac{1}{2}))$ can be structurally represented as $\frac{|S_X|}{|S_Y|}$. Thus $\chi_1$ can be represented as the ratio of square of the volumes of two paralleloptopes.

**Proof.** If $|S_Y|$ and $\frac{|S_X|}{|S_Y|}$ are independently distributed then we can write

$$|S_X| = |S_Y| \times \frac{|S_X|}{|S_Y|}$$

So $b^{th}$ moment of $|S_X|$ is given by

$$E_x[|S_X|^p] = E_x[|S_Y|^p] \times E_x \left[ \frac{|S_X|}{|S_Y|} \right]^b$$

(16)
due to independence. Therefore we can write it as

$$E_x \left[ \frac{|S_X|^p}{|S_Y|} \right] = \frac{E_x[|S_X|^p]}{E_x[|S_Y|^p]}$$

$$= \frac{\Gamma_p(\alpha + \frac{n}{2} + b)}{\Gamma_p(\alpha + \frac{n}{2})} \times \frac{\Gamma_p(\alpha + \beta + \frac{n}{2})}{\Gamma_p(\alpha + \beta + \frac{n}{2} + b)}$$

$$= \prod_{j=1}^{p} \frac{\Gamma(\alpha + \frac{n}{2} - \frac{j-1}{2} + b)}{\Gamma(\alpha + \frac{n}{2} - \frac{j-1}{2})} \frac{\Gamma(\alpha + \beta + \frac{n}{2} - \frac{j-1}{2})}{\Gamma(\alpha + \beta + \frac{n}{2} - \frac{j-1}{2} + b)}$$

(17)
Consider (6) with \(p\) variables and let \(\alpha_1 = \alpha + \frac{n}{2}, \alpha_2 = \alpha_3 = \ldots = \alpha_p = \beta, \beta_2 = \beta_3 = \ldots = \beta_p = -(\beta + \frac{1}{2})\). Then from (9) we get

\[
E(\chi_1^b) = \prod_{j=1}^{p} \frac{\Gamma(\alpha + \frac{n - j - 1}{2} + b) \Gamma(\alpha + \beta + \frac{n - j - 1}{2})}{\Gamma(\alpha + \frac{n}{2} - \frac{j - 1}{2}) \Gamma(\alpha + \beta + \frac{n}{2} - \frac{j - 1}{2} + b)}.
\]

(18)

Since (17) and (18) are the same we may conclude that \(\chi_1\) with the specified set of parameters can be written as \(\begin{bmatrix} S_X \\ S_Y \end{bmatrix}\) if \(|S_Y|\) and \(|S_X|\) are independently distributed.

From (17) we can get \(\chi_1\) as the product of \(p\) independent real type-1 beta random variables. This type of structure appears in many situations such as distribution of random volume (see Mathai, 1999, 1999a, 2007), the distribution of the likelihood ratio test statistics when testing the hypotheses on the parameters in multivariate normal and other distributions, likelihood ratio criterion for testing hypotheses concerning multivariate analysis of variance (MANOVA), multivariate analysis of covariance (MANCOVA) etc. The importance of this result is that we can study all the above structural representations, likelihood ratio tests, random volumes etc by studying \(\chi_1\). For example, consider a one-one function of the \(\lambda\) -criterion of the likelihood ratio test of the above mentioned hypotheses, denoted by \(\mu\). Let

\[
\mu = \frac{|S_2|}{|S_1 + S_2|}
\]

where \(S_1\) and \(S_2\) are independently distributed Wishart matrices with parameters \((m, \Sigma)\) and \((n, \Sigma)\) respectively, where \(m, n\) are degrees of freedoms and \(\Sigma\) is a symmetric positive definite matrix. It is known that Wishart density is a particular case of a real matrix-variate gamma density. So we may note that \(S_1\) follows a matrix-variate gamma, that is, \(S_1 \sim G_p\left(\frac{m}{2}, \frac{1}{2} \Sigma^{-1}\right)\) and \(S_2 \sim G_p\left(\frac{n}{2}, \frac{1}{2} \Sigma^{-1}\right)\).

When \(S_1, S_2\) are independently distributed then many general properties follow.

For example, \(S_1 + S_2\) and \((S_1 + S_2)^{-1/2} S_2 (S_1 + S_2)^{-1/2}\) are independently distributed, which means \(|S_1 + S_2|\) and \(|S_2|\) are independently distributed. Therefore
It is noted that (19) has the same form of (16); the difference is that, in (19), the denominator \( S_Y \) is represented as a sum.

4. STATISTICAL APPLICATIONS

In this section regression of \( x_k \) on \( x_1, \ldots, x_{k-1} \) and Bayes' estimates are calculated, since these can be obtained explicitly. In order to simplify the calculations we use hypergeometric series. Some basic results and notations used in the derivations are the following:

(i) A hypergeometric series with \( p \) upper parameters and \( q \) lower parameters is defined as

\[
\sum_{r=0}^{\infty} \frac{(a_1)_{r} \cdots (a_p)_{r} \; \zeta^r}{(b_1)_{r} \cdots (b_q)_{r} \; r!}
\]

where \((a_j)_r\) and \((b_j)_r\) are the Pochhammer symbols. For example, for a non-negative integer \( r \),

\[
(a)_r = (a)(a+1)\cdots(a+r-1); \quad (a)_0 = 1, \quad a \neq 0,
\]

\[
= \frac{\Gamma(a + r)}{\Gamma(a)} \text{, when } \Gamma(a) \text{ is defined}
\]

The series in (21) is defined when none of the \( b_j \) s, \( j=1, \ldots, q \), is a negative integer or zero. Its convergence properties are available from books on special functions.

(ii) For \( |\zeta| < 1 \)

\[
(1 - \zeta)^{-a} = \sum_{r=0}^{\infty} \frac{(a)_r \; \zeta^r}{r!} = _1F_0(a; ; \zeta).
\]

This is the binomial series.
4.1. Regression

Let $x_1, \ldots, x_k$ have the joint density as in (6). Then the joint density of $(x_1, \ldots, x_{k-1})$ denoted by $f_4(x_1, \ldots, x_{k-1})$ is given by

$$f_4(x_1, \ldots, x_{k-1}) = C_k x_1^{a_{k-1}} \ldots x_{k-1}^{a_{k-1}} (x_1 + x_2)^{\beta_1} \ldots (x_1 + \ldots + x_{k-1})^{\beta_{k-1}}$$

$$\times \int_{x_i = 0}^{1-x_1-\ldots-x_{k-1}} x_k^{a_{k-1}} (x_1 + \ldots + x_k)^{\beta_k} (1-x_1-\ldots-x_k)^{a_{k+1}-1} dx_k$$  \hspace{1cm} (23)

Now we can write

$$(x_1 + \ldots + x_k)^{\beta_k} = (1-(1-x_1-\ldots-x_k))^{(1-\beta_k)}$$

$$= \sum_{r=0}^{\infty} \frac{(-\beta_k)^r}{r!} (1-x_1-\ldots-x_k)^r \text{ for } |1-x_1-\ldots-x_k| < 1$$

Using this result and (21) we obtain

$$\int_{x_i = 0}^{1-x_1-\ldots-x_{k-1}} x_k^{a_{k-1}} (x_1 + \ldots + x_k)^{\beta_k} (1-x_1-\ldots-x_k)^{a_{k+1}-1} dx_k$$

$$= \sum_{r=0}^{\infty} \frac{(-\beta_k)^r}{r!} \int_{x_i = 0}^{1-x_1-\ldots-x_{k-1}} x_k^{a_{k-1}} (1-x_1-\ldots-x_k)^{a_{k+1}+r-1} dx_k$$

$$= \sum_{r=0}^{\infty} \frac{(-\beta_k)^r}{r!} \frac{\Gamma(\alpha_k) \Gamma(\alpha_{k+1}+r)}{\Gamma(\alpha_k + \alpha_{k+1}+r)} (1-x_1-\ldots-x_{k-1})^{a_k+a_{k+1}+r-1}$$

$$= \frac{\Gamma(\alpha_k) \Gamma(\alpha_{k+1})}{\Gamma(\alpha_k + \alpha_{k+1})} (1-x_1-\ldots-x_{k-1})^{a_k+a_{k+1}-1}$$

$$\times \text{ } _2F_1 \left(-\beta_k, \alpha_{k+1}; \alpha_k + \alpha_{k+1}; (1-x_1-\ldots-x_{k-1}) \right).$$  \hspace{1cm} (24)

Substitute (24) in (23) then we get

$$f_4(x_1, \ldots, x_{k-1}) = C_k \sum_{r=0}^{\infty} \frac{(-\beta_k)^r}{r!} \frac{\Gamma(\alpha_k) \Gamma(\alpha_{k+1}+r)}{\Gamma(\alpha_k + \alpha_{k+1}+r)} (1-x_1-\ldots-x_{k-1})^{a_k+a_{k+1}+r-1}$$

$$\times \text{ } _2F_1 \left(-\beta_k, \alpha_{k+1}; \alpha_k + \alpha_{k+1}; (1-x_1-\ldots-x_{k-1}) \right).$$

If the regression function of $x_k$ on $(x_1, \ldots, x_{k-1})$ is needed then that can be easily obtained.
Here the integration procedure is the same as above. Finally, we obtain

\[
E(x_k | x_1, \ldots, x_{k-1}) = \int_{x_j=0}^{1-x_1-\cdots-x_{k-1}} x_k f(x_k | x_1, \ldots, x_{k-1}) \, dx_k \\
= \frac{\Gamma(a_k + a_{k+1})}{\Gamma(a_k)\Gamma(a_{k+1})} \frac{1}{(1-x_1-\cdots-x_{k-1})^{a_k+a_{k+1}-1}} \times \frac{1}{2} \text{F}_1(-\beta_k, a_{k+1}; a_k + a_{k+1}; (1-x_1-\cdots-x_{k-1})) \\
\times \int_{x_j=0}^{1-x_1-\cdots-x_{k-1}} x_k a_k (x_1 + \cdots + x_k)^{\beta_k} \\
\times (1-x_1-\cdots-x_{k-1})^{a_{k+1}-1} \, dx_k 
\]

(25)

Hence the best predictor of \( x_k \) at preassigned values of \( x_1, \ldots, x_{k-1} \) is given in (26).

4.4. Bayesian analysis

Dirichlet distribution is usually used as the prior distribution for multinomial probabilities. Let \((x_1, \ldots, x_k)\) follow multinomial distribution with probability mass function

\[
f_S(x_1, \ldots, x_k) = \frac{n!}{x_1! \cdots x_k!} \theta_1^{x_1} \ldots \theta_k^{x_k} 
\]

(27)

for \( \theta_i \geq 0, x_j = 0, 1, \ldots, n, \text{ and } i = 1, \ldots, n \), such that \( x_1 + \cdots + x_k = n \) and \( \theta_1 + \cdots + \theta_k = 1 \).

Let the prior distribution of \((\theta_1, \ldots, \theta_k)\) be a generalized Dirichlet density in (6). Then the posterior density is given by the formula

\[
b(\theta | x) = \frac{f_S(\theta) f_S(x | \theta)}{\prod_{\theta_i} \int f_S(\theta) f_S(x | \theta) \, d\theta_k \ldots d\theta_1}.
\]

Now
\[
\int_{\theta_1} \ldots \int_{\theta_k} f_\theta(x) f_\theta(x) \left| \theta \right| d\theta_1 \ldots d\theta_k
= C_k \int_{\theta_1 = 0}^{1-\theta_1-\cdots-\theta_k} \theta_1^{a_1+x_1-1} \cdots \theta_k^{a_k+x_k-1} \times (\theta_1 + \theta_2)^{\beta_2} \cdots (\theta_1 + \cdots + \theta_k)^{\beta_k} \times (1-\theta_1-\cdots-\theta_k)^{a_{k+1}-1} d\theta_k \ldots d\theta_1
= C_k \int_{x_1}^{x_1} \ldots \int_{x_k}^{x_k} \frac{\Gamma(a_1 + x_1) \cdots \Gamma(a_k + x_k)\Gamma(\alpha_{k+1})}{\Gamma(a_1 + \alpha_2 + x_1 + x_2) \cdots \Gamma(a_1 + \alpha_2 + x_1 + x_2 + \alpha_3 + x_3 + \alpha_4 + \beta_2) \cdots \Gamma(a_1 + \cdots + \alpha_{k+1} + x_1 + \cdots + x_k + \beta_2 + \cdots + \beta_k)} \times \theta_1^{a_1+x_1-1} \cdots \times \theta_k^{a_k+x_k-1}(\theta_1 + \theta_2)^{\beta_2} \cdots (\theta_1 + \cdots + \theta_k)^{\beta_k}(1-\theta_1-\cdots-\theta_k)^{a_{k+1}-1},
\]

\(b(\theta | x) = \frac{\Gamma(a_1 + \alpha_2 + x_1 + x_2)\Gamma(a_1 + \alpha_2 + x_2 + \beta_2) \cdots}{\Gamma(a_1 + x_1) \cdots \Gamma(a_k + x_k)\Gamma(\alpha_{k+1})} \times \frac{\Gamma(a_1 + \cdots + \alpha_{k+1} + x_1 + \cdots + x_k + \beta_2 + \cdots + \beta_k)\theta_1^{a_1+x_1-1} \cdots \times \theta_k^{a_k+x_k-1}(\theta_1 + \theta_2)^{\beta_2} \cdots (\theta_1 + \cdots + \theta_k)^{\beta_k}(1-\theta_1-\cdots-\theta_k)^{a_{k+1}-1},\)

for \(0 < \theta_1 + \cdots + \theta_k < 1, \alpha_j + x_j > 0, i = 1, \ldots, k, \alpha_{k+1} > 0, \alpha_1 + \cdots + \alpha_j + x_1 + \cdots + x_j + \beta_2 + \cdots + \beta_j > 0, j = 1, \ldots, k\).

If we replace \(\alpha_j + x_j\) by \(\gamma_j, j = 1, \ldots, k\), then it can be seen that the posterior density (28) has the form of generalized Dirichlet density (6) with different parameters. Therefore we can say that generalized Dirichlet density in (6) is conjugate to multinomial density.

Now the Bayes' estimate for \(\theta_1\), with quadratic loss, is

\[
E(\theta_1 | x) = \int_{\theta_1 = 0}^{1} \int_{\theta_2 = 0}^{1-\theta_1-\theta_2-\cdots-\theta_k} \theta_1 b(\theta | x) d\theta_1 
= \frac{(a_1 + x_1)(a_1 + \alpha_2 + x_1 + x_2 + \beta_2) \cdots}{(a_1 + \alpha_2 + x_1 + x_2)(a_1 + \alpha_2 + \alpha_3 + x_1 + x_2 + \alpha_3 + x_3 + \beta_2) \cdots \times \frac{(a_1 + \cdots + \alpha_{k+1} + x_1 + \cdots + x_k + \beta_2 + \cdots + \beta_k)}{(a_1 + \cdots + \alpha_{k+1} + x_1 + \cdots + x_k + \beta_2 + \cdots + \beta_k)} 
\]
Bayes’ estimate for $\theta_2$ is

$$E(\theta_2|x) = \frac{(\alpha_2 + x_2)(\alpha_1 + \alpha_2 + \alpha_3 + x_1 + x_2 + \beta_2)}{(\alpha_1 + \alpha_2 + x_1 + x_2)(\alpha_1 + \alpha_2 + \alpha_3 + x_1 + x_2 + x_3 + \beta_2)} \times \frac{\alpha_1 + \ldots + \alpha_k + x_1 + \ldots + x_k + \beta_2 + \ldots + \beta_k}{(\alpha_1 + \ldots + \alpha_{k+1} + x_1 + \ldots + x_k + \beta_2 + \ldots + \beta_k)},$$

and so on, and finally Bayes’ estimate for $\theta_k$ is

$$E(\theta_k|x) = \frac{(\alpha_k + x_k)}{(\alpha_1 + \ldots + \alpha_k + x_1 + \ldots + x_k + \beta_2 + \ldots + \beta_{k-1})} \times \frac{\alpha_1 + \ldots + \alpha_k + x_1 + \ldots + x_k + \beta_2 + \ldots + \beta_k}{(\alpha_1 + \ldots + \alpha_{k+1} + x_1 + \ldots + x_k + \beta_2 + \ldots + \beta_k)}.$$
Some properties of a generalized type-1 Dirichlet distribution

This paper deals with a generalization of type-1 Dirichlet density by incorporating partial sums of the component variables. We study various proportions, structural decompositions, connections to random volumes and \( p \)-parallelotopes. We will also look into the regression function of \( x_k \) on \( x_1,...,x_{k-1} \), Bayes’ estimates for the probabilities of a multinomial distribution by using this generalized Dirichlet model as the prior density are given. Other results illustrate the importance of the study of variable \( x_1 \) in this model. It is found that the variable \( x_1 \) in this model can be represented as the ratio of squares of volumes of two parallelotopes. Under certain conditions, \( x_1 \) can be used to study the structural representations of the likelihood ratio criteria in MANOVA, MANCOVA etc.