THE EXACT AND NEAR-EXACT DISTRIBUTIONS FOR THE WILKS LAMBDA STATISTIC USED IN THE TEST OF INDEPENDENCE OF TWO SETS OF VARIABLES

Luís Miguel Grilo
Mathematics Department of the Polytechnic Institute of Tomar
Center for Mathematics and its Applications to the Faculty of Sciences and Technology, New University of Lisbon, Portugal.
lgrilo@ipt.pt

Carlos Agra Coelho
Mathematics Department and Center for Mathematics and its Applications
Faculty of Sciences and Technology, New University of Lisbon, Portugal.
cmac@fct.unl.pt

ABSTRACT: We develop the exact distribution of the Wilks Lambda statistic to test the independence of two sets of variables, both with an odd number of variables, under the form of an infinite mixture of Generalized Integer Gamma distributions. Based on truncations of the exact characteristic function, for the product of independent Beta random variables, we obtain near-exact distributions for such product and then by direct application of these results, and once again based on truncations, we develop near-exact distributions for the Wilks Lambda statistic. These near-exact distributions are finite mixtures of Generalized Integer Gamma and Generalized Near-Integer Gamma distributions. By construction, the two first moments of these approximations are equal to the exact moments. These distributions are manageable and relatively easy to implement computationally, allowing for the computation of near-exact quantiles which may indeed be regarded as virtually exact, given the good convergence properties of the series involved, mainly when the difference between the sample size and the overall number of variables involved is rather small. We assess the proximity between these near-exact distributions and the exact distribution by using two measures based on the Berry-Esseen bounds.

Key words and Phrases: Beta and Gamma random variables, sum of log Beta and sum of Gamma random variables, mixtures, Wilks Lambda statistic, proximity measures.

1. INTRODUCTION

The purpose of this paper is to study the distribution of the Wilks Λ statistic, (Wilks, 1932, 1935) used in the test of independence of two sets of random variables (r.v.'s), both with an odd number of variables, in cases where the exact distributions are not known, or being known are too complicated to handle for practical use.

Let $\underline{X}$ be a random vector with dimension *p*, where the r.v.'s have a joint Normal *p*-multivariate distribution, $N_p(\mu,\Sigma)$. Let us consider $\underline{X}$ split in two subvectores, where the *k*-th subvector has $p_k\,(k=1,2)$ variables and $p=p_1+p_2$ is the overall number of variables. Then, each subvector $\underline{X}_k\,(k=1,2)$ will have a joint Normal p_k - multivariate distribution, $N_{p_k}(\mu_k,\Sigma_{kk})$. Symbolically,

$$\underline{X}=[\underline{X}'_1,\underline{X}'_2]' \sim N_p\left(\underline{\mu},\Sigma\right) \tag{1}$$

where

$$\underline{\mu}=[\underline{\mu}'_1,\underline{\mu}'_2]', \qquad\qquad \Sigma=\begin{bmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{bmatrix}.$$

The Wilks Λ statistic is then defined as

$$\Lambda=\frac{|V|}{|V_{11}||V_{22}|},$$

where $|\,.\,|$ stands for the determinant and *V* is either the Maximum Likelihood Estimator (MLE) of Σ or the sample variance-covariance matrix of $\underline{X}$, and V_{kk} is either the MLE of Σ_{kk} or the sample variance-covariance matrix of $\underline{X}_k\,(k=1,2)$. For a sample of size *n*, the Wilks Λ statistic is the (2/*n*)th power of the likelihood ratio test statistic to test the null hypothesis of independence of the two sets of variables, that is

$$H_0:\Sigma=\mathrm{diag}(\Sigma_{11},\Sigma_{22})\,. \tag{2}$$

According to Theorem 9.3.3 in Anderson (1984), under the null hypothesis (2), the Wilks Λ statistic has the same distribution as $\prod_{j=1}^{p_1}X_j$; where, for a sample of size *n* + 1, and $n\geq p_1+p_2$, $X_j\ (j=1,\dots,p_1)$ are p_1 independent Beta r.v.'s with,

$$X_j \sim Beta\left(\frac{n+1-p_2-j}{2},\frac{p_2}{2}\right),\quad j=1,\dots,p_1\,. \tag{3}$$

Then, for a sample of size *n* + 1, we may write the *h*-th moment of the Wilks Λ statistic, under (2), as

$$E(\Lambda^h)=\prod_{j=1}^{p_1}\frac{\Gamma\left(\frac{n+1-j}{2}\right)}{\Gamma\left(\frac{n+1-j}{2}+h\right)}\frac{\Gamma\left(\frac{n+1-p_2-j}{2}+h\right)}{\Gamma\left(\frac{n+1-p_2-j}{2}\right)} \tag{4}$$

and, since the Gamma functions in (4) are still valid for any strictly complex *h*, for a sample of size *n* + 1, the characteristic function (c.f.) of the r.v. $W=-\ln\Lambda$ is given by

$$\varphi_W(t)=E(\mathrm{e}^{\mathrm{i}tW})=E(\mathrm{e}^{-\mathrm{i}t\ln\Lambda})=E(\Lambda^{-\mathrm{i}t})=\prod_{j=1}^{p_1}\frac{\Gamma\left(\frac{n+1-j}{2}\right)}{\Gamma\left(\frac{n+1-j}{2}-\mathrm{i}t\right)}\frac{\Gamma\left(\frac{n+1-p_2-j}{2}-\mathrm{i}t\right)}{\Gamma\left(\frac{n+1-p_2-j}{2}\right)}, \tag{5}$$

where $\mathrm{i}=(-1)^{1/2}$ and $t\in\mathbb{R}$ (where $\mathbb{R}$ is the set of reals).

We start by developing the exact distribution for the product of an odd number of independent Beta r.v.'s with distributions given by (3). By direct application of this result, and taking the c.f. in (5) as a basis, we then develop the exact distribution for the Wilks Λ statistic (Grilo, 2005). These distributions are manageable and expressed under the form of an infinite mixture of Generalized Integer Gamma (GIG) distributions, where the associate series have better convergence properties than the mixtures of GIG distributions obtained by Coelho et al. (2006). We also obtain a family of near-exact distributions, based on truncations of the exact c.f., as a finite mixture of GIG distributions and Generalized Near-Integer

Gamma (GNIG) distributions. These near-exact distributions are built in the following way: one of the factors of the original c.f., that corresponds to the c.f. of a log Beta r.v. (a r.v. which Exponential has a Beta distribution), is expressed as an infinite mixture of Exponential distributions; then we truncate the associated series and we approximate the rest by the c.f. of a Gamma distribution that matches the two first moments; by joining this changed factor with the remaining unchanged part of the original c.f., we get what we call a near-exact c.f. that corresponds to a near-exact distribution expressed under the form of a finite mixture (Grilo, 2005).

These approximations lay closer to the exact distributions than the near-exact distributions based on factorizations of the exact c.f. (Coelho, 2003, 2004; Grilo, 2005; Grilo and Coelho, 2007; Alberto and Coelho, 2007) and correspond to a probability density function (p.d.f.) and a cumulative distribution function (c.d.f.) which are practical to use, allowing for an easy determination of quantiles. We analyze the behavior of these near-exact distributions, based on truncations of the exact c.f., by using two measures of proximity and we also compare them with the recent near-exact distributions, based on factorizations of the exact c.f., obtained by Grilo and Coelho (2007) for the product of an odd number of particular independent Beta r.v.'s and Alberto and Coelho (2007) for the Wilks Λ statistic.

2. Some useful distributions

Since our near-exact distributions are finite mixtures of GIG and GNIG distributions we introduce them now and also the useful log Beta distribution.

Let $X_1,...,X_g$ be g independent Gamma r.v.'s with shape parameters $r_1,...,r_g \in \mathbb{N}$ (where $\mathbb{N}$ is the set of positive integers) and all different rate parameters $\lambda_1,...,\lambda_g \in \mathbb{R}^+$, then $Z=\sum_{i=1}^{g} X_i$ is a r.v. with a GIG distribution of depth g (Coelho, 1998, 2003), denoted by

$$Z \sim GIG(r_1,...r_g;\lambda_1,...,\lambda_g)\,.$$

The p.d.f. of Z is given by

$$f_Z(z)=K\sum_{i=1}^{g} P_i(z)\mathrm{e}^{-\lambda_i z}, \qquad (z>0) \tag{6}$$

where

$$K=\prod_{i=1}^{g}\lambda_i^{r_i} \tag{7}$$

and $P_i(z)$ is a polynomial of degree r_i-1 in z, which may be written as

$$P_i(z)=\sum_{k=1}^{r_i} c_{i,k}\, z^{k-1} \tag{8}$$

where

$$c_{i,r_i}=\frac{1}{(r_i-1)!}\prod_{\substack{j=1\\ j\neq i}}^{g}(\lambda_j-\lambda_i)^{-r_j} \tag{9}$$

and, for $k=1,...,r_i-1$,

$$c_{i,r_i-k}=\frac{1}{k}\sum_{j=1}^{k}\frac{(r_i-k+j-1)!}{(r_i-k-1)!}R(j-1,i)\,c_{i,r_i-(k-j)}\,, \tag{10}$$

where

$$R(n,j)=\sum_{\substack{i=1\\ i\neq j}}^{g} r_i(\lambda_j-\lambda_i)^{-n-1}, \quad (n=0,...,r_i-1)\,. \tag{11}$$

The c.d.f. of Z is given by

$$F_Z(z)=K\sum_{i=1}^{g} P_i^*(z), \qquad (z>0) \tag{12}$$

with K given by (7) and where

$$P_i^*(z) = \sum_{k=1}^{r_i} c_{i,k} \frac{(k-1)!}{\lambda_i^k} \left[1 - \left(\sum_{j=0}^{k-1} \frac{\lambda_i^j z^j}{j!} \right) e^{-\lambda_i z} \right] \tag{13}$$

with $c_{i,k}$ $(i = 1,\dots,g; k = 1,\dots,r_i)$ given by (9) through (11).

If the r.v. Z has a distribution that is a mixture, with k components, each of which is a GIG distribution, the j-th component with weight π_j and depth g_j, we denote this fact by

$$Z \sim MkGIG(\pi_1; r_{11},\dots,r_{g_1 1}; \lambda_{11},\dots,\lambda_{g_1 1} \mid \dots \mid \pi_k; r_{1k},\dots,r_{g_k k}; \lambda_{1k},\dots,\lambda_{g_k k}) .$$

Let us consider, now, $Z \sim GIG(r_1,\dots r_g; \lambda_1,\dots,\lambda_g)$ and $X \sim Gamma(r,\lambda)$ two independent r.v.'s with $r \in \mathbb{R}^+ \setminus \mathbb{N}$ and $\lambda \neq \lambda_j$, $\forall_j \in \{ j = 1,\dots,g \}$. Then the r.v. $W = Z + X$ has a GNIG distribution with depth $g + 1$ (Coelho, 2004). Symbolically,

$$W \sim GNIG(r_1,\dots r_g, r; \lambda_1,\dots,\lambda_g, \lambda) . \tag{14}$$

The p.d.f. of W is given by

$$f_W(w) = K\lambda^r \sum_{j=1}^{g} e^{-\lambda_j w} \sum_{k=1}^{r_j} \left\{ c_{j,k} \frac{\Gamma(k)}{\Gamma(k+r)} w^{k+r-1} \, {}_1F_1(r, k+r, -(\lambda-\lambda_j)w) \right\}, \qquad (w > 0) \tag{15}$$

and the c.d.f. by

$$\begin{aligned} F_W(w) = {} & \lambda^r \frac{w^r}{\Gamma(r+1)} \, {}_1F_1(r, r+1, -\lambda w) \\ & - K\lambda^r \sum_{j=1}^{g} e^{-\lambda_j w} \sum_{k=1}^{r_j} c_{j,k}^* \sum_{i=0}^{k-1} \frac{w^{r+i} \lambda_j^i}{\Gamma(r+1+i)} \, {}_1F_1(r, r+1+i, -(\lambda-\lambda_j)w), \qquad (w > 0) \end{aligned} \tag{16}$$

where

$$K = \prod_{j=1}^{g} \lambda_j^{r_j} \quad \text{and} \quad c_{jk}^* = \frac{c_{jk}}{\lambda_j^k} \Gamma(k)$$

with $c_{j,k}$ given by (9) through (11). In the above expressions

$$\begin{aligned} {}_1F_1(a,b,z) &= \frac{\Gamma(b)}{\Gamma(a)} \sum_{j=0}^{\infty} \frac{\Gamma(a+j)}{\Gamma(b+j)} \frac{z^j}{j!} \\ &= \frac{\Gamma(b)}{\Gamma(b-a)\Gamma(a)} \int_0^1 e^{zt} t^{a-1} (1-t)^{b-a-1} dt \qquad (a \neq b) \end{aligned}$$

is the Kummer confluent hypergeometric function (Abramowitz and Stegun, 1974) that has good convergence properties and is nowadays handled by a number of software packages, like Mathematica.

The c.f. of W is given by

$$\varphi_W(t) = \lambda^r (\lambda - \mathrm{i}t)^{-r} \prod_{j=1}^{g} \lambda_j^{r_j} (\lambda_j - \mathrm{i}t)^{-r_j} , \tag{17}$$

where $r \in \mathbb{R}^+ \setminus \mathbb{N}$, $\lambda \in \mathbb{R}^+$, $r_j \in \mathbb{N}$ and $\lambda \neq \lambda_j$, $\forall j \in \{1,\dots,g\}$. If $r \in \mathbb{N}$ then the GNIG distribution of depth $g+1$ reduces to a GIG distribution of depth $g+1$. This way we may look at the GNIG distribution as a generalization of the GIG distribution.

Let X be a r.v. with a Beta distribution, with parameters $\alpha > 0$ and $\beta > 0$, what we denote by

$$X \sim Beta(\alpha,\beta) .$$

The h-th moment of X is

$$E(X^h) = \frac{B(\alpha+h,\beta)}{B(\alpha,\beta)} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)} \frac{\Gamma(\alpha+h)}{\Gamma(\alpha+\beta+h)} \qquad (h > -\alpha) . \tag{18}$$

Then $Y = -\ln X$ is a r.v. with a log Beta distribution with parameters α and β (Johnson et al., 1995), denoted by

$$Y \sim \log Beta(\alpha,\beta)\,. \tag{19}$$

The p.d.f. of Y is

$$f_Y(y)=\frac{1}{B(\alpha,\beta)}\mathrm{e}^{-\alpha y}(1-\mathrm{e}^{-y})^{\beta-1}, \qquad (y>0)\,. \tag{20}$$

Since the Gamma functions in (18) are still defined for any strictly complex h, the c.f. of Y is given by

$$\varphi_Y(t)=E(\mathrm{e}^{\mathrm{i}tY})=E(\mathrm{e}^{-\mathrm{i}t\ln X})=E(X^{-\mathrm{i}t})=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)}\frac{\Gamma(\alpha-\mathrm{i}t)}{\Gamma(\alpha+\beta-\mathrm{i}t)} \tag{21}$$

where $\mathrm{i}=(-1)^{1/2}$ and $t\in\mathbb{R}$.

We may also write the c.f. of Y in (19) as (Coelho et al., 2006),

$$\varphi_Y(t)=\frac{1}{B(\alpha,\beta)}\frac{1}{\Gamma(1-\beta)}\sum_{j=0}^{\infty}\frac{\Gamma(1-\beta+j)}{j!\;(\alpha+j)}(\alpha+j)(\alpha+j-\mathrm{i}t)^{-1}\,, \tag{22}$$

which is the c.f. of an infinite mixture of $Gamma(1,\alpha+j)\equiv Exponential(\alpha+j)$ distributions with weights given by,

$$\pi_j=\frac{1}{B(\alpha,\beta)}\frac{\Gamma(1-\beta+j)}{\Gamma(1-\beta)\,j!(\alpha+j)} \quad (j=0,1,\ldots).$$

3. The exact and near-exact distributions of the product of an odd number of particular independent Beta random variables

3.1. The exact distribution for the product of variables with Beta distribution

In Theorem 1 we present the exact distribution of the product of an odd number of independent r.v.'s with Beta distribution, expressed under the form of an infinite mixture of GIG distributions.

Theorem 1. *Let*

$$X_j \sim Beta\left(a_j,\frac{b}{2}\right), \quad j=1,\ldots,p \tag{23}$$

be p independent r.v.'s where p and b are both positive odd integers, with $b\geq 3$, and $a_j=c+\frac{p}{2}-\frac{j}{2}$ $(j=1,\ldots,p)$, with $c\in\mathbb{R}^+$. Let us consider the r.v.'s

$$W_1'=\prod_{j=1}^{p}X_j \qquad \textit{and} \qquad W_1=-\ln W_1'=-\sum_{j=1}^{p}\ln X_j\,.$$

Let us, also, consider b independent r.v.'s

$$X_j^* \sim Beta\left(a_j^*,\frac{p}{2}\right), \quad j=1,\ldots,b\,, \tag{24}$$

where $a_j^*=c^*+\frac{b}{2}-\frac{j}{2}$ $(j=1,\ldots,b)$, with $c^*\in\mathbb{R}^+$, and let

$$W_2'=\prod_{j=1}^{b}X_j^* \qquad \textit{and} \qquad W_2=-\ln W_2'=-\sum_{j=1}^{b}\ln X_j^*\,.$$

Then the exact distribution of W_1 and W_2 is

$$W_l \sim MGIG\left(\frac{1}{B\left(c,\frac{b}{2}\right)\Gamma\left(1-\frac{b}{2}\right)}\frac{\Gamma\left(1-\frac{b}{2}+k\right)}{k!(c+k)};r_{k,1}^*,\ldots,r_{k,p+b-3}^*;\lambda_1,\ldots,\lambda_{p+b-3}\left(k=1,\ldots,\frac{p+b}{2}-2\right);\right.$$
$$\left. r_1,\ldots,r_{p+b-2};\lambda_1,\ldots,\lambda_{p+b-2};\left(k=0,k\geq\frac{p+b}{2}-1\right)\right),\quad (l=1,2) \tag{25}$$

which is an infinite mixture of GIG distributions, $\frac{p+b}{2}-2$ *of them with depth* $p+b-3$ *and the remaining with depth* $p+b-2$*. For* $k=1,\ldots,\frac{p+b}{2}-2$ *the rate parameters are*

$$\lambda_j = c+\frac{j}{2} \quad (j=1,\ldots,p+b-3)\,, \tag{26}$$

and the shape parameters are

$$r^*_{k\,j} = \begin{cases} r_j+1 & j=2k \\ r_j & j=1,\ldots,p+b-3;\ j\neq 2k \end{cases} \tag{27}$$

where

$$r_j = \begin{cases} h_j & j=1,2 \\ r_{j-2}+h_j & j=3,\ldots,p+b-3 \end{cases} \tag{28}$$

with

$$h_j = \begin{cases} 1 & j=1,\ \ldots,\min(p-1,b) \\ 0 & j=1+\min(p-1,b),\ \ldots,\ \max(p-1,b) \\ \text{-}1 & j=1+\max(p-1,b),\ \ldots,\ p+b-3 \end{cases} \tag{29}$$

or, alternatively,

$$h_j = (\#\text{ of elements of } \{p-1,b\}\geq j)-1\,; \tag{30}$$

and, for $k=0$ *and* $k\geq\frac{p+b}{2}-1$ *the shape parameters* r_j $(j=1,\ldots,p+b-3)$ *are given by (28) through (30) and* $r_{p+b-2}=1$*, the rate parameters* λ_j $(j=1,\ldots,p+b-3)$ *are given by* (26) *and* $\lambda_{p+b-2}=c+k$.

The distribution of W_2 *is the same as the distribution of* W_1 *and thus, also the distribution of* W_2' *is the same as the distribution of* W_1'.

Proof. Based on (18), we have

$$E(X_j^h)=\frac{\Gamma\left(a_j+\frac{b}{2}\right)}{\Gamma(a_j)}\frac{\Gamma(a_j+h)}{\Gamma\left(a_j+\frac{b}{2}+h\right)},\quad (h>-a_j),\quad j=1,\ldots,p$$

and given the independence of the *p* r.v.'s X_j $(j=1,\ldots,p)$,

$$E(W_1'^h)=E\left(\prod_{j=1}^p X_j^h\right)=\prod_{j=1}^p E(X_j^h)=\prod_{j=1}^p\frac{\Gamma\left(a_j+\frac{b}{2}\right)}{\Gamma(a_j)}\frac{\Gamma(a_j+h)}{\Gamma\left(a_j+\frac{b}{2}+h\right)}\quad (h>-a_p). \tag{31}$$

Since the Gamma functions in (31) are still defined for any strictly complex *h*, the c.f. of W_1 is then given by

$$\varphi_{W_1}(t)=E(\mathrm{e}^{\mathrm{i}tW_1})=E(\mathrm{e}^{-\mathrm{i}t\ln W_1'})=E(W_1'^{-\mathrm{i}t})=\prod_{j=1}^p\frac{\Gamma\left(a_j+\frac{b}{2}\right)}{\Gamma(a_j)}\frac{\Gamma(a_j-\mathrm{i}t)}{\Gamma\left(a_j+\frac{b}{2}-\mathrm{i}t\right)}, \tag{32}$$

where $\mathrm{i}=(-1)^{1/2}$ and $t\in\mathbb{R}$.

Replacing in (32), the c.f. of the *p*-th log Beta by (22) with parameters $\alpha=c$ and $\beta=b/2$ and considering the result of Section 4 in Coelho (2004), a by-product of the proof of Theorem 2 in Coelho (1998), which for even *p* states that for $a_j=c+\frac{p}{2}-\frac{j}{2}$,

$$\prod_{j=1}^{p}\frac{\Gamma\left(a_j+\frac{b}{2}\right)}{\Gamma(a_j)}=\prod_{j=1}^{p}\frac{\Gamma\left(c+\frac{p}{2}-\frac{j}{2}+\frac{b}{2}\right)}{\Gamma\left(c+\frac{p}{2}-\frac{j}{2}\right)}=\prod_{j=1}^{p+b-2}\left(c+\frac{j}{2}-\frac{1}{2}\right)^{r_j}$$

where

$$r_j=\begin{cases}h_j & j=1,2\\ r_{j-2}+h_j & j=3,\dots,p+b-2\end{cases} \tag{33}$$

with

$$h_j=(\#\text{ of elements of }\{p,b\}\geq j)-1\,, \tag{34}$$

what for odd p yields

$$\prod_{j=1}^{p-1}\frac{\Gamma\left(a_j+\frac{b}{2}\right)}{\Gamma(a_j)}=\prod_{j=1}^{p-1}\frac{\Gamma\left(c+\frac{1}{2}+\frac{p-1}{2}-\frac{j}{2}+\frac{b}{2}\right)}{\Gamma\left(c+\frac{1}{2}+\frac{p-1}{2}-\frac{j}{2}\right)}=\prod_{j=1}^{p+b-3}\left(c+\frac{j}{2}\right)^{r_j}\,, \tag{35}$$

with r_j given by (33) and (34), with p replaced by $p-1$, the c.f. of W_1 may be written as

$$\begin{aligned}
\varphi_{W_1}(t)&=\prod_{j=1}^{p}\frac{\Gamma\left(a_j+\frac{b}{2}\right)}{\Gamma(a_j)}\frac{\Gamma(a_j-\mathrm{i}t)}{\Gamma\left(a_j+\frac{b}{2}-\mathrm{i}t\right)}\\
&=\frac{\Gamma\left(a_p+\frac{b}{2}\right)}{\Gamma(a_p)}\frac{\Gamma(a_p-\mathrm{i}t)}{\Gamma\left(a_p+\frac{b}{2}-\mathrm{i}t\right)}\prod_{j=1}^{p-1}\frac{\Gamma\left(a_j+\frac{b}{2}\right)}{\Gamma(a_j)}\frac{\Gamma(a_j-\mathrm{i}t)}{\Gamma\left(a_j+\frac{b}{2}-\mathrm{i}t\right)}\\
&=\frac{1}{B\left(c,\frac{b}{2}\right)}\frac{1}{\Gamma\left(1-\frac{b}{2}\right)}\sum_{k=0}^{\infty}\frac{\Gamma\left(1-\frac{b}{2}+k\right)}{k!(c+k)}(c+k)(c+k-\mathrm{i}t)^{-1}\prod_{j=1}^{p+b-3}\left(c+\frac{j}{2}\right)^{r_j}\left(c+\frac{j}{2}-\mathrm{i}t\right)^{-r_j}\\
&=\frac{1}{B\left(c,\frac{b}{2}\right)}\frac{1}{\Gamma\left(1-\frac{b}{2}\right)}\left[\sum_{k=1}^{\frac{p+b}{2}-2}\frac{\Gamma\left(1-\frac{b}{2}+k\right)}{k!(c+k)}\prod_{j=1}^{p+b-3}\left(c+\frac{j}{2}\right)^{r^*_{kj}}\left(c+\frac{j}{2}-\mathrm{i}t\right)^{-r^*_{kj}}\right.\\
&\qquad\left.+\sum_{k=0,k\geq\frac{p+b}{2}-1}^{\infty}\frac{\Gamma\left(1-\frac{b}{2}+k\right)}{k!(c+k)}(c+k)(c+k-\mathrm{i}t)^{-1}\prod_{j=1}^{p+b-3}\left(c+\frac{j}{2}\right)^{r_j}\left(c+\frac{j}{2}-\mathrm{i}t\right)^{-r_j}\right]
\end{aligned} \tag{36}$$

which is the c.f. of an infinite mixture of GIG distributions, $\frac{p+b}{2}-2$ of them with depth $p+b-3$, with shape parameters r^*_{kj} given by (27) through (30) and rate parameters λ_j given by (26) $(k=1,\dots,\frac{p+b}{2}-2;\ j=1,\dots,p+b-3)$, and the remaining GIG distributions with depth $p+b-2$, with shape parameters r_j given by (28) through (30) and $r_{p+b-2}=1$, and the rate parameters λ_j given by (26) and $\lambda_{p+b-2}=c+k$ $(k=0,k\geq\frac{p+b}{2}-1;\ j=1,\dots,p+b-3)$.

To confirm that the c.f. of W_2 is the same as the c.f. of W_1 (and, consequently, p and b are interchangeable) we just have to consider the equality (Coelho, 1998, 1999),

$$\prod_{j=1}^{p}\frac{\Gamma\left(c+\frac{p}{2}-\frac{j}{2}+\frac{b}{2}\right)}{\Gamma\left(c+\frac{p}{2}-\frac{j}{2}\right)}=\prod_{j=1}^{b}\frac{\Gamma\left(c+\frac{b}{2}-\frac{j}{2}+\frac{p}{2}\right)}{\Gamma\left(c+\frac{b}{2}-\frac{j}{2}\right)},$$

which is valid for p and b positive integers and $c\in\mathbb{R}^+$. Then, a direct application of this result to the c.f. of W_2 leads us immediately to the c.f. of W_1,

$$\begin{aligned}\varphi_{W_2}(t)=E(\mathrm{e}^{\mathrm{i}tW_2})&=\prod_{j=1}^{b}\frac{\Gamma\left(a_j^*+\frac{p}{2}\right)}{\Gamma(a_j^*)}\frac{\Gamma(a_j^*-\mathrm{i}t)}{\Gamma\left(a_j^*+\frac{p}{2}-\mathrm{i}t\right)}\\
&=\prod_{j=1}^{b}\frac{\Gamma\left(c+\frac{p}{2}-\frac{j}{2}+\frac{b}{2}\right)}{\Gamma\left(c+\frac{p}{2}-\frac{j}{2}\right)}\frac{\Gamma\left(c+\frac{p}{2}-\frac{j}{2}-\mathrm{i}t\right)}{\Gamma\left(c+\frac{p}{2}-\frac{j}{2}+\frac{b}{2}-\mathrm{i}t\right)}\\
&=\prod_{j=1}^{p}\frac{\Gamma\left(c+\frac{b}{2}-\frac{j}{2}+\frac{p}{2}\right)}{\Gamma\left(c+\frac{b}{2}-\frac{j}{2}\right)}\frac{\Gamma\left(c+\frac{b}{2}-\frac{j}{2}-\mathrm{i}t\right)}{\Gamma\left(c+\frac{b}{2}-\frac{j}{2}+\frac{p}{2}-\mathrm{i}t\right)}\\
&=\prod_{j=1}^{p}\frac{\Gamma\left(a_j+\frac{b}{2}\right)}{\Gamma(a_j)}\frac{\Gamma(a_j-\mathrm{i}t)}{\Gamma\left(a_j+\frac{b}{2}-\mathrm{i}t\right)}=\varphi_{W_1}(t).\end{aligned}\tag{37}$$

Since $\varphi_{W_1}(t)=\varphi_{W_2}(t)$ then W_1 and W_2 have the same distribution and consequently W_1' and W_2' also have the same distribution. ▮

Based on (25) and (6), the expression for the exact p.d.f. of W_1 is then given by

$$f_{W_1}(w)=\frac{1}{B\left(c,\frac{b}{2}\right)\Gamma\left(1-\frac{b}{2}\right)}\left[\sum_{k=1}^{\frac{p+b}{2}-2}\frac{\Gamma\left(1-\frac{b}{2}+k\right)}{k!(c+k)}K_k\sum_{j=1}^{p+b-3}P_{kj}(w)\mathrm{e}^{-\lambda_j w}+\sum_{k=0,k\geq\frac{p+b}{2}-1}^{\infty}\frac{\Gamma\left(1-\frac{b}{2}+k\right)}{k!(c+k)}K\sum_{j=1}^{p+b-2}P_j(w)\mathrm{e}^{-\lambda_j w}\right]$$

where,

$$K_k=\prod_{j=1}^{p+b-3}\lambda_j^{r_{kj}^*},\qquad P_{kj}(w)=\sum_{i=1}^{r_{kj}^*}c_{kj,i}w^{i-1},$$

and

$$K=\prod_{j=1}^{p+b-2}\lambda_j^{r_j},\qquad P_j(w)=\sum_{m=1}^{r_j}c_{j,m}w^{m-1},$$

and, based on (12), the expression for the exact c.d.f. of W_1 is given by

$$F_{W_1}(w) = \frac{1}{B\left(c, \frac{b}{2}\right)\Gamma\left(1-\frac{b}{2}\right)} \left[\sum_{k=1}^{\frac{p+b}{2}-2} \frac{\Gamma\left(1-\frac{b}{2}+k\right)}{k!(c+k)} K_k \sum_{j=1}^{p+b-3} P^*_{k\,j}(w) + \sum_{k=0, k\geq\frac{p+b}{2}-1}^{\infty} \frac{\Gamma\left(1-\frac{b}{2}+k\right)}{k!(c+k)} K \sum_{j=1}^{p+b-2} P^*_j(w) \right]$$

where,

$$K_k = \prod_{j=1}^{p+b-3} \lambda_j^{r^*_{k\,j}}, \qquad P^*_{k\,j}(w) = \sum_{i=1}^{r^*_{k\,j}} c_{k\,j,i} \frac{(i-1)!}{\lambda_j^i}\left[1-\left(\sum_{i^*=0}^{i-1} \frac{\lambda_j^{i^*} w^{i^*}}{i^*!}\right) \mathrm{e}^{-\lambda_j w}\right],$$

and

$$K = \prod_{j=1}^{p+b-2} \lambda_j^{r_j}, \qquad P^*_j(w) = \sum_{m=1}^{r_j} c_{j,m} \frac{(m-1)!}{\lambda_j^m}\left[1-\left(\sum_{i=0}^{m-1} \frac{\lambda_j^i w^i}{i!}\right) \mathrm{e}^{-\lambda_j w}\right].$$

For $k=0$ and $k \geq \frac{p+b}{2}-1$ the shape parameters r_j $(j=1,\ldots,p+b-3)$ are given by (28) through (30) and $r_{p+b-2}=1$; the rate parameters λ_j $(j=1,\ldots,p+b-3)$ are given by (26); $\lambda_{p+b-2}=c+k$ and $c_{j,m}$ $(j=1,\ldots,p+b-2; m=1,\ldots,r_j)$ are given by (9) through (11). For $k=1,\ldots,\frac{p+b}{2}-2$ the shape parameters $r^*_{k\,j}$ $(j=1,\ldots,p+b-3)$ are given by (27) through (30) and the rates λ_j $(j=1,\ldots,p+b-3)$ are given by (26); and, with $c_{k\,j,i}$ $(j=1,\ldots,p+b-3; i=1,\ldots,r^*_{k\,j})$ based on (9) through (11), given by

$$c_{kj,r^*_{kj}} = \frac{1}{(r^*_{kj}-1)!} \prod_{\substack{i=1\\ i\neq j}}^{p+b-3} (\lambda_i - \lambda_j)^{-r^*_{ki}} \tag{38}$$

and, for $\tau = 1,\ldots,r^*_{kj}-1$,

$$c_{kj,r^*_{kj}-\tau} = \frac{1}{\tau}\sum_{i=1}^{\tau} \frac{(r^*_{kj}-\tau+i-1)!}{(r^*_{kj}-\tau-1)!} R(i-1,j)\, c_{kj,r^*_{kj}-(\tau-i)}\,, \tag{39}$$

where

$$R(n,j) = \sum_{\substack{i=1\\ i\neq j}}^{p+b-3} r^*_{ki}(\lambda_j-\lambda_i)^{-n-1}, \quad (n=0,\ldots,r^*_{ki}-1)\,. \tag{40}$$

The exact distributions of W'_1 and W'_2 could be easily obtained through the transformation $W'_k = \mathrm{e}^{-W_k}$ $(k=1,2)$.

3.2. NEAR-EXACT DISTRIBUTIONS FOR THE PRODUCT OF VARIABLES WITH BETA DISTRIBUTION

Based on truncations of the exact distribution we may obtain near-exact distributions, expressed as finite mixtures, that equate the two first exact moments and which allow for the computation of near-exact quantiles. These distributions are obtained in Theorem 2.

Theorem 2. *Let X_j and X^*_j be defined as in Theorem 1, in (23) and (24), respectively, and let*

$$W'_1 = \prod_{j=1}^{p} X_j, \qquad W_1 = -\ln W'_1 = -\sum_{j=1}^{p} \ln X_j\,, \tag{41}$$

and also

$$W_2' = \prod_{j=1}^{b} X_j^*\,, \qquad W_2 = -\ln W_2' = -\sum_{j=1}^{b} \ln X_j^*.$$

Then, a family of near-exact distributions for W_1 and W_2 is given by

$$W_l \overset{ne}{\sim} M\,(n^*+1)GIG+GNIG(\pi_k; r_{k1}^*,\dots,r_{k,p+b-2}^*; \lambda_1,\dots,\lambda_{p+b-2}; k=0,\dots,n^*+1), \qquad l=1,2$$

with

$$\pi_k = \frac{\Gamma\left(1-\frac{b}{2}+k\right)}{B\left(c,\frac{b}{2}\right)\Gamma\left(1-\frac{b}{2}\right)k!(c+k)} \quad (k=0,\dots,n^*) \qquad and \qquad \pi_{n^*+1} = 1-\sum_{k=0}^{n^*}\pi_k\,,$$

which is a finite mixture of n^*+2 distributions, where n^*+1 components are GIG distributions and the last component is a GNIG distribution with depth $p+b-2$.

For $n^* < \frac{p+b}{2}-2$ we have a mixture of n^*+1 GIG distributions (n^* of them with depth $p+b-3$ and one with depth $p+b-2$) and one GNIG distribution with depth $p+b-2$. For $k=1,\dots,n^* < \frac{p+b}{2}-2$, the shape parameters r_{kj}^* are given by (27) through (30) and the rates λ_j $(j=1,\dots,p+b-3)$ are given by (26)*; for $k=0$, the $r_{kj}^* = r_j$ are given by (28) through (30) and $r_{p+b-2}=1$, since the rates λ_j $(j=1,\dots,p+b-3)$ are given by* (26) *and $\lambda_{p+b-2}=c$.*

For $n^* \ge \frac{p+b}{2}-2$ we have a mixture of n^*+2 distributions, where n^*+1 components are GIG distributions ($\frac{p+b}{2}-2$ with depth $p+b-3$ and $n^*+1-\left(\frac{p+b}{2}-2\right)$ with depth $p+b-2$) and the last one is a GNIG distribution of depth $p+b-2$. For $k=1,\dots,\frac{p+b}{2}-2$ the parameters r_{kj}^* are given by (27) through (30) and the rates λ_j $(j=1,\dots,p+b-3)$ are given by (26)*; for $k=0$ and $k \ge \frac{p+b}{2}-1$, the $r_{kj}^*=r_j$ are given by (28) through (30) and $r_{p+b-2}=1$, since the rates λ_j are given by* (26) *and $\lambda_{p+b-2}=c+k$.*

Both for $n^* < \frac{p+b}{2}-2$ and for $n^* \ge \frac{p+b}{2}-2$ we have in the last component of the mixture (*i.e., the GNIG distribution with depth $p+b-2$) the $r_{kj}^*=r_j$ given by (28) through (30) and the rates λ_j $(j=1,\dots,p+b-3)$ given by* (26)*. The parameters $r_{k,p+b-2}^*=r$ and $\lambda_{p+b-2}=\lambda$ of GNIG distribution are obtained in such way that the two first moments of the near-exact distributions are equal to the two first exact moments.*

Proof. From (36) we may write the c.f. of W_1 as

$$\varphi_{W_1}(t) = \underbrace{\frac{1}{B\left(c,\frac{b}{2}\right)\Gamma\left(1-\frac{b}{2}\right)}\sum_{k=0}^{\infty}\frac{\Gamma\left(1-\frac{b}{2}+k\right)}{k!(c+k)}(c+k)(c+k-\mathrm{i}t)^{-1}}_{\varphi'(t)}\;\underbrace{\prod_{j=1}^{p+b-3}\left(c+\frac{j}{2}\right)^{r_j}\left(c+\frac{j}{2}-\mathrm{i}t\right)^{-r_j}}_{\varphi''(t)}\,,$$

where we will truncate the infinite mixture of Exponential distributions relative to the p-th log Beta and write,

$$\varphi'(t) = \sum_{k=0}^{\infty}\pi_k\varphi_{X_k}(t) = \underbrace{\sum_{k=0}^{n^*}\pi_k\varphi_{X_k}(t)}_{T_{n^*}(t)} + \underbrace{\sum_{k=n^*+1}^{\infty}\pi_k\varphi_{X_k}(t)}_{R_{n^*}(t)}\,,$$

where $\varphi_{X_k}(t)$ is the c.f. of $X_k \sim Exponential(c+k)$ $(k=0,1,\dots)$ and the weights π_k are given by

$$\pi_k = \frac{\Gamma\left(1-\frac{b}{2}+k\right)}{B\left(c,\frac{b}{2}\right)\Gamma\left(1-\frac{b}{2}\right)k!(c+k)} \quad (k=0,1,...)\,.$$

Then we approximate $R_{n^*}(t)$ by $\theta\varphi_1(t)$, where $\varphi_1(t)=\lambda^r(\lambda-\mathrm{i}t)^{-r}$ is the c.f. of a $Gamma(r,\lambda)$ r.v. and the weigth θ is given by $\theta=\sum_{k=n^*+1}^{\infty}\pi_k = 1-\sum_{k=0}^{n^*}\pi_k$. The parameters r and λ of the Gamma distribution are obtained in such a way that the two first derivatives of $R_{n^*}(t)$ and $\theta\varphi_1(t)$ with respect to t, at $t=0$, are equal, i.e., in such a way that

$$\left.\frac{d^h}{dt^h}R_{n^*}(t)\right|_{t=0} = \theta\left.\frac{d^h}{dt^h}\varphi_1(t)\right|_{t=0} \;\Leftrightarrow\; \left.\frac{d^h}{dt^h}\varphi'(t)\right|_{t=0} - \left.\frac{d^h}{dt^h}T_{n^*}(t)\right|_{t=0} = \theta\left.\frac{d^h}{dt^h}\varphi_1(t)\right|_{t=0} \qquad (h=1,2),$$

where

$$\left.\frac{d^h}{dt^h}\varphi'(t)\right|_{t=0} = \frac{d^h}{dt^h}\left.\left[\sum_{k=0}^{n^*}\pi_k\varphi_{X_k}(t)+\theta\varphi_1(t)\right]\right|_{t=0} \qquad (h=1,2),$$

that is, the two first moments of the exact and near-exact distributions are equal.

This way, we approximate $\varphi'(t)$ by

$$T_{n^*}(t)+\theta\varphi_1(t) = \left(\frac{1}{B\left(c,\frac{b}{2}\right)}\frac{1}{\Gamma\left(1-\frac{b}{2}\right)}\sum_{k=0}^{n^*}\frac{\Gamma\left(1-\frac{b}{2}+k\right)}{k!(c+k)}(c+k)(c+k-\mathrm{i}t)^{-1}\right)+\theta\lambda^r(\lambda-\mathrm{i}t)^{-r}\,,$$

and consequently, the near-exact c.f. of W_1 is given by

$$\begin{aligned}
&\left[\frac{1}{B\left(c,\frac{b}{2}\right)}\frac{1}{\Gamma\left(1-\frac{b}{2}\right)}\left(\sum_{k=0}^{n^*}\frac{\Gamma\left(1-\frac{b}{2}+k\right)}{k!(c+k)}(c+k)(c+k-\mathrm{i}t)^{-1}\right)+\theta\lambda^r(\lambda-\mathrm{i}t)^{-r}\right]\times\varphi''(t)\\
&=\left[\left(\frac{1}{B\left(c,\frac{b}{2}\right)}\frac{1}{\Gamma\left(1-\frac{b}{2}\right)}\sum_{k=0}^{n^*}\frac{\Gamma\left(1-\frac{b}{2}+k\right)}{k!(c+k)}(c+k)(c+k-\mathrm{i}t)^{-1}\right)+\theta\lambda^r(\lambda-\mathrm{i}t)^{-r}\right]\\
&\qquad\times\prod_{j=1}^{p+b-3}\left(c+\frac{j}{2}\right)^{r_j}\left(c+\frac{j}{2}-\mathrm{i}t\right)^{-r_j} \qquad\qquad (42)\\
&=\frac{1}{B\left(c,\frac{b}{2}\right)}\frac{1}{\Gamma\left(1-\frac{b}{2}\right)}\left[\sum_{k=0}^{n^*}\frac{\Gamma\left(1-\frac{b}{2}+k\right)}{k!(c+k)}(c+k)(c+k-\mathrm{i}t)^{-1}\prod_{j=1}^{p+b-3}\left(c+\frac{j}{2}\right)^{r_j}\left(c+\frac{j}{2}-\mathrm{i}t\right)^{-r_j}\right]\\
&\qquad+\theta\lambda^r(\lambda-\mathrm{i}t)^{-r}\prod_{j=1}^{p+b-3}\left(c+\frac{j}{2}\right)^{r_j}\left(c+\frac{j}{2}-\mathrm{i}t\right)^{-r_j},
\end{aligned}$$

which is the c.f. of a finite mixture of n^*+2 distributions, where n^*+1 of them are GIG distributions and the last one is a GNIG distribution. We denote these near-exact distributions by M(n^*+1)GIG+GNIG.

Then, we have to consider two situations:

- $n^* < \frac{p+b}{2} - 2$, where (42) may be written as

$$\frac{1}{B\left(c,\frac{b}{2}\right)}\frac{1}{\Gamma\left(1-\frac{b}{2}\right)}\left[\sum_{k=1}^{n^*}\frac{\Gamma\left(1-\frac{b}{2}+k\right)}{k!(c+k)}\prod_{j=1}^{p+b-3}\left(c+\frac{j}{2}\right)^{r^*_{kj}}\left(c+\frac{j}{2}-\mathrm{i}t\right)^{-r^*_{kj}}\right.$$
$$\left.+\frac{\Gamma\left(1-\frac{b}{2}\right)}{c}c(c-\mathrm{i}t)^{-1}\prod_{j=1}^{p+b-3}\left(c+\frac{j}{2}\right)^{r_j}\left(c+\frac{j}{2}-\mathrm{i}t\right)^{-r_j}\right]+\theta\lambda^r(\lambda-\mathrm{i}t)^{-r}\prod_{j=1}^{p+b-3}\left(c+\frac{j}{2}\right)^{r_j}\left(c+\frac{j}{2}-\mathrm{i}t\right)^{-r_j} \tag{43}$$

which is the c.f. of a finite mixture of n^*+2 distributions, where n^*+1 components are GIG distributions (n^* of them with depth $p+b-3$ and one with depth $p+b-2$) and the last one is a GNIG distribution of depth $p+b-2$, with shape and rate parameters mentioned in the body of Theorem 2;

- $n^* \geq \frac{p+b}{2} - 2$, where (42) may be written as

$$\frac{1}{B\left(c,\frac{b}{2}\right)}\frac{1}{\Gamma\left(1-\frac{b}{2}\right)}\left[\sum_{k=1}^{\frac{p+b}{2}-2}\frac{\Gamma\left(1-\frac{b}{2}+k\right)}{k!(c+k)}\prod_{j=1}^{p+b-3}\left(c+\frac{j}{2}\right)^{r^*_{kj}}\left(c+\frac{j}{2}-\mathrm{i}t\right)^{-r^*_{kj}}\right.$$
$$\left.+\sum_{k=0,k\geq\frac{p+b}{2}-1}^{n^*}\frac{\Gamma\left(1-\frac{b}{2}+k\right)}{k!(c+k)}(c+k)(c+k-\mathrm{i}t)^{-1}\prod_{j=1}^{p+b-3}\left(c+\frac{j}{2}\right)^{r_j}\left(c+\frac{j}{2}-\mathrm{i}t\right)^{-r_j}\right] \tag{44}$$
$$+\theta\lambda^r(\lambda-\mathrm{i}t)^{-r}\prod_{j=1}^{p+b-3}\left(c+\frac{j}{2}\right)^{r_j}\left(c+\frac{j}{2}-\mathrm{i}t\right)^{-r_j}$$

which is the c.f. of a finite mixture of n^*+2 distributions, where n^*+1 of them are GIG distributions, ($\frac{p+b}{2}-2$ with depth $p+b-3$ and $n^*+1-\left(\frac{p+b}{2}-2\right)$ with depth $p+b-2$) and the last one is a GNIG distribution of depth $p+b-2$, with shape and rate parameters mentioned in the body of Theorem 2. ∎

The near-exact c.f.'s obtained in this way are asymptotic for increasing values of n^*, in the sense that they converge to the exact c.f., $\varphi_{W_1}(t)$, when $n^* \to +\infty$.

4. The exact and near-exact distributions for the Wilks Lambda statistic for two sets of variables with an odd number of variables

4.1. The exact distribution for the Wilks Lambda statistic

In Theorem 3 we present the exact distribution for the Wilks Lambda statistic for the case of two sets of variables, both with an odd number of variables. This distribution, expressed under the form of an infinite mixture of GIG distributions, is developed by direct application of the exact distribution of the product of an odd number of independent r.v.'s with Beta distributions (Theorem 1).

Theorem 3. *If the two sets of variables both have an odd number of variables, then, under (2) and for a sample of size $n+1$, the exact distribution of $W=-\ln\Lambda$ is an infinite mixture of GIG distributions, $\frac{p}{2}-2$ with depth $p-3$ and the remaining with depth $p-2$:*

$$W \sim MGIG\left(\frac{1}{B\left(c,\frac{p_2}{2}\right)\Gamma\left(1-\frac{p_2}{2}\right)}\frac{\Gamma\left(1-\frac{p_2}{2}+k\right)}{k!(c+k)};r_{k1}^*,\dots,r_{k,p-3}^*;\lambda_1,\dots,\lambda_{p-3}\left(k=1,\dots,\frac{p}{2}-2\right);\right.$$
$$\left. r_1,\dots,r_{p-2};\lambda_1,\dots,\lambda_{p-2}\left(k=0,k\geq\frac{p}{2}-1\right)\right)$$

where $c=\frac{n+1-p}{2}$ *and* $p=p_1+p_2$. *For* $k=1,\dots,\frac{p}{2}-2$ *the rate parameters are given by*

$$\lambda_j=c+\frac{j}{2},\qquad j=1,\dots,p-3 \tag{45}$$

and the shape parameters are

$$r_{kj}^*=\begin{cases} r_j+1 & j=2k\\ r_j & j\in\{1,\dots,p-3\}\setminus\{2k\}\end{cases} \tag{46}$$

where

$$r_j=\begin{cases} h_j & j=1,2\\ r_{j-2}+h_j & j=3,\dots,p-3\end{cases} \tag{47}$$

with

$$h_j=(\#\text{ of elements of }\{p_1-1,p_2\}\geq j)-1\,. \tag{48}$$

For $k=0$ *and* $k\geq\frac{p}{2}-1$ *the shape parameters* r_j $(j=1,\dots,p-3)$ *are given by (47) and (48) and* $r_{p-2}=1$*; the rate parameters* λ_j $(j=1,\dots,p-3)$ *are given by* (45) *and* $\lambda_{p-2}=c+k$.

Proof. In order to prove this theorem we just have to consider Theorem 1 and establish the following equivalencies (bearing in mind that while in Theorem 1, *p* denotes the number of Beta r.v.'s involved, in this theorem *p* denotes the overall number of variables):

Product of Beta variables		Wilks Λ statistic
p	↔	p_1
b	↔	p_2
$p+b$	↔	$p_1+p_2=p$
a_j	↔	$\frac{n+1-p_2-j}{2}$
$c=a_j-\frac{p}{2}+\frac{j}{2}$	↔	$\frac{n+1-p_1-p_2}{2}=\frac{n+1-p}{2}$.

The c.f. of *W* may then be written as

$$\varphi_W(t)=\frac{1}{B\left(c,\frac{p_2}{2}\right)}\frac{1}{\Gamma\left(1-\frac{p_2}{2}\right)}\left[\sum_{k=1}^{\frac{p}{2}-2}\frac{\Gamma\left(1-\frac{p_2}{2}+k\right)}{k!(c+k)}\prod_{j=1}^{p-3}\left(c+\frac{j}{2}\right)^{r_{kj}^*}\left(c+\frac{j}{2}-\mathrm{i}t\right)^{-r_{kj}^*}\right.$$
$$\left.+\sum_{k=0,k\geq\frac{p}{2}-1}^{\infty}\frac{\Gamma\left(1-\frac{p_2}{2}+k\right)}{k!(c+k)}(c+k)(c+k-\mathrm{i}t)^{-1}\prod_{j=1}^{p-3}\left(c+\frac{j}{2}\right)^{r_j}\left(c+\frac{j}{2}-\mathrm{i}t\right)^{-r_j}\right], \tag{49}$$

which is the c.f. of an infinite mixture of GIG distributions, $\frac{p}{2}-2$ with depth $p-3$, with shape parameters r_{kj}^* given by (46) through (48) and rates λ_j given by (45) $(k=1,\dots,\frac{p}{2}-2;\ j=1,\dots,p-3)$, and the remaining GIG distributions have depth $p-2$, with shape parameters r_j given by (47) and (48) and $r_{p-2}=1$, since the rates λ_j are given by (45) and $\lambda_{p-2}=c+k$ $(k=0,k\geq\frac{p}{2}-1;\ j=1,\dots,p-3)$. ∎

4.2. NEAR-EXACT DISTRIBUTIONS FOR THE WILKS LAMBDA STATISTIC

Once again, based on truncations of the exact distribution and following the procedure used in Theorem 2 we obtain near-exact distributions, expressed as finite mixtures, that equate the two first exact moments and allows for the computation of near-exact quantiles. This way, we have the Theorem 4.

Theorem 4. *If the two sets of variables both have an odd number of variables, then, under (2) and for a sample of size* $n+1$, *a family of near-exact distributions for* $W=-\ln\Lambda$ *may be obtained under the form*

$$W \overset{ne}{\sim} M(n^*+1)GIG+GNIG(\pi_k;r^*_{k1},\dots,r^*_{k,p-2};\lambda_1,\dots,\lambda_{p-2};k=0,\dots,n^*+1)$$

with

$$\pi_k=\frac{1}{B\left(c,\frac{p_2}{2}\right)}\frac{\Gamma\left(1-\frac{p_2}{2}+k\right)}{\Gamma\left(1-\frac{p_2}{2}\right)k!(c+k)}\quad(k=0,\dots,n^*)\qquad and \qquad \pi_{n^*+1}=1-\sum_{k=0}^{n^*}\pi_k\ ,$$

which is a finite mixture of n^*+2 *distributions, where* n^*+1 *components are GIG distributions and the last component is a GNIG distribution.*

For $n^*<\frac{p}{2}-2$ *we have a finite mixture of* n^*+1 *GIG distributions (*n^* *with depth* $p-3$ *and one with depth* $p-2$*) and one GNIG distribution of depth* $p-2$. *For* $k=1,\dots,n^*<\frac{p}{2}-2$ *the shape parameters* r^*_{kj} *are given by (46) through (48) and the rates* λ_j $(j=1,\dots,p-3)$ *are given by* (45)*; for* $k=0$, *the* $r^*_{kj}=r_j$ *are given by (47) and (48) and* $r_{p-2}=1$, *where the rates* λ_j $(j=1,\dots,p-3)$ *are given by* (45) *and* $\lambda_{p-2}=c$.

For $n^*\geq\frac{p}{2}-2$ *we have a finite mixture of* n^*+1 *GIG distributions (*$\frac{p}{2}-2$ *with depth* $p-3$ *and* $n^*+1-\left(\frac{p}{2}-2\right)$ *with depth* $p-2$*) and one GNIG distribution of depth* $p-2$. *For* $k=1,\dots,\frac{p}{2}-2$ *the parameters* r^*_{kj} *are given by (46) through (48) and the rates* λ_j $(j=1,\dots,p-3)$ *are given by* (45)*; for* $k=0$ *and* $k\geq\frac{p}{2}-1$, *the* $r^*_{kj}=r_j$ *are given by (47) and (48) and* $r_{p-2}=1$, *with rates* λ_j *given by* (45) *and* $\lambda_{p-2}=c+k$.

Both for $n^*<\frac{p}{2}-2$ *and* $n^*\geq\frac{p}{2}-2$ *we have for the last component of the mixture (i.e., in the GNIG distribution of depth* $p-2$*) the* $r^*_{kj}=r_j$ *given by (47) and (48) and the rate parameters* λ_j $(j=1,\dots,p-3)$ *given by* (45)*. The parameters* $r^*_{k,p-2}=r$ *and* $\lambda_{p-2}=\lambda$ *of the GNIG distribution are obtained in such way that the two first exact moments are equal.*

Proof. Appling the result in (42) and using the equivalences in the proof of Theorem 3, we may write the c.f. of W under the form

$$\frac{1}{B\left(c,\frac{p_2}{2}\right)}\frac{1}{\Gamma\left(1-\frac{p_2}{2}\right)}\left[\sum_{k=0}^{n^*}\frac{\Gamma\left(1-\frac{p_2}{2}+k\right)}{k!(c+k)}(c+k)(c+k-\mathrm{i}t)^{-1}\prod_{j=1}^{p-3}\left(c+\frac{j}{2}\right)^{r_j}\left(c+\frac{j}{2}-\mathrm{i}t\right)^{-r_j}\right]$$
$$+\theta\lambda^r(\lambda-\mathrm{i}t)^{-r}\prod_{j=1}^{p-3}\left(c+\frac{j}{2}\right)^{r_j}\left(c+\frac{j}{2}-\mathrm{i}t\right)^{-r_j}, \tag{50}$$

wich is the c.f. of a finite mixture of n^*+2 distributions, where n^*+1 components are GIG distributions and the last one is a GNIG distribution.

Again, as in Theorem 2, we have to consider two situations:

- $n^* < \frac{p}{2} - 2$, where (50) may be written as

$$\frac{1}{B\left(c,\frac{p_2}{2}\right)}\frac{1}{\Gamma\left(1-\frac{p_2}{2}\right)}\left[\sum_{k=1}^{n^*}\frac{\Gamma\left(1-\frac{p_2}{2}+k\right)}{k!(c+k)}\prod_{j=1}^{p-3}\left(c+\frac{j}{2}\right)^{r^*_{kj}}\left(c+\frac{j}{2}-\mathrm{i}t\right)^{-r^*_{kj}}\right.$$
$$\left.+\frac{\Gamma\left(1-\frac{p_2}{2}\right)}{c}c(c-\mathrm{i}t)^{-1}\prod_{j=1}^{p-3}\left(c+\frac{j}{2}\right)^{r_j}\left(c+\frac{j}{2}-\mathrm{i}t\right)^{-r_j}\right]+\theta\lambda^r(\lambda-\mathrm{i}t)^{-r}\prod_{j=1}^{p-3}\left(c+\frac{j}{2}\right)^{r_j}\left(c+\frac{j}{2}-\mathrm{i}t\right)^{-r_j}, \tag{51}$$

which is the c.f. of a finite mixture of n^*+2 distributions, where n^*+1 components are GIG distributions (n^* with depth $p-3$ and one with depth $p-2$) and the last one is a GNIG distribution of depth $p-2$, with shape and rate parameters mentioned in the body of Theorem 4;

- $n^* \geq \frac{p}{2} - 2$, where (50) may be written as

$$\frac{1}{B\left(c,\frac{p_2}{2}\right)}\frac{1}{\Gamma\left(1-\frac{p_2}{2}\right)}\left[\sum_{k=1}^{\frac{p}{2}-2}\frac{\Gamma\left(1-\frac{p_2}{2}+k\right)}{k!(c+k)}\prod_{j=1}^{p-3}\left(c+\frac{j}{2}\right)^{r^*_{kj}}\left(c+\frac{j}{2}-\mathrm{i}t\right)^{-r^*_{kj}}\right.$$
$$\left.+\sum_{k=0,k\geq\frac{p}{2}-1}^{n^*}\frac{\Gamma\left(1-\frac{p_2}{2}+k\right)}{k!(c+k)}(c+k)(c+k-\mathrm{i}t)^{-1}\prod_{j=1}^{p-3}\left(c+\frac{j}{2}\right)^{r_j}\left(c+\frac{j}{2}-\mathrm{i}t\right)^{-r_j}\right] \tag{52}$$
$$+\theta\lambda^r(\lambda-\mathrm{i}t)^{-r}\prod_{j=1}^{p-3}\left(c+\frac{j}{2}\right)^{r_j}\left(c+\frac{j}{2}-\mathrm{i}t\right)^{-r_j}$$

which is the c.f. of a finite mixture of n^*+2 distributions, where n^*+1 components are GIG distributions ($\frac{p}{2}-2$ with depth $p-3$ and $n^*+1-\left(\frac{p}{2}-2\right)$ with depth $p-2$) and the last one is a GNIG distribution of depth $p-2$, with shape and rate parameters mentioned in the body of Theorem 4. ∎

The near-exact c.f.'s obtained in such a way are asymptotic (as the ones presented in Theorem 2) for increasing values of n^*, in a sense that they converge for the exact c.f., $\varphi_W(t)$, when $n^* \to +\infty$.

5. Comparative numerical studies

In order to assess the behaviour of the near-exact distributions developed we use two proximity measures, Δ_1 and Δ_2, already used by Grilo and Coelho (2007). These two measures are directly derived from the inversion formulas respectively for the p.d.f. and the c.d.f., and their expressions are

$$\Delta_1 = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left|\varphi_W(t)-\varphi(t)\right|dt \tag{53}$$

and

$$\Delta_2 = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left|\frac{\varphi_W(t)-\varphi(t)}{t}\right|dt \tag{54}$$

where $\varphi_W(t)$ represents the exact c.f. of the r.v. W and $\varphi(t)$ the near-exact c.f. under study. The measure Δ_2 in (54) may be seen as directly related to the Berry-Esseen bound. The use of the measures Δ_1 and Δ_2

enables us to obtain an upper bound on the absolute value of the difference of the density or the cumulative function, respectively. More precisely,

$$\max_{w>0} \left| f_W(w) - f(w) \right| \le \Delta_1 \quad \text{and} \quad \max_{w>0} \left| F_W(w) - F(w) \right| \le \Delta_2 ,$$

where $f_W(w)$ and $F_W(w)$ are, respectively, the exact p.d.f. and c.d.f. of *W* evaluated at $w > 0$ and $f(w)$ and $F(w)$ are, respectively, the near-exact p.d.f. and c.d.f. of *W*.

Smaller values of the measures Δ_1 and Δ_2 correspond to better approximations; this way, these measures are an useful tool for evaluating and comparing the performance of the near-exact distributions proposed.

5.1 Near-exact distributions for the product of variables with Beta distribution

At this stage we use the two proximity measures to evaluate the proximity of the near-exact distributions M(n^*+1)GIG+GNIG, which equate the first 2 moments, to the exact distribution and also to compare them with the near-exact distributions GNIG and M2GNIG developed by Grilo and Coelho (2007), which equate respectively the first 3 and 4 exact moments. We also analyze the precision of these near-exact distributions in terms of quantiles.

In Table 1 we have the values of both proximity measures for two members of the family of near-exact distributions M(n^*+1)GIG+GNIG, developed in Theorem 2 in Section 3. For the truncations n^*+1 = 24 and n^*+1 = 49 we obtain respectively the distributions M24GIG+GNIG and M49GIG+GNIG. In those calculations we use the exact c.f. in (32) and the near-exact c.f. in (42) and we consider some different values for the parameters *p*, *b* and *c*. As we expected, the values of the measures confirm that the near-exact distributions M(n^*+1)GIG+GNIG are asymptotic for increasing values of n^* (number of terms), with the M49GIG+GNIG distribution beating the M24GIG+GNIG distribution, for any combination of the parameters *p*, *b* and *c*. The quality of the near-exact distributions M(n^*+1)GIG+GNIG improve when *p* or *b* increase. But, when the value of *c* increases, from 0.5 to 1.0 or from 0.5 to 2.5, keeping fixed the values of the parameters *p* and *b*, the values of the measures increase, i.e., we have a decrease in the quality of the near-exact distributions.

In Table 2 we have the proximity measures for the near-exact distributions GNIG and M2GNIG used by Grilo and Coelho (2007). Computations are done for the same combinations of values of the parameters *p*, *b* and *c* as in Table 1. In terms of the variations of the values of these parameters we have a similar behavior as the near-exact distributions M24GIG+GNIG and M49GIG+GNIG. However, when the value of *c* increases above a certain level, the quality of the near-exact distributions GNIG and M2GNIG improves, with this last one always showing lower values for the proximity measures (see Table 4), in agreement with the conclusions obtained in the comparative analysis of moments in Grilo and Coelho (2007).

From the comparison of the values of the measures in the Tables 1 and 2 we can conclude that, for any combination of the parameters involved, the near-exact distributions based on truncations always show lower values of the proximity measures. This way, we can say that the near-exact distributions based on truncations are more precise (i.e., closer to the exact distribution) than the near-exact distributions based on factorizations, although these last ones equated more moments. All the near-exact distributions show an asymptotic behavior for increasing number of Beta r.v.'s.

Table 1. Values of measures Δ_1 and Δ_2, obtained with *p* and *b* odd, for some examples of near-exact distributions, based on truncations of the exact c.f.: M24GIG+MGNIG and M49GIG+MGNIG.

			Near-exact M24GIG+GNIG		Near-exact M49GIG+GNIG	
p	*b*	*c*	Δ_1	Δ_2	Δ_1	Δ_2
3	15	0.5	1.836E-15	1.123E-15	6.187E-19	3.770E-19
3	15	1.0	1.868E-14	9.569E-15	6.641E-18	3.388E-18
5	15	0.5	5.430E-16	4.834E-16	1.820E-19	1.622E-19
15	15	0.5	1.806E-16	2.213E-16	6.043E-20	7.402E-20
25	15	0.5	1.423E-16	1.862E-16	4.762E-20	6.228E-20
5	25	0.5	1.545E-16	1.474E-17	4.150E-23	3.904E-23
5	25	2.5	2.086E-14	1.196E-14	6.870E-20	3.930E-20

Table 2. Values of measures Δ_1 and Δ_2, obtained with p and b odd, for some examples of near-exact distributions, based on factorization of the exact c.f.: GNIG and M2GNIG.

			Near-exact GNIG		Near-exact M2GNIG	
p	b	c	Δ_1	Δ_2	Δ_1	Δ_2
3	15	0.5	4.087E-10	3.251E-10	9.544E-13	6.232E-13
3	15	1.0	1.083E-09	6.917E-10	2.838E-12	1.519E-12
5	15	0.5	4.898E-11	4.854E-11	1.191E-13	9.927E-14
15	15	0.5	9.249E-13	1.153E-12	7.449E-16	7.964E-16
25	15	0.5	8.895E-14	1.168E-13	3.648E-17	4.121E-17
5	25	0.5	2.578E-12	2.667E-12	1.922E-15	1.677E-15
5	25	2.5	2.764E-11	1.594E-11	2.821E-14	1.415E-14

In Table 3 we have the quantiles computed for the same combinations of values of the parameters p, b and c as in Table 1 and 2. In these examples the weights in the mixture will vanish rather quickly so that a few terms in the mixture may be enough to obtain a very accurate result. In the last column of the table we have the number of terms needed to stabilize the first fifteen decimal places of quantiles of the near-exact distributions M(n^*+1)GIG+GNIG. Taking this fact into account, these near-exact quantiles may indeed be regarded as virtually exact, given the good convergence properties of the series involved. When we increase the value of c (from 0.5 to 1.0 or from 0.5 to 2.5) the number of terms needed to obtain the desired convergence of the series also increases.

Using the virtually exact quantiles of the M(n^*+1)GIG+GNIG, we can see that the near-exact quantiles of the GNIG and M2GNIG display a good quality of approximation (once again, the near-exact distribution based on a mixture, M2GNIG, is more precise).

Table 3. Quantiles (q_α) of some examples, obtained with p and b odd, for the near-exact distributions: GNIG, M2GNIG and M(n^*+1)GIG+GNIG.

p	b	c	q_α	Near-exact GNIG	Near-exact M2GNIG	Near-exact M(n^*+1)GIG+GNIG	No. of terms (n^*+1)
3	15	0.5	0.90	12.232137405284773	12.232137405198539	12.232137405199049	
			0.95	13.689451147392434	13.689451146907433	13.689451146907453	21
			0.99	16.964548406041502	16.964548405149140	16.964548405148528	
3	15	1.0	0.90	8.768751109405600	8.768751109981488	8.768751109984258	
			0.95	9.605166238056734	9.605166237391850	9.605166237393618	39
			0.99	11.392758529180804	11.392758526021177	11.392758526018796	
5	15	0.5	0.90	15.766703028422185	15.766703028417300	15.766703028417470	
			0.95	17.247066255829877	17.247066255666547	17.247066255666580	27
			0.99	20.542428502938455	20.542428502588221	20.542428502588027	
15	15	0.5	0.90	25.962744221484464	25.962744221485059	25.962744221485061	
			0.95	27.480762583918854	27.480762583914082	27.480762583914083	25
			0.99	30.811290143362668	30.811290143350415	30.811290143350412	
25	15	0.5	0.90	31.832735492207600	31.832735492207671	31.832735492207671	
			0.95	33.362684554131993	33.362684554131506	33.362684554131506	25
			0.99	36.704902646544632	36.704902646543340	36.704902646543339	
5	25	0.5	0.90	18.135351985604587	18.135351985604009	18.135351985604012	
			0.95	19.620476655328868	19.620476655322552	19.620476655322552	19
			0.99	22.920056799971283	22.920056799958249	22.920056799958245	
5	25	2.5	0.90	9.890322906372548	9.890322906404629	9.890322906373449	
			0.95	10.401843841474107	10.401843841474313	10.401843841471693	27
			0.99	11.424649925108542	11.424649924998074	11.424649924998035	

5.2 NEAR-EXACT DISTRIBUTIONS FOR THE WILKS LAMBDA STATISTIC

Here we consider variations on the number of variables per set and on the sample size and we compare the near-exact distributions M(n^*+1)GIG+GNIG, GNIG and M2GNIG.

The exact distribution of the Wilks Λ statistic used to test the independence of two sets of variables both with an odd number of variables (say p_1 and p_2) and also the near-exact distributions M(n^*+1)GIG+GNIG, were developed in Section 4, by direct application of the results obtained for the distribution of the product of independent number of r.v.'s Beta, in Section 3. This way, if we note that the parameters p, b and c of Section 3 are now, respectively, p_1, p_2 and $\frac{n+1-(p_1+p_2)}{2}$, we may consider the quantiles for different examples in Table 3, as quantiles for the Wilks Λ statistic, taking $p_1 = p,\ p_2 = b$ and $n = 2c + (p_1 + p_2) - 1$. In order to assess the quality of the near-exact distributions M(n^*+1)GIG+GNIG we use the exact distribution in Alberto (1998) and Alberto and Coelho (2007), in a manageable form, which makes possible the computation of quantiles for the particular case of $p_1 = 3$ and p_2 odd. Then, we verify in Table 3 that for $p_1 = 3$ the quantiles obtained for the family members of the near-exact distribution M(n^*+1)GIG+GNIG, which we designated as virtually exact quantiles, have the first fifteen decimal places equal to the exact quantiles.

In table 4 we may see the computed values of the two proximity measures for two other examples where the exact distribution is known, which correspond to situations where both sets have an odd number of variables (close to each other), one of them with three variables. We may confirm that, when the sample size increases (and, consequently, the value of c) the quality of approximation of the near-exact distribution M499GIG+GNIG becomes a bit worse (the values of both measures increase). However, we can surpass this minor drawback increasing the number of terms considered in the truncations. As expected, when we increase the number of variables in one of the sets, keeping constant the difference n - p, the performance of all near-exact distributions becomes even better.

We may see the better performance of the near-exact distributions based on truncations, with lower values for both proximity measures, when the values of n and p are close or even equal (Table 1 and 4).

Table 4. Values of measures Δ_1 and Δ_2, obtained with $p_1 = 3$ and p_2 odd, for some examples of near-exact distributions: M499GIG+MGNIG, GNIG and M2GNIG.

				Near-exact M499GIG+GNIG		Near-exact GNIG		Near-exact M2GNIG	
p_1	p_2	c	n	Δ_1	Δ_2	Δ_1	Δ_2	Δ_1	Δ_2
3	5	1	9	4.296E-17	8.843E-18	6.214E-07	3.102E-07	4.322E-09	1.633E-09
3	5	46	99	1.844E-08	2.275E-10	3.162E-08	8.039E-10	6.945E-11	1.556E-12
3	7	1	11	6.440E-20	2.277E-20	8.601E-08	4.745E-08	8.269E-10	3.754E-10
3	7	45	99	2.324E-10	5.014E-12	1.225E-08	3.947E-10	2.341E-11	6.650E-13

In Table 5 we have the quantiles for the same examples studied in Table 4. Once again, we use the exact quantiles obtained with the exact distribution, in Alberto (1998) and Alberto and Coelho (2007), to compare with our virtually exact quantiles obtained with the near-exact distributions M(n^*+1)GIG+GNIG. For the cases where p_2 = 5 or 7 and $n - p$ = 1, we used as a stop criterion the number of terms needed to equate the first fifteen decimal places, which we obtain with 500 terms. However, with one hundred terms we already equated at least the first ten decimal places. For the cases where $n - p$ = 91 and $n - p$ = 89, the virtually exact quantiles equate at most the first ten decimal places of the exact quantiles, considering the same number of terms.

Based on different family members of the near-exact distribution M(n^*+1)GIG+GNIG evidence is given that, when we have p_1 and p_2 small and close to each other or when the difference between the total number of variables and the sample size, $n - p$, is relatively high, we have to consider more terms in the series involved in order to increase the quality of the quantiles (see Table 3 and 5). The results obtained for the proximity measures, in Tables 1 and 4, corroborate these conclusions.

Table 5. Quantiles (q_α) of some examples, obtained with p_1 = 3 and p_2 odd, for near-exact distributions: M499GIG+GNIG, GNIG and M2GNIG.

p_1	p_2	c	n	q_α	Near-exact M499GIG+GNIG	Near-exact GNIG	Near-exact M2GQIG
3	5	1	9	0.90	5.902424173900701	5.902424705542200	5.902424172750858
				0.95	6.708991141654191	6.708991203754051	6.708991143494366
				0.99	8.458264467172885	8.458263267281493	8.458264470394018
3	5	46	99	0.90	0.236128176914884	0.236128176955616	0.236128176816062
				0.95	0.264594184788194	0.264594184882131	0.264594184679287
				0.99	0.323698431665835	0.323698431554972	0.323698431562073
3	7	1	11	0.90	6.737900958585013	6.737901023198875	6.737900957949009
				0.95	7.556390637123642	7.556390623356051	7.556390636602448
				0.99	9.320950830762371	9.320950634927288	9.320950831144527
3	7	45	99	0.90	0.316922153184000	0.316922153194397	0.316922153096966
				0.95	0.349631778898221	0.349631778491605	0.349631778378447
				0.99	0.416670444515805	0.416670444275518	0.416670444318288

Since for $p_1 \geq 3$ and p_2 odd the exact distribution is not available in a manageable form, adequate for the expedite computation of quantiles, the near-exact distributions presented are then an excellent option. The quantiles obtained with the near-exact distributions GNIG and M2GNIG show a very regular behavior for different variations in the number of variables per set and on the sample size, where we notice the quality of approximation of the near-exact distribution M2GNIG. But, we have to point out the excellent performance of the family members of the near-exact distribution M(n^*+1)GIG+GNIG which allows us to obtain virtually exact quantiles. They are particularly adequate and useful for cases where $n - p$ is small.

6. CONCLUSIONS AND FINAL REMARKS

Once the exact distribution of the Wilks Lambda statistic is expressed under the form of an infinite mixture of GIG distributions, where the associated series has good convergence properties, we are able to obtain the family of near-exact distributions M(n^*+1)GIG+GNIG, based on truncations of the exact c.f., which is relatively easy to implement computationally and allows for the computation of virtually exact quantiles. These near-exact distributions are obtained in a manageable form and by construction the first two moments are equal to the exact ones.

Based on cases where the exact distribution is known, some evidence is given that the measures Δ_1 and Δ_2 are accurate to evaluate the proximity of quantiles (smaller values of the proposed measures are associated with smaller differences among quantiles).

The family members of the near-exact distribution M(n^*+1)GIG+GNIG proposed in this paper displaying an asymptotic behavior for increasing number of variables. We also have to stress the outstanding performance of these near-exact distributions for small values of n, or should we rather say, for small values of $n - p$, what makes these approximations particularly useful in situations of small sample sizes, situations in which even the best asymptotic distributions available perform not too well. For higher values of $n - p$, we may consider the near-exact distributions GNIG and M2GNIG as an alternative to the near-exact distributions M(n^*+1)GIG+GNIG. These approximations lay very close to the exact distribution in terms of c.f.'s, p.d.f.'s, c.d.f.'s, moments and quantiles.

ACKNOWLEDGMENTS

The authors wish to thank the financial support of the Portuguese Foundation for Science and Technology (FCT).

REFERENCES

Abramowitz, M. and Stegun, I. A. (1974). *Handbook of Mathematical Functions.* (eds.). 9.ª ed., Dover, New York.

Alberto, R. P. (1998). *The exact distribution of the Wilks Lambda statistic* (in Portuguese). Mathematics M.Sc. Thesis, Lusiada University, Lisboa, Portugal.

Alberto, R. P. and Coelho, C. A. (2007). Study of the quality of several asymptotic and near-exact approximations based on moments for the distribution of the Wilks Lambda statistic. *Journal of Statistical Planning and Inference*, **137**, 1612-1626.

Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis.* 2.ª ed., J. Wiley & Sons, New York.

Coelho, C. A. (1998). The generalized integer Gamma distribution – a basis for distributions in Multivariate Statistics. *Journal of Multivariate Analysis*, **64**, 86-102.

Coelho, C. A. (1999). Addendum to the paper "The generalized integer Gamma distribution – a basis for distributions in Multivariate Statistics". *Journal of Multivariate Analysis*, **69**, 281-285.

Coelho, C. A. (2003). A generalized Integer Gamma distribution as an asymptotic replacement for a Logbeta distribution & applications. *American Journal of Mathematical Management Sciences*, **23**, 383-399.

Coelho, C. A. (2004). The generalized near-integer Gamma distribution: a basis for 'near-exact' approximations to the distribution of statistics which are the product of an odd number of independent Beta random variables. *Journal of Multivariate Analysis*, **89**, 191-218.

Coelho, C. A., Alberto, R. P. and Grilo, L. M. (2006). A mixture of Generalized Integer Gamma distribution as the exact distribution of the product of an odd number of independent Beta random variables with first parameter evolving by 1/2. Applications. *Journal of Interdisciplinary Mathematics*, **9**, 2, 228-248.

Grilo, L. M. (2005). *Development of near-exact distributions for different scenarios of application of the Wilks Lambda statistic* (in Portuguese). Ph.D. thesis, Thecnical University of Lisbon, Portugal.

Grilo, L. M. and Coelho, C. A. (2007). Development and Comparative Study of two Near-exact Approximations to the Distribution of the Product of an Odd Number of Independent Beta Random Variables. *Journal of Statistical Planning and Inference*, **137**, 1560-1575.

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distributions*, vol.2, J. Wiley & Sons, New York.

Wilks, S. S. (1932). Certain generalizations in the analysis of variance. *Biometrika*, **24**, 471-494.

Wilks, S. S. (1935). On the independence of *k* sets of normally distributed statistical variables. *Econometrica*, **3**, 309-326.