STATISTICA, anno LXVIII, nn. 3-4, 2008

DISTRIBUTION OF THE LR CRITERION $U_{p,m,n}$ AS A MARGINAL DISTRIBUTION OF A GENERALIZED DIRICHLET MODEL

S. Thomas, A. Thannippara

1. INTRODUCTION

It is well known that most of the testing hypotheses procedures based on random samples from multivariate normal distributions are in terms of Wilks' likelihood ratio statistic. The testing of hypotheses in multivariate regression analysis, multivariate analysis of variance, multivariate analysis of covariance, canonical correlations etc. are based on the null distribution of the likelihood ratio statistic. In addition to these areas we use likelihood ratio criterion for testing the independence of sets of variates and also for testing the significance of a subvector in the T^2-test of H_0: $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$. The construction of confidence region or simultaneous confidence intervals also make use of the distribution of the likelihood ratio criterion. Following (Anderson, 2003) we shall denote the above mentioned likelihood ratio criterion as $U_{p,m,n}$ where p is the dimension, m is the degrees of freedom for hypothesis and n is the degrees of freedom for error. The details of the wide range of applications of $U_{p,m,n}$ can be seen in standard textbooks on multivariate analysis, see for example, (Anderson, 2003) and (Rencher, 1998). The exact distribution of $U_{p,m,n}$ has been investigated by several authors such as (Schatzoff, 1966), (Pillai and Gupta, 1969), (Mathai, 1971) and (Coelho, 1998). All the results have been obtained under the form of series expansions or in terms of complicated expressions so that computation of p-values of $U_{p,m,n}$ from its exact density is a tedious task. Hence one has to rely on simulation procedures for the computation of p-values. The main purpose of this paper is to show the connection of a generalized Dirichlet model to the distribution of $U_{p,m,n}$ and represent the density of $U_{p,m,n}$ in a tractable form which enables the computation of p-values of $U_{p,m,n}$ directly from its exact density. A similar study in the case of the distribution of the likelihood ratio criterion for sphericity test can be seen from (Thomas and Thannippara, 2008). We shall first consider a few preliminary results. In Section 3, we express the exact density of $U_{p,m,n}$ as a marginal distribution of the generalized Dirichlet model. Section 4 provides closed form of the density of $U_{2,m,n}$ and $U_{4,m,n}$. Finally, computation of p-values of $U_{p,m,n}$ is illustrated.

2. PRELIMINARY RESULTS

Lemma 2.1

The likelihood ratio criterion $U_{p,m,n}$ based on observations from a p-variate normal distribution has the distribution of

$$U_{p,m,n} = \frac{|\mathbf{G}|}{|\mathbf{G}+\mathbf{H}|} \tag{1}$$

where **G** and **H** are independent and each distributed according to Wishart with n and m degrees of freedom respectively.

Theorem 2.1 (Anderson, 2003, Theorem 8.4.3)

The t-th moment of $U_{p,m,n}$ for $t > -\frac{1}{2}(n+1-p)$ is

$$E(U^t) = \prod_{j=1}^{p} \frac{\Gamma\left[\frac{1}{2}(n+1-j)+t\right]\Gamma\left[\frac{1}{2}(n+m+1-j)\right]}{\Gamma\left[\frac{1}{2}(n+1-j)\right]\Gamma\left[\frac{1}{2}(n+m+1-j)+t\right]}. \tag{2}$$

We can write (2) as

$$E(U^t) = \prod_{j=1}^{p} E(V_j^t)$$

where V_j has the type-1 beta distribution with the parameters $\left[\frac{1}{2}(n+1-j), \frac{1}{2}m)\right]$.

Hence it follows that the distribution of $U_{p,m,n}$ is that of the product $\prod_{j=1}^{p} V_j$ where $V_1, \ldots, V_p$ are independent and V_j has the type-1 beta distribution with the parameters $\left[\frac{1}{2}(n+1-j), \frac{1}{2}m)\right]$.

Now consider the real variables $0 < x_i < 1,\ i = 1,\ldots,p,\ 0 < x_1 + \cdots + x_p < 1$ such that

$$z_1 = \frac{x_1}{x_1 + x_2}, \ldots, z_{p-1} = \frac{x_1 + \cdots + x_{p-1}}{x_1 + \cdots + x_p}, z_p = x_1 + \cdots + x_p \tag{3}$$

are independently distributed. If we further assume that z_j's in (3) are type-1 beta random variables with z_j having the parameters $(\alpha_1 + \cdots + \alpha_j + \beta_2 + \cdots + \beta_j, \alpha_{j+1})$ for $j = 2,\ldots,p$ and z_1 has the parameters (α_1, α_2), then (Thomas and George, 2004) have shown that $(x_1,\ldots,x_p)$ has the density function of the following structure:

$$\begin{aligned} f(x_1,\ldots,x_p) &= c_p x_1^{\alpha_1 - 1} x_2^{\alpha_2 - 1} \ldots x_p^{\alpha_p - 1} (x_1 + x_2)^{\beta_2} \ldots \\ &\quad \times (x_1 + \cdots + x_p)^{\beta_p} (1 - x_1 - \cdots - x_p)^{\alpha_{p+1} - 1} \end{aligned} \tag{4}$$

where the normalizing constant c_p is such that

$$\begin{aligned} c_p^{-1} &= \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)\cdots\Gamma(\alpha_{p+1})}{\Gamma(\alpha_1 + \alpha_2)} \frac{\Gamma(\alpha_1 + \alpha_2 + \beta_2)}{\Gamma(\alpha_1 + \alpha_2 + \alpha_3 + \beta_2)} \cdots \\ &\quad \times \frac{\Gamma(\alpha_1 + \cdots + \alpha_p + \beta_2 + \cdots + \beta_p)}{\Gamma(\alpha_1 + \cdots + \alpha_{p+1} + \beta_2 + \cdots + \beta_p)} \end{aligned} \tag{5}$$

for $\alpha_j > 0,\ j = 1,\ldots,p+1,\ \alpha_1 + \cdots + \alpha_j + \beta_2 + \cdots + \beta_j > 0,\ j = 2,\ldots,p.$ The model (4) was introduced by (Thomas and George, 2004) as a generalization of type-1 Dirichlet model. Note that the β_j's can take negative values also in this model. Similar extensions of type-1 Dirichlet model can be seen in the literature, see for example, (Connor and Mosimann, 1969) and (Lochner, 1975).

Definition 2.1

A general G-function is defined as the following Mellin-Barnes integral:

$$G_{p,q}^{m,n}\left[z \middle| \begin{matrix} a_1,\ldots,a_p \\ b_1,\ldots,b_q \end{matrix} \right]$$

$$= (2\pi i)^{-1} \int_L \frac{\{\Pi_{j=1}^{m} \Gamma(b_j + s)\}\{\Pi_{j=1}^{n} \Gamma(1 - a_j - s)\}}{\{\Pi_{j=m+1}^{q} \Gamma(1 - b_j - s)\}\{\Pi_{j=n+1}^{p} \Gamma(a_j + s)\}} z^{-s} ds$$

where $i = \sqrt{-1}$ and L is a suitable contour.

The existence of different types of contours, general existence conditions, properties and applications are available in (Mathai, 1993).

Theorem 2.2

x_1 is structurally a product of p independent real type-1 beta random variables and its density can be written in terms of a Meijer's G-function of the type $G_{p,p}^{p,0}(\cdot)$.

Proof. From (3) by taking the product we see that

$$x_1 = z_1 z_2 \ldots z_p.$$

$$\begin{aligned} E(x_1^t) &= \prod_{j=1}^{p} E(z_j^t) \\ &= c_p' \prod_{j=1}^{p} \frac{\Gamma(\alpha_1 + \cdots + \alpha_j + \beta_2 + \cdots + \beta_j + t)}{\Gamma(\alpha_1 + \cdots + \alpha_{j+1} + \beta_2 + \cdots + \beta_j + t)} \end{aligned} \tag{6}$$

where

$$c_p' = \prod_{j=1}^{p} \frac{\Gamma(\alpha_1 + \cdots + \alpha_{j+1} + \beta_2 + \cdots + \beta_j)}{\Gamma(\alpha_1 + \cdots + \alpha_j + \beta_2 + \cdots + \beta_j)}.$$

Treating (6) as a Mellin transform of the density of x_1 the density is available by taking the inverse Mellin transform. Denoting the density of x_1 by $g(x_1)$ we have,

$$\begin{aligned} g(x_1) &= c_p' x_1^{-1} \frac{1}{2\pi i} \int_L \frac{\Gamma(\alpha_1 + t)}{\Gamma(\alpha_1 + \alpha_2 + t)} \frac{\Gamma(\alpha_1 + \alpha_2 + \beta_2 + t)}{\Gamma(\alpha_1 + \alpha_2 + \alpha_3 + \beta_2 + t)} \cdots \\ &\quad \times \frac{\Gamma(\alpha_1 + \cdots + \alpha_p + \beta_2 + \cdots + \beta_p + t)}{\Gamma(\alpha_1 + \cdots + \alpha_{p+1} + \beta_2 + \cdots + \beta_p + t)} x_1^{-t} dt \\ &= c_p' x_1^{-1} G_{p,p}^{p,0} \left[x_1 \middle| \begin{matrix} \alpha_1 + \alpha_2, \ldots, \alpha_1 + \cdots + \alpha_{p+1} + \beta_2 + \cdots + \beta_p \\ \alpha_1, \ldots, \alpha_1 + \cdots + \alpha_p + \beta_2 + \cdots + \beta_p \end{matrix} \right] \end{aligned} \tag{7}$$

for $0 < x_1 < 1$ and zero elsewhere.

3. DENSITY OF $U_{p,m,n}$ AS A GENERALIZED DIRICHLET MARGINAL

Now let us consider (6) and put

$$\alpha_1 = \frac{n}{2},\ \alpha_2 = \alpha_3 = \cdots = \alpha_{p+1} = \frac{m}{2},\ \beta_2 = \beta_3 = \cdots = \beta_p = -\left(\frac{m+1}{2}\right). \tag{8}$$

On simplification we have

$$E(x_1^t) = \prod_{j=1}^{p} \frac{\Gamma\left[\frac{1}{2}(n+1-j)+t\right]\Gamma\left[\frac{1}{2}(n+m+1-j)\right]}{\Gamma\left[\frac{1}{2}(n+1-j)\right]\Gamma\left[\frac{1}{2}(n+m+1-j)+t\right]}$$

which is the same as (2).

From the above considerations and since arbitrary moments in this case will determine the density uniquely we can write the following theorem.

Theorem 3.1

When $(x_1,\ldots,x_p)$ has a real generalized type-1 Dirichlet density of (4) with the parameters as given in (8), then the distribution of $U_{p,m,n}$ in (1) and the marginal distribution of x_1 are identical. The exact density of $U_{p,m,n}$ is given by

$$\begin{aligned} g(x_1) \quad = \quad & c_p \int_0^{1-x_1}\int_0^{1-x_1-x_2}\cdots\int_0^{1-x_1-\cdots-x_{p-1}} x_1^{\frac{n}{2}-1} x_2^{\frac{m}{2}-1}\cdots x_p^{\frac{m}{2}-1} \times (x_1+x_2)^{-\left(\frac{m+1}{2}\right)}\cdots \\ & \times (x_1+\cdots+x_p)^{-\left(\frac{m+1}{2}\right)}(1-x_1-\cdots-x_p)^{\frac{m}{2}-1}\, dx_p \ldots dx_2 \end{aligned} \tag{9}$$

for $0 < x_i < 1, i = 1,\ldots,p,\ 0 < x_1 + \cdots + x_p < 1,\ n \ge p$ and zero otherwise; where

$$c_p = \frac{\Pi_{j=1}^{p}\Gamma\left[\frac{1}{2}(n+m+1-j)\right]}{\left[\Gamma\left(\frac{m}{2}\right)\right]^p \Pi_{j=1}^{p}\Gamma\left[\frac{1}{2}(n+1-j)\right]}.$$

Remark 3.1

On performing the integrations in (9) we can express $g(x_1)$ in terms of the following multiple series:

$$g(x_1) = c_p \frac{\left[\Gamma\left(\frac{m}{2}\right)\right]^p}{\Gamma\left(\frac{pm}{2}\right)} x_1^{\frac{n}{2}-1}(1-x_1)^{\frac{pm}{2}-1} \sum_{r_1=0}^{\infty} \frac{\left(\frac{m+1}{2}\right)_{r_1}\left(\frac{m}{2}\right)_{r_1}}{\left(\frac{2m}{2}\right)_{r_1}} \frac{(1-x_1)^{r_1}}{r_1!}$$

$$\times \sum_{r_2=0}^{\infty} \frac{\left(\frac{m+1}{2}\right)_{r_2}\left(\frac{2m}{2}\right)_{r_1+r_2}}{\left(\frac{3m}{2}\right)_{r_1+r_2}} \frac{(1-x_1)^{r_2}}{r_2!}\cdots$$

$$\times \sum_{r_{p-1}=0}^{\infty} \frac{\left(\frac{m+1}{2}\right)_{r_{p-1}}\left(\frac{(p-1)m}{2}\right)_{r_1+\cdots+r_{p-1}}}{\left(\frac{pm}{2}\right)_{r_1+\cdots+r_{p-1}}} \frac{(1-x_1)^{r_{p-1}}}{r_{p-1}!},$$

where $(a)_r$ is defined in (A1).

The way in which the above representation of $g(x_1)$ comes can be seen from the intermediate steps in the derivation of the density of $U_{2,m,n}$ given in Section 4.

Remark 3.2

On replacing the α_j s and β_j s in (7) by (8) we can also obtain the exact distribution of $U_{p,m,n}$ in terms of Meijer's G-function as follows:

$$g(u) = c_p' u^{-1} G_{p,p}^{p,0}\left[u \middle| \begin{matrix} a_1,\ldots,a_p \\ b_1,\ldots,b_p \end{matrix}\right] \tag{10}$$

where

$$b_j = \frac{1}{2}(n+1-j) \text{ for } j=1,\ldots,p, \quad a_j = b_j + \frac{m}{2} \text{ for } j=1,\ldots,p \text{ and}$$

$$c_p' = \prod_{j=1}^{p} \frac{\Gamma(a_j)}{\Gamma(b_j)}.$$

The derivation of the density of $U_{4,m,n}$ by using Meijer's G-function is illustrated in Section 4.

The likelihood ratio test procedure rejects the null hypothesis H_0 if the computed value of $U_{p,m,n}$ is less than $u_{p,m,n}(\alpha)$, the α significant point for $U_{p,m,n}$.

The values of $u_{p,m,n}(\alpha)$ are tabulated for certain values of p,m,n and α. When p or m is small, there are good approximations in terms of the F-distribution (Anderson, 2003, Section 8.5.4). In other situations there are other methods using asymptotic distribution of $U_{p,m,n}$ or using the distribution of some functions of $U_{p,m,n}$ or Monte Carlo simulation are also available in the literature.

Remark 3.3

Theorem 3.1 enables us to express the distribution of $U_{p,m,n}$ in a simple manageable form and we can directly obtain the exact p-value corresponding to a computed value u of $U_{p,m,n}$. Hence there is no need to rely on simulation or tabulated critical points or asymptotic distribution of some functions of $U_{p,m,n}$. The exact p-value for a computed value u of $U_{p,m,n}$ can be obtained very easily by evaluating the following with the help of Mathematica or Maple:

$$P(U_{p,m,n} \le u) = \int_0^u g(x_1)dx_1, \tag{11}$$

where $g(x_1)$ is given in (9).

4. SOME SPECIAL CASES

We have seen the multiple integral representation of the density of $U_{p,m,n}$ in theorem 3.1. In this section we shall obtain the explicit form of the density of $U_{2,m,n}$ and $U_{4,m,n}$.

Density of $U_{2,m,n}$

From (9) it follows that the marginal density of x_1 is

$$g(x_1) = c_2 x_1^{\frac{n}{2}-1} \int_{x_2=0}^{1-x_1} x_2^{\frac{m}{2}-1} (x_1+x_2)^{-\left(\frac{m+1}{2}\right)} (1-x_1-x_2)^{\frac{m}{2}-1} dx_2;\ 0 < x_1 < 1, \text{ and}$$

zero otherwise; where

$$c_2 = \frac{\Gamma\left(\frac{n+m}{2}\right)\Gamma\left(\frac{n+m-1}{2}\right)}{\left[\Gamma\left(\frac{m}{2}\right)\right]^2 \Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n-1}{2}\right)}.$$

By applying (A3) we can write c_2 as

$$c_2 = \frac{\Gamma(n+m-1)}{2^m \left[\Gamma\left(\frac{m}{2}\right)\right]^2 \Gamma(n-1)}.$$

On writing

$$(x_1+x_2)^{-\left(\frac{m+1}{2}\right)} = [1-(1-x_1-x_2)]^{-\left(\frac{m+1}{2}\right)}$$

and using (A2) we have

$$(x_1+x_2)^{-\left(\frac{m+1}{2}\right)} = \sum_{r=0}^{\infty} \frac{\left(\frac{m+1}{2}\right)_r}{r!} (1-x_1-x_2)^r.$$

Now

$$\int_{x_2=0}^{1-x_1} x_2^{\frac{m}{2}-1} (x_1+x_2)^{-\left(\frac{m+1}{2}\right)} (1-x_1-x_2)^{\frac{m}{2}-1} dx_2$$

$$= \sum_{r=0}^{\infty} \frac{\left(\frac{m+1}{2}\right)_r}{r!} \int_{x_2=0}^{1-x_1} x_2^{\frac{m}{2}-1} (1-x_1-x_2)^{\frac{m}{2}+r-1} dx_2$$

$$= \sum_{r=0}^{\infty} \frac{\left(\frac{m+1}{2}\right)_r}{r!} (1-x_1)^{\frac{m}{2}+r-1} \int_{x_2=0}^{1-x_1} x_2^{\frac{m}{2}-1} \left[1-\frac{x_2}{1-x_1}\right]^{\frac{m}{2}+r-1} dx_2$$

$$= \sum_{r=0}^{\infty} \frac{\left(\frac{m+1}{2}\right)_r}{r!} (1-x_1)^{m+r-1} \int_{y=0}^{1} y^{\frac{m}{2}-1} (1-y)^{\frac{m}{2}+r-1} dy$$

$$= \sum_{r=0}^{\infty} \frac{\left(\frac{m+1}{2}\right)_r}{r!} (1-x_1)^{m+r-1} \frac{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{m}{2}+r\right)}{\Gamma(m+r)}$$

$$= \sum_{r=0}^{\infty} \frac{\left(\frac{m+1}{2}\right)_r}{r!} (1-x_1)^{m+r-1} \frac{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{m}{2}\right)\left(\frac{m}{2}\right)_r}{\Gamma(m)(m)_r} \quad \text{by using } (A1).$$

Thus

$$g(x_1) = \frac{\Gamma(n+m-1)}{2^m\Gamma(n-1)\Gamma(m)} x_1^{\frac{n}{2}-1}(1-x_1)^{m-1}\sum_{r=0}^{\infty}\frac{\left(\frac{m}{2}\right)_r\left(\frac{m+1}{2}\right)_r}{(m)_r}\frac{(1-x_1)^r}{r!}$$

for $0<x_1<1$. Now by using (A5) it can be written as

$$g(x_1) = \frac{1}{2^m B(n-1,m)} x_1^{\frac{n}{2}-1}(1-x_1)^{m-1}\ {}_2F_1\left(\frac{m}{2},\frac{m+1}{2};m;1-x_1\right) \quad (12)$$

for $0<x_1<1, n\geq 2$ and zero elsewhere.

As an immediate consequence we can write the following remarks.

Remark 4.1

For $0<x_i<1,\ i=1,2,\ 0<x_1+x_2<1$,

$$\int_{x_2=0}^{1-x_1} x_2^{\frac{m}{2}-1}(x_1+x_2)^{-\left(\frac{m+1}{2}\right)}(1-x_1-x_2)^{\frac{m}{2}-1}dx_2 =$$
$$\frac{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{m}{2}\right)}{\Gamma(m)}(1-x_1)^{m-1}\ {}_2F_1\left(\frac{m}{2},\frac{m+1}{2};m;1-x_1\right).$$

Remark 4.2

Since $\int_0^1 g(x_1)dx_1 = 1$, then from (12) it follows that

$$\int_0^1 x^{\frac{n}{2}-1}(1-x)^{m-1}\ {}_2F_1\left(\frac{m}{2},\frac{m+1}{2};m;1-x\right)dx = 2^m B(n-1,m).$$

Alternatively, we can obtain the density (12) from (10) in terms of Meijer's G-function as follows.

$$g(u) = c_2^{'}u^{-1}G_{2,2}^{2,0}\left(u\,\middle|\,\begin{matrix}a_1,a_2\\ b_1,b_2\end{matrix}\right)$$

where

$$a_1=\frac{1}{2}(n+m),\ a_2=\frac{1}{2}(n+m-1),\ b_1=\frac{1}{2}n,\ b_2=\frac{1}{2}(n-1)$$

and

$$c_2^{'} = \frac{\Gamma\left[\frac{1}{2}(n+m)\right]\Gamma\left[\frac{1}{2}(n+m-1)\right]}{\Gamma\left[\frac{1}{2}n\right]\Gamma\left[\frac{1}{2}(n-1)\right]} = \frac{\Gamma(n+m-1)}{2^m\Gamma(n-1)}.$$

Now $G_{2,2}^{2,0}\left(u\middle|\begin{matrix}a_1,a_2\\b_1,b_2\end{matrix}\right)$ is of the form $G_{2,2}^{2,0}\left(u\middle|\begin{matrix}\gamma_1+\delta_1-1,\gamma_2+\delta_2-1\\\gamma_1-1,\gamma_2-1\end{matrix}\right)$ with $\gamma_1-1=\frac{1}{2}(n-1),\quad \gamma_2-1=\frac{1}{2}n,\ \delta_1=\delta_2=\frac{1}{2}m.$

Hence we can use the result given in (A4) and write the density function as

$$\begin{aligned} g(u) &= \frac{\Gamma(n+m-1)}{2^m\Gamma(n-1)}u^{-1}\frac{u^{\frac{n}{2}}(1-u)^{m-1}}{\Gamma(m)}\ {}_2F_1\left(\frac{m+1}{2},\frac{m}{2};m;1-u\right);\ 0<u<1 \\ &= \frac{1}{2^m B(n-1,m)}u^{\frac{n}{2}-1}(1-u)^{m-1}\ {}_2F_1\left(\frac{m}{2},\frac{m+1}{2};m;1-u\right);\quad 0<u<1, \end{aligned} \tag{13}$$

and zero elsewhere; which is the same as (12).

If we use the form $G_{2,2}^{2,0}\left(u\middle|\begin{matrix}\alpha+\beta-1,\alpha+\beta-\frac{1}{2}\\\alpha-1,\alpha-\frac{1}{2}\end{matrix}\right)$ with $\alpha=\frac{1}{2}(n+1)$ and $\beta=\frac{1}{2}m$ instead of $G_{2,2}^{2,0}\left(u\middle|\begin{matrix}\gamma_1+\delta_1-1,\gamma_2+\delta_2-1\\\gamma_1-1,\gamma_2-1\end{matrix}\right)$ we obtain a different form of the density function as given below.

$$g(u)=c_2^{'}u^{-1}G_{2,2}^{2,0}\left(u\middle|\begin{matrix}\alpha+\beta-1,\alpha+\beta-\frac{1}{2}\\\alpha-1,\alpha-\frac{1}{2}\end{matrix}\right)$$

where $\alpha=\frac{1}{2}(n+1),\ \beta=\frac{1}{2}m$ and $c_2^{'}=\frac{\Gamma(n+m-1)}{2^m\Gamma(n-1)}$.

Now apply (A6) and obtain $g(u)$ as

$$g(u)=c_2^{'}u^{-1}2^{m-1}G_{1,1}^{1,0}\left(u^{\frac{1}{2}}\middle|\begin{matrix}n+m-1\\n-1\end{matrix}\right).$$

Again, by using (A7) we can write $g(u)$ as

$$g(u) = c_2^{'} u^{-1} 2^{m-1} \frac{(u^{\frac{1}{2}})^{n-1}(1-u^{\frac{1}{2}})^{m-1}}{\Gamma(m)}; \ 0 < u < 1.$$

That is,

$$g(u) = \frac{1}{2B(n-1,m)}(u^{\frac{1}{2}})^{n-3}(1-u^{\frac{1}{2}})^{m-1}; \ 0 < u < 1, \tag{14}$$

and zero elsewhere. As the density function is unique, (13) and (14) must be equal. Therefore we obtain the following relation which we shall write as a remark.

Remark 4.3

$${}_2F_1\left(\frac{m}{2}, \frac{m+1}{2}; m; 1-u\right) = 2^{m-1} u^{-\frac{1}{2}} \left[1 + u^{\frac{1}{2}}\right]^{-(m-1)}; \ 0 < u < 1.$$

Density of $U_{4,m,n}$

We can write the required density function using (10) as

$$g(u) = c_4^{'} u^{-1} G_{4,4}^{4,0}\left(u \middle| \begin{matrix} a_1, a_2, a_3, a_4 \\ b_1, b_2, b_3, b_4 \end{matrix}\right)$$

Note that

$$G_{4,4}^{4,0}\left(u \middle| \begin{matrix} a_1, a_2, a_3, a_4 \\ b_1, b_2, b_3, b_4 \end{matrix}\right) = (2\pi i)^{-1} \int_L \phi(s) u^{-s} ds$$

where

$$\phi(s) = \prod_{j=1}^{4} \left[\frac{\Gamma(b_j + s)}{\Gamma(a_j + s)}\right] = \prod_{j=1}^{4} \frac{\Gamma\left[\frac{1}{2}(n+1-j)+s\right]}{\Gamma\left[\frac{1}{2}(n+m+1-j)+s\right]}.$$

Now on combining the two consecutive gammas in the numerator and denominator of $\phi(s)$ by using (A3) we obtain

$$\phi(s) = 2^{2m} \frac{\Gamma(n-1+2s)\Gamma(n-3+2s)}{\Gamma(n+m-1+2s)\Gamma(n+m-3+2s)}.$$

On putting $2s = s'$ the above Mellin-Barnes type integral becomes

$$\begin{aligned} G_{4,4}^{4,0}\left(u\left|\begin{matrix} a_1, a_2, a_3, a_4 \\ b_1, b_2, b_3, b_4 \end{matrix}\right.\right) \\ &= 2^{2m-1}(2\pi i)^{-1}\int_L \frac{\Gamma(n-1+s')\Gamma(n-3+s')}{\Gamma(n+m-1+s')\Gamma(n+m-3+s')}(u^{\frac{1}{2}})^{-s'}\,ds' \\ &= 2^{2m-1}G_{2,2}^{2,0}\left(u^{\frac{1}{2}}\left|\begin{matrix} n+m-1, n+m-3 \\ n-1, n-3 \end{matrix}\right.\right). \end{aligned}$$

On using (A4) with $\gamma_1 = n-2,\ \gamma_2 = n,\ \delta_1 = \delta_2 = m$ and simplifying $c_4^{'}$ using (A3) we obtain the final form of the density function as

$$\begin{aligned} g(u) &= \frac{\Gamma(n+m-1)\Gamma(n+m-3)}{2\Gamma(n-1)\Gamma(n-3)\Gamma(2m)}(u^{\frac{1}{2}})^{n-3}(1-u^{\frac{1}{2}})^{2m-1} \\ &\quad \times {}_2F_1(m+2, m; 2m; 1-u^{\frac{1}{2}});\quad 0<u<1,\ n\geq 4 \end{aligned} \tag{15}$$

and zero elsewhere.

Since (15) is a density we obtain the following relation.

Remark 4.4

$$\int_0^1 (u^{\frac{1}{2}})^{n-3}(1-u^{\frac{1}{2}})^{2m-1}\ {}_2F_1(m+2, m; 2m; 1-u^{\frac{1}{2}})du = \frac{2\Gamma(n-1)\Gamma(n-3)\Gamma(2m)}{\Gamma(n+m-1)\Gamma(n+m-3)}.$$

Remark 4.5

In some cases the deduction of density function using (10) may be quite tedious. For example, the case when $p=3$ requires evaluation of residues at poles of orders one and two. Hence in such cases we may end up with a density involving psi and gamma functions. When $p>3$ it can involve zeta functions also. But for practical purposes we need only to compute the values of cumulative distribution function of $U_{p,m,n}$ and this can be done very easily by using (11) for any p.

5. COMPUTATIONS

When $p=2$ we have seen that the density function of $U_{2,m,n}$ is of the form (14) and hence we can replace $g(x_1)$ in (11) by (14). Then putting $u^{\frac{1}{2}} = v$ we can write (11) as

$$P(U_{2,m,n} \le u) = \frac{1}{B(n-1,m)} \int_0^{\sqrt{u}} v^{(n-1)-1}(1-v)^{m-1}\,dv. \tag{16}$$

Therefore, the computation of the above probability is the same as the evaluation of cumulative distribution function of a type-1 beta random variable with the parameters $(n-1,m)$ and it can be done very easily using MS Excel. When $p \ge 3$, the computation of p-value for an observed value u of $U_{p,m,n}$ can be performed with widely used software packages such as Mathematica and Maple. For example, let us compute (11) for an observed value $u = 0.121$ of $U_{4,6,17}$ using Mathematica. The Mathematica input:

$$
\begin{aligned}
p: &= 4\\
m: &= 6\\
n: &= 17\\
a &= \prod_{i=1}^{p} (Gamma[(n+m+1-i)/2])/(Gamma[m/2]Gamma[(n+1-i)/2])\\
b &= \int_0^{0.121}\int_0^{1-w}\int_0^{1-w-x}\int_0^{1-w-x-y} w^{\frac{n}{2}-1} x^{\frac{m}{2}-1} y^{\frac{m}{2}-1} z^{\frac{m}{2}-1} (w+x)^{-\left(\frac{m+1}{2}\right)}\\
&\quad (w+x+y)^{-\left(\frac{m+1}{2}\right)} (w+x+y+z)^{-\left(\frac{m+1}{2}\right)} (1-w-x-y-z)^{\frac{m}{2}-1}\, dz dy dx dw\\
c &= a * b
\end{aligned}
$$

gives the output value 0.0503688. Note that 0.121 is an entry in the table of lower critical values of likelihood ratio criterion corresponding to $\alpha = 0.05, p = 4, m = 6$ and $n = 17$ (Rencher, 1998, Table B4). We can also evaluate (11) for the above mentioned case by using (15) instead of (9).

Remark 5.1.

Since the distributions of $U_{p,m.n}$ and $U_{m,p,n+m-p}$ are same (Bilodeau and Brenner, 1999, Corollary 11.1) we can use this fact in computations of the p-value especially when $m < p$.

For illustrative purposes the lower critical values corresponding to $\alpha = 0.05$ and $p = 2$ are computed from the representation (16) by using MS Excel is given in table 1. Since the significance points of $U_{p,m,n}$ for $\alpha = 0.05$ and $p = 1(1)8$ are available in the literature and the table values obtained using the density (9) coincides with that of Table B4 of (Rencher, 1998), we are not reproducing the entire table values here.

TABLE 1

Lower Critical Values of $U_{p,m,n}$ *for* $\alpha = 0.05$ *and* $p = 2$

	m								
n	1	2	3	4	5	6	7	8	9
2	$0.0^{2}25000$	$0.0^{3}64113$	$0.0^{3}28739$	$0.0^{3}16234$	$0.0^{3}10417$	$0.0^{4}72462$	$0.0^{4}53303$	$0.0^{4}40847$	$0.0^{3}32297$
3	0.05000	0.01832	$0.0^{2}95280$	$0.0^{2}58431$	$0.0^{2}39501$	$0.0^{2}28489$	$0.0^{2}21520$	$0.0^{2}16829$	$0.0^{2}13521$
4	0.13572	0.06180	0.03582	0.02346	0.01658	0.01235	$0.0^{2}95544$*	$0.0^{2}76151$	$0.0^{2}62126$
5	0.22361	0.11737	0.07362	0.05077	0.03721	0.02848	0.02251	0.01825	0.01509
6	0.30171	0.17489	0.11646	0.08366	0.06319	0.04948	0.03983	0.03277	0.02744
7	0.36840	0.22973	0.16025	0.11898	0.09213	0.07358	0.06017	0.05016	0.04247
8	0.42489*	0.28018	0.20282	0.15474	0.12237	0.09937	0.08240	0.06948	0.05940
9	0.47287	0.32589	0.24315	0.18977	0.15277	0.12588	0.10564	0.08999	0.07762
10	0.51390	0.36704	0.28081	0.22342	0.18265	0.15242	0.12929	0.11114	0.09662
11	0.54928	0.40404	0.31573	0.25538	0.21159	0.17855	0.15289	0.13250	0.11601
12	0.58003	0.43734	0.34798	0.28552	0.23936*	0.20399	0.17615	0.15379	0.13551
13	0.60696	0.46739	0.37775	0.31384	0.26585	0.22856	0.19888	0.17478	0.15491
14	0.63073	0.49458	0.40521	0.34039	0.29101	0.25218	0.22093	0.19533	0.17405
15	0.65184	0.51927	0.43057	0.36525	0.31485	0.27479	0.24224	0.21536	0.19284
16	0.67070	0.54176	0.45402	0.38853	0.33742	0.29639	0.26277	0.23478	0.21118
17	0.68766	0.56231	0.47574	0.41034	0.35877	0.31700	0.28250*	0.25358	0.22904
18	0.70297	0.58116	0.49589	0.43078	0.37895	0.33663	0.30143	0.27172	0.24637
19	0.71687	0.59850	0.51464	0.44996	0.39805	0.35534	0.31957	0.28922	0.26317
20	0.72954	0.61449	0.53209	0.46798	0.41611	0.37315	0.33696	0.30607	0.27943
21	0.74113	0.62929	0.54839	0.48492	0.43322	0.39012	0.35360	0.32228	0.29514
22	0.75178	0.64301	0.56363	0.50088	0.44942	0.40629	0.36954	0.33788	0.31032
23	0.76160	0.65577	0.57790	0.51592	0.46479	0.42169	0.38481	0.35288	0.32498
24	0.77067	0.66766	0.59130	0.53013	0.47938	0.43639	0.39943	0.36730	0.33912
25	0.77908	0.67878	0.60389	0.54355	0.49324	0.45041	0.41344	0.38117	0.35277
26	0.78690	0.68918	0.61575	0.55626	0.50641	0.46380	0.42686	0.39451	0.36594
27	0.79418	0.69894	0.62694	0.56831	0.51895	0.47659	0.43974	0.40734	0.37865
28	0.80099	0.70811	0.63751	0.57973	0.53090	0.48882	0.45209	0.41969	0.39091
29	0.80736	0.71675	0.64750	0.59059	0.54229	0.50052	0.46394	0.43158	0.40275
30	0.81334	0.72489	0.65697	0.60091	0.55316	0.51173	0.47532	0.44303	0.41418
40	0.85759	0.78644	0.72982	0.68163	0.63943*	0.60186	0.56807	0.53743	0.50948
60	0.90344	0.85259	0.81066	0.77381	0.74058	0.71019	0.68215	0.65610	0.63180
80	0.92696	0.88750	0.85435	0.82474	0.79764	0.77249	0.74897	0.72684	0.70593
100	0.94128	0.90905	0.88168	0.85699	0.83418	0.81284	0.79270	0.77360	0.75541

* the corresponding value given in Table B4 of (Rencher, 1998) differs slightly from the exact value shown in this table

APPENDIX

The basic notations and properties of elementary functions that we used in the study are given here. The derivations and proofs of these results can be seen from (Mathai, 1993).

(i) $B(\alpha,\beta)=\dfrac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$.

(ii) For a non-negative integer r,

$$(a)_r=(a+r-1)(a+r-2)\ldots(a)=\frac{\Gamma(a+r)}{\Gamma(a)};\ (a)_0=1, a\neq 0, \tag{A1}$$

when $\Gamma(a)$ is defined.

(iii) For $|z|<1$,

$$(1-z)^{-a}=\sum_{r=0}^{\infty}\frac{(a)_r}{r!}z^r. \tag{A2}$$

(iv) $\Gamma(z)\Gamma(z+\frac{1}{2})=\pi^{\frac{1}{2}}2^{1-2z}\Gamma(2z)$. (A3)

(v) For $|z|<1$,

$$G_{2,2}^{2,0}\left(z\middle|\begin{matrix}\gamma_1+\delta_1-1,\gamma_2+\delta_2-1\\ \gamma_1-1,\gamma_2-1\end{matrix}\right)$$

$$=\frac{z^{\gamma_2-1}(1-z)^{\delta_1+\delta_2-1}}{\Gamma(\delta_1+\delta_2)}\,{}_2F_1(\gamma_2+\delta_2-\gamma_1,\delta_1;\delta_1+\delta_2;1-z), \tag{A4}$$

where

$${}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q;z)=\sum_{r=0}^{\infty}\frac{(a_1)_r\ldots(a_p)_r}{(b_1)_r\ldots(b_q)_r}\frac{z^r}{r!}. \tag{A5}$$

(vi)

$$G_{2,2}^{2,0}\left(z\middle|\begin{matrix}\alpha+\beta-1,\alpha+\beta-\frac{1}{2}\\ \alpha-1,\alpha-\frac{1}{2}\end{matrix}\right)=2^{2\beta-1}G_{1,1}^{1,0}\left(z^{\frac{1}{2}}\middle|\begin{matrix}2\alpha+2\beta-2\\ 2\alpha-2\end{matrix}\right). \tag{A6}$$

(vii) For $|z|<1, \beta>-1$,

$$G_{1,1}^{1,0}\left(z\left|\begin{matrix}\alpha+\beta+1\\ \alpha\end{matrix}\right.\right)=z^{\alpha}G_{1,1}^{1,0}\left(z\left|\begin{matrix}\beta+1\\ 0\end{matrix}\right.\right)=\frac{z^{\alpha}(1-z)^{\beta}}{\Gamma(\beta+1)}. \quad \text{(A7)}$$

Department of Statistics SEEMON THOMAS
St. Thomas College, Pala ALEX THANNIPPARA
Mahatma Gandhi University, Kottayam
Kerala-686 574, India

REFERENCES

T.W. ANDERSON, (2003), *An Introduction to Multivariate Statistical Analysis*, John Wiley & Sons, New Jersey.

M. BILODEAU, D. BRENNER, (1999), *Theory of Multivariate Statistics*, Springer-Verlag, New York.

C.A. COELHO, (1998), *The generalized integer gamma distribution-A basis for distributions in multivariate statistics*, "Journal of Multivariate Analysis", 64, pp. 86-102.

R.J. CONNOR, J.E. MOSIMANN, (1969), *Concepts of independence for proportions with a generalization of the Dirichlet distribution*, "Journal of American Statistical Association.", 64, pp. 194-206.

R.H.A. LOCHNER, (1975), *A generalized Dirichlet distribution in Bayesian life testing*, "Journal of the Royal Statistical Society", series B, 37, pp. 103-113.

A.M. MATHAI, (1971), *On the distribution of the likelihood ratio criterion for testing linear hypotheses on regression coefficients*, "The Annals of Institute of Statistical Mathematics", 23, pp. 181-197.

A.M. MATHAI, (1993), *A Handbook of Generalized Special Functions for Statistical and Physical Sciences*, Oxford University Press, Oxford.

K.C.S. PILLAI, A.K. GUPTA, (1969), *On the exact distribution of Wilks' criterion*, "Biometrika", 56, pp. 109-118.

A.C. RENCHER, (1998), *Multivariate Statistical Inference and Applications*, John Wiley & Sons, New York.

M. SCHATZOFF, (1966), *Exact distributions of Wilks' likelihood ratio criterion*, "Biometrika", 53, pp. 347-358.

S. THOMAS, S. GEORGE, (2004), *A review of Dirichlet distribution and its generalizations*, "Journal of the Indian Society for Probability and Statistics", 8, pp. 72-91.

S. THOMAS, A. THANNIPPARA, (2008), *Distribution of the* Λ *-criterion for sphericity test and its connection to a generalized Dirichlet model*, "Communicatoins in Statistics - Simulation and Computation", 37, pp. 1385-1395.

SUMMARY

Distribution of the LR criterion $U_{p,m,n}$ *as a marginal distribution of a generalized Dirichlet model*

The density of the likelihood ratio criterion $U_{p,m,n}$ is expressed in terms of a marginal density of a generalized Dirichlet model having a specific set of parameters. The exact distribution of the likelihood ratio criterion so obtained has a very simple and general format for every p. It provides an easy and direct method of computation of the exact p-value of $U_{p,m,n}$. Various types of properties and relations involving hypergeometric series are also established.