Near-exact distributions for the independence and sphericity likelihood ratio test statistics

Carlos A. Coelho * Filipe J. Marques

cmac@fct.unl.pt *fjm@fct.unl.pt*

Mathematics Department, Faculty of Sciences and Technology
The New University of Lisbon, Portugal

Abstract

In this paper we will show how a suitably developed decomposition of the characteristic function of the logarithm of the likelihood ratio test statistic to test independence in a set of variates may be used to obtain extremely well-fit near-exact distributions both for this test statistic as well as for the likelihood ratio test statistic for sphericity, based on a decomposition of this latter test in two independent tests. For the independence test statistic, numerical studies and comparisons with asymptotic distributions proposed by other authors show the extremely high closeness of the near-exact distributions developed to the exact distribution. Concerning the sphericity test statistic, comparisons with the near-exact distributions developed in [19] show the advantages of these new near-exact distributions.

Key words: Wilks Lambda statistic, independence test, sphericity test, Generalized Near-Integer Gamma distribution, mixtures.

1 Introduction

Let $\underline{X}$ be a $p \times 1$ vector with a p-multivariate Normal distribution with expected value $\underline{\mu}$ and variance-covariance matrix Σ, that is, let

$$\underline{X} \sim N_p(\underline{\mu}, \Sigma)\,. \tag{1}$$

* Corresponding author. Address: The New University of Lisbon, Faculty of Sciences and Technology, Mathematics Department, Quinta da Torre, 2829-516 Caparica, Portugal; Tel:(351) 21 294 8388; Fax:(351) 21 294 8391.

Preprint submitted to Elsevier Science *9 November 2007*

Then, the power $2/(n+1)$ of the likelihood ratio test statistic to test the null hypothesis

$$H_{01}: \Sigma = diag(\sigma_1^2, \sigma_2^2, \ldots, \sigma_p^2) \tag{2}$$

based on a sample of size $n+1$, is the statistic

$$\Lambda_1 = \frac{|V|}{\prod_{j=1}^{p} V_j} \tag{3}$$

where the $p{\times}p$ matrix V is either the MLE (Maximum Likelihood Estimator) of Σ, the sample matrix of sum of squares and products of deviations from the sample mean or the sample variance-covariance matrix of the p variables in $\underline{X}$ and V_j is the j-th diagonal element of V. The statistic in (3) is a particular case of the generalized Wilks Λ statistic used to test the independence of p groups of variables, each with one only variable (see [7],[8],Ch. 9 in [3], Ch. 10 in [17], Ch. 11 in [21]). Near-exact distributions for this statistic are thus readily available from the results in [9,10,2,15]. However, given the specificity of the case under consideration, some further developments may be sought, namely a simpler way to obtain and a simpler formulation for the shape parameters of the Gamma distributions involved in the part of the distribution of Λ_1 left untouched. These details will be addressed in Section 2.

On the other hand, the power $2/(n+1)$ of the likelihood ratio test statistic to test the sphericity hypothesis on Σ, based on a sample of size $n+1$, that is, to test the null hypothesis

$$H_0: \Sigma = \sigma^2 I_p \qquad (\sigma^2 \text{ unspecified}) \tag{4}$$

is the statistic (see Ch. 10 in [3], Ch. 10 in [17], Ch. 8 in [21])

$$\Lambda = p^p \frac{|V|}{(trV)^p}, \tag{5}$$

where V is the matrix in (3).

Well-fit near-exact distributions have alread been developed for this statistic by Marques and Coelho in [19]. However, and somehow unexpectedly, in this paper we will show that even better near-exact distributions may be obtained for this statistic by taking as a basis the near-exact distributions developed for the statistic Λ_1 in (3) and the decomposition performed on its characteristic function. These near-exact distributions for Λ will be obtained from a decomposition of the statistic Λ in (5), which may be written as

$$\Lambda = \Lambda_1 \, \Lambda_2 \,, \tag{6}$$

where Λ_1 is the statistic in (3) and

$$\Lambda_2 = p^p \, \frac{\prod_{j=1}^{p} V_j}{(trV)^p} \,, \tag{7}$$

is the power $2/(n{+}1)$ of the likelihood ratio test statistic to test the hypothesis

$$H_{02|01} : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_p^2 \qquad \text{(given that, or, assuming that the } p \text{ variables in } \underline{X} \text{ are independent)} \tag{8}$$

based on p independent estimates of the variances of the variables in $\underline{X}$, one for each σ_j^2, based on samples of size $n+1$. In (7), V and V_j are the same as in (3). We may note that V_j is either the MLE of σ_j^2 ($j = 1, \ldots, p$), the sample variance of the j-th variable in $\underline{X}$ in (1), or the sum of squares of the deviations from the sample mean for the j-th variable in $\underline{X}$, according to the choice of V.

The statistic in (7) may be derived from the likelihood ratio test statistic for the equality of p variance-covariance matrices (see Ch. 10 in [3], Ch. 10 in [17], Ch. 8 in [21]), taking each matrix to have dimensions $1{\times}1$ (there they are again, the p groups of one variable each).

We may write for H_0 in (4), H_{01} in (2) and H_{02} in (8),

$$H_0 = H_{02|01} \circ H_{01} \,, \tag{9}$$

to be read as "$H_{02|01}$ after H_{01}", meaning that we may test H_0 in two steps: (i) testing first H_{01}, that is, if the p variables in $\underline{X}$ are independent and (ii) once the hypothesis of independence of the p variables is not rejected, testing then if they all have the same variance. Under H_0 in (4) the two test statistics Λ_1 and Λ_2 in (6) are independent (see Ch. 10, subsec. 10.7.3 in [3]). This way to look at this test will enable us to obtain even better near-exact distributions than the already much well-fit ones in [19].

As a side note we may stress that the test statistic in (7) may be used, under a slightly different setting, to test the null hypothesis of equality of variances in (8), without any conditioning if the p estimators V_j are based on p independent samples, in which case those samples may have different sizes.

Indeed in the test statistic in (7) the requirement is that the p estimators V_j have to be independent (among other reasons, because one has to be able to easily derive the distribution of trV, their sum). We may note that this will be the case even if the p estimators V_j come from a multivariate sample of size $n+1$ of the p variables in $\underline{X}$, once the null hypothesis of independence of the p variables is not rejected, since then the matrix V will have a Wishart distribution with n degrees of freedom and parameter matrix the matrix Σ in (2), so that the diagonal elements of V are independent (through a simple extension of Theorem 3.2.7 in [21]).

2 Near-exact distributions for the likelihood ratio test statistic of independence

In order to obtain the c.f. (characteristic function) of $W_1 = -\log \Lambda_1$ we may consider Theorems 9.3.2 and 9.3.3 of [3] which state that, for a sample of size $n+1$,

$$\Lambda_1 \sim \prod_{j=1}^{p-1} Y_j \,, \tag{10}$$

(where '$\sim$' is to be read as 'is distributed as') with

$$Y_j \sim B\left(\frac{n-p+j}{2}, \frac{p-j}{2}\right) \qquad (j=1,\ldots,p) \tag{11}$$

where Λ_1 is the statistic in (3), and where, under H_{01} in (2), the $p-1$ random variables Y_j in (10) and (11) are independent. Then, since we known that

$$E\left(Y_j^h\right) = \frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n-p+j}{2}+h\right)}{\Gamma\left(\frac{n}{2}+h\right)\Gamma\left(\frac{n-p+j}{2}\right)}\,, \qquad \left(h > -\frac{n-p+j}{2}\right)$$

we have, for $i=(-1)^{1/2}$,

$$\Phi_{W_1}(t) = E\left(e^{itW_1}\right) = \prod_{j=1}^{p-1} E\left(Y_j^{-it}\right) = \prod_{j=1}^{p-1} \frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n-p+j}{2}-it\right)}{\Gamma\left(\frac{n}{2}-it\right)\Gamma\left(\frac{n-p+j}{2}\right)}\,. \tag{12}$$

Now, in order to be able to obtain a suitable decomposition of the c.f. of W_1 we may either consider the results and developments in section 5 of [10], taking $p_k=1$ for $k=1,\ldots,m$ and $k^*=\lfloor p/2 \rfloor$, or we may take a different approach which will indeed enable us to obtain simpler expressions for the shape parameters of the Gamma distributions involved in the part of the distribution of W_1 which will be left unchanged. We will take this second approach.

The following Lemma gives the c.f. of W_1, for both even and odd p, under a form that is suitable for the development of near-exact distributions for both W_1 and Λ_1.

Lemma 1 *Under H_0 in (2), taking $k^* = \lfloor p/2 \rfloor$ and "$a \perp\!\!\!\perp b$", with $a, b, \in I\!N$, as representing the remainder of the integer ratio of a by b, the c.f. of $W_1 = -\log \Lambda_1$ (where Λ_1 is the statistic in (3), used to test the independence among the p variables in $\underline{X}$), may be written under the form*

$$\Phi_{W_1}(t) = \underbrace{\left(\frac{\Gamma\left(\frac{n}{2}\right)\,\Gamma\left(\frac{n}{2}-\frac{1}{2}-it\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\,\Gamma\left(\frac{n}{2}-it\right)}\right)^{k^*}}_{\Phi_1^*(t)} \tag{13}$$
$$\times \underbrace{\prod_{k=1}^{p-2}\left(\frac{n-1-k}{2}\right)^{k^*-\left\lfloor\frac{k+1-(p\perp\!\!\!\perp 2)}{2}\right\rfloor}\left(\frac{n-1-k}{2}-it\right)^{-k^*+\left\lfloor\frac{k+1-(p\perp\!\!\!\perp 2)}{2}\right\rfloor}}_{\Phi_2^*(t)}\,,$$

where $\Phi_1^*(t)$ is the c.f. of the sum of k^* independent Logbeta r.v.'s with parameters $(n-1)/2$ and $1/2$ and $\Phi_2^*(t)$ is the c.f. of a GIG (Generalized Integer Gamma) distribution [8] of depth $p-2$ with rate parameters $(n-1-k)/2$ and shape parameters $k^* - \left\lfloor\frac{k+1-(p\perp\!\!\!\perp 2)}{2}\right\rfloor$ $(k = 1, \ldots, p-2)$, that is the distribution of the sum of $p-2$ independent Gamma r.v.'s with the given rate and integer shape parameters.

Proof: We will need to consider separately the two cases of even and odd p. For even p, from (12) we may write

$$\begin{aligned}\Phi_{W_1}(t) &= \underbrace{\prod_{\substack{j=1\\ \text{step } 2}}^{p-1}\frac{\Gamma\left(\frac{n}{2}\right)\,\Gamma\left(\frac{n-p+j}{2}-it\right)}{\Gamma\left(\frac{n}{2}-it\right)\,\Gamma\left(\frac{n-p+j}{2}\right)}}_{\Phi_1(t)}\;\underbrace{\prod_{\substack{j=2\\ \text{step } 2}}^{p-2}\frac{\Gamma\left(\frac{n}{2}\right)\,\Gamma\left(\frac{n-p+j}{2}-it\right)}{\Gamma\left(\frac{n}{2}-it\right)\,\Gamma\left(\frac{n-p+j}{2}\right)}}_{\Phi_2(t)}\\ &= \Phi_1(t)\;\Phi_2(t)\,,\end{aligned}$$

where, for even j, $(p-j)/2 \in I\!N$, so that, using for $z \in \mathbb{C}\backslash\{0,-1,-2,\ldots\}$ and $n \in I\!N$,

$$\frac{\Gamma(z+n)}{\Gamma(z)} = \prod_{k=0}^{n-1}(z+k)\,,$$

we may write

$$\begin{aligned}
\Phi_2(t) &= \prod_{\substack{j=2\\ \text{step } 2}}^{p-2} \prod_{k=0}^{\frac{p-j}{2}-1} \left(\frac{n-p+j}{2}+k\right)\left(\frac{n-p+j}{2}+k-it\right)^{-1} \\
&= \prod_{\substack{j=2\\ \text{step } 2}}^{p-2} \prod_{k=0}^{\frac{p-j}{2}-1} \left(\frac{n-p+j}{2}+\frac{p-j}{2}-1-k\right)\left(\frac{n-p+j}{2}+\frac{p-j}{2}-1-k-it\right)^{-1} \\
&= \prod_{\substack{j=2\\ \text{step } 2}}^{p-2} \prod_{k=1}^{\frac{p-j}{2}} \left(\frac{n}{2}-k\right)\left(\frac{n}{2}-k-it\right)^{-1} = \prod_{k=1}^{\frac{p-2}{2}} \left(\frac{n}{2}-k\right)^{\frac{p}{2}-k}\left(\frac{n}{2}-k-it\right)^{-\left(\frac{p}{2}-k\right)}
\end{aligned}$$

and

$$\begin{aligned}
\Phi_1(t) &= \prod_{\substack{j=1\\ \text{step } 2}}^{p-1} \frac{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\Gamma\left(\frac{n-p+j}{2}-it\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}-it\right)\Gamma\left(\frac{n-p+j}{2}\right)} \frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}-\frac{1}{2}-it\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\Gamma\left(\frac{n}{2}-it\right)} \\
&= \left(\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}-\frac{1}{2}-it\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\Gamma\left(\frac{n}{2}-it\right)}\right)^{p/2} \prod_{\substack{j=1\\ \text{step } 2}}^{p-1} \prod_{k=0}^{\frac{p-j-1}{2}-1} \left(\frac{n-p+j}{2}+k\right)\left(\frac{n-p+j}{2}+k-it\right)^{-1} \\
&= \left(\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}-\frac{1}{2}-it\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\Gamma\left(\frac{n}{2}-it\right)}\right)^{p/2} \prod_{\substack{j=1\\ \text{step } 2}}^{p-1} \prod_{k=0}^{\frac{p-j-1}{2}-1} \left(\frac{n-p+j}{2}+\frac{p-j-1}{2}-1-k\right) \\
&\qquad\qquad \left(\frac{n-p+j}{2}+\frac{p-j-1}{2}-1-k-it\right)^{-1} \qquad (14) \\
&= \left(\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}-\frac{1}{2}-it\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\Gamma\left(\frac{n}{2}-it\right)}\right)^{p/2} \prod_{k=1}^{\frac{p-2}{2}} \left(\frac{n-1}{2}-k\right)^{\frac{p}{2}-k}\left(\frac{n-1}{2}-k-it\right)^{-\left(\frac{p}{2}-k\right)},
\end{aligned}$$

so that we may finally write, for even p,

$$\Phi_{W_1}(t) = \left(\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}-\frac{1}{2}-it\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\Gamma\left(\frac{n}{2}-it\right)}\right)^{p/2} \prod_{k=1}^{p-2} \left(\frac{n-1-k}{2}\right)^{\frac{p}{2}-\lfloor\frac{k+1}{2}\rfloor}\left(\frac{n-1-k}{2}-it\right)^{-\frac{p}{2}+\lfloor\frac{k+1}{2}\rfloor}.$$

For odd p, we may write

$$\Phi_{W_1}(t) = \underbrace{\prod_{\substack{j=1\\ \text{step } 2}}^{p-2} \frac{\Gamma\left(\frac{n}{2}\right)\,\Gamma\left(\frac{n-p+j}{2}-it\right)}{\Gamma\left(\frac{n}{2}-it\right)\,\Gamma\left(\frac{n-p+j}{2}\right)}}_{\Phi_1(t)} \underbrace{\prod_{\substack{j=2\\ \text{step } 2}}^{p-1} \frac{\Gamma\left(\frac{n}{2}\right)\,\Gamma\left(\frac{n-p+j}{2}-it\right)}{\Gamma\left(\frac{n}{2}-it\right)\,\Gamma\left(\frac{n-p+j}{2}\right)}}_{\Phi_2(t)}$$

$$= \Phi_1(t)\,\Phi_2(t)\,,$$

where now it is for odd j that $(p-j)/2 \in I\!N$, so that following similar steps to the ones used above to handle $\Phi_2(t)$, we may write

$$\Phi_1(t) = \prod_{\substack{j=1\\ \text{step } 2}}^{p-2} \prod_{k=0}^{\frac{p-j}{2}-1} \left(\frac{n-p+j}{2}+k\right)\left(\frac{n-p+j}{2}+k-it\right)^{-1}$$

$$= \prod_{\substack{j=1\\ \text{step } 2}}^{p-2} \prod_{k=1}^{\frac{p-j}{2}} \left(\frac{n}{2}-k\right)\left(\frac{n}{2}-k-it\right)^{-1}$$

$$= \prod_{k=1}^{\frac{p-1}{2}} \left(\frac{n}{2}-k\right)^{\frac{p-1}{2}-k}\left(\frac{n}{2}-k-it\right)^{-\left(\frac{p-1}{2}-k\right)}$$

and

$$\Phi_2(t) = \prod_{\substack{j=2\\ \text{step } 2}}^{p-1} \frac{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\,\Gamma\left(\frac{n-p+j}{2}-it\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}-it\right)\,\Gamma\left(\frac{n-p+j}{2}\right)} \frac{\Gamma\left(\frac{n}{2}\right)\,\Gamma\left(\frac{n}{2}-\frac{1}{2}-it\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\,\Gamma\left(\frac{n}{2}-it\right)}$$

$$= \left(\frac{\Gamma\left(\frac{n}{2}\right)\,\Gamma\left(\frac{n}{2}-\frac{1}{2}-it\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\,\Gamma\left(\frac{n}{2}-it\right)}\right)^{(p-1)/2} \prod_{\substack{j=2\\ \text{step } 2}}^{p-1} \prod_{k=0}^{\frac{p-j-1}{2}-1} \left(\frac{n-p+j}{2}+k\right)\left(\frac{n-p+j}{2}+k-it\right)^{-1} \quad (15)$$

$$= \left(\frac{\Gamma\left(\frac{n}{2}\right)\,\Gamma\left(\frac{n}{2}-\frac{1}{2}-it\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\,\Gamma\left(\frac{n}{2}-it\right)}\right)^{(p-1)/2} \prod_{k=1}^{\frac{p-3}{2}} \left(\frac{n-1}{2}-k\right)^{\frac{p-1}{2}-k}\left(\frac{n-1}{2}-k-it\right)^{-\left(\frac{p-1}{2}-k\right)},$$

so that we may finally write, for odd p,

$$\Phi_{W_1}(t) = \left(\frac{\Gamma\left(\frac{n}{2}\right)\,\Gamma\left(\frac{n}{2}-\frac{1}{2}-it\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\,\Gamma\left(\frac{n}{2}-it\right)}\right)^{(p-1)/2}$$
$$\prod_{k=1}^{p-2}\left(\frac{n-1-k}{2}\right)^{\frac{p-1}{2}-\lfloor\frac{k}{2}\rfloor}\left(\frac{n-1-k}{2}-t\right)^{-\frac{p-1}{2}+\lfloor\frac{k}{2}\rfloor},$$

and so that, taking $k^* = \lfloor p/2 \rfloor$, we may write $\Phi_W(t)$ under the form in (13), for any even or odd p. □

Then, taking into account that since a single Logbeta distribution may be represented under the form of an infinite mixture of either Exponential or GIG distributions ([12]), a sum of independent Logbeta r.v.'s with either the same or different parameters may thus be represented under the form of an infinite mixture of sums of independent Exponentials or GIG distributions, which are all GIG distributions and that the GIG distribution itself may be seen as a mixture of Gamma distributions ([11]), the replacement of the sum of independent Logbeta r.v.'s by a single Gamma distribution or by a (finite) mixture of Gamma distributions seems to be most adequate.

Thus, near-exact distributions for W_1 may then be obtained under the form of a (Generalized Near-Integer Gamma) distribution ([10]) or mixtures of GNIG distributions by replacing $\Phi_1^*(t)$ by the c.f. of a Gamma distribution or the c.f. of a mixture of Gamma distributions (see Appendix A for details on the GNIG distribution). These near-exact distributions will match, by construction, the first two, four and six exact moments of W_1.

Theorem 2 *Using for $\Phi_1^*(t)$ in (13), the approximations:*

- $\lambda^s(\lambda-\mathrm{i}t)^{-s}$ *with $s,\lambda>0$, such that*

$$\frac{\partial^h}{\partial t^h}\,\lambda^s(\lambda-\mathrm{i}t)^{-s}\Bigg|_{t=0} = \frac{\partial^h}{\partial t^h}\,\Phi_1^*(t)\Bigg|_{t=0} \qquad \textit{for} \quad h=1,2\,; \tag{16}$$

- $\sum_{k=1}^{2}\theta_k\,\mu^{s_k}(\mu-\mathrm{i}t)^{-s_k}$, *where $\theta_2=1-\theta_1$ with $\theta_k,\ s_k,\ \mu>0$, such that*

$$\frac{\partial^h}{\partial t^h}\sum_{k=1}^{2}\theta_k\,\mu^{s_k}(\mu-\mathrm{i}t)^{-s_k}\Bigg|_{t=0} = \frac{\partial^h}{\partial t^h}\,\Phi_1^*(t)\Bigg|_{t=0} \qquad \textit{for} \quad h=1,\ldots,4\,; \tag{17}$$

- $\sum_{k=1}^{3} \theta_k^* \, \nu^{s_k^*} (\nu - \mathrm{i}t)^{-s_k^*}$, *where* $\theta_3^* = 1 - \theta_1^* - \theta_2^*$ *with* θ_k^*, s_k^*, $\nu > 0$, *such that*

$$\frac{\partial^h}{\partial t^h} \sum_{k=1}^{3} \theta_k^* \, \nu^{s_k^*} (\nu - \mathrm{i}t)^{-s_k^*} \bigg|_{t=0} = \frac{\partial^h}{\partial t^h} \, \Phi_1^*(t) \bigg|_{t=0} \qquad \text{for} \quad h = 1, \ldots, 6 \, ; \quad (18)$$

we obtain as near-exact distributions for W_1*, respectively,*

i) a GNIG distribution of depth $p-1$ *with cdf (cumulative distribution function) (using the notation in* (42) *in Appendix A)*

$$F(w | r_1, \ldots, r_{p-2}, s; \lambda_1, \ldots, \lambda_{p-2}, \lambda) \, , \qquad (19)$$

where

$$r_j = \left\lfloor \frac{p}{2} \right\rfloor - \left\lfloor \frac{j+1-(p \perp\!\!\!\perp 2)}{2} \right\rfloor, \quad \lambda_j = \frac{n-1-j}{2} \, , \quad (j = 1, \ldots, p-2), \quad (20)$$

and

$$\lambda = \frac{m_1}{m_2 - m_1^2} \qquad \text{and} \qquad s = \frac{m_1^2}{m_2 - m_1^2} \qquad (21)$$

with

$$m_h = \mathrm{i}^{-h} \, \frac{\partial^h}{\partial t^h} \, \Phi_1^*(t) \bigg|_{t=0} \, , \qquad h = 1, 2 \, ;$$

ii) a mixture of two GNIG distributions of depth $p-1$*, with cdf (using the notation in* (42) *in Appendix A)*

$$\sum_{k=1}^{2} \theta_k \, F(w | r_1, \ldots, r_{p-2}, s_k; \lambda_1, \ldots, \lambda_{p-2}, \mu) \, , \qquad (22)$$

where r_j *and* λ_j *(*$j = 1, \ldots, p-2$*) are given by* (20) *above and* θ_1*,* μ*,* r_1 *and* r_2 *are obtained from the numerical solution of the system of four equations*

$$\sum_{k=1}^{2} \theta_k \, \frac{\Gamma(r_k + h)}{\Gamma(r_k)} \, \mu^{-h} = \mathrm{i}^{-h} \, \frac{\partial^h}{\partial t^h} \, \Phi_1^*(t) \bigg|_{t=0} \qquad (h = 1, \ldots, 4) \qquad (23)$$

for these parameters, with $\theta_2 = 1 - \theta_1$*;*

iii) or a mixture of three GNIG distributions of depth $p-1$*, with cdf (using the notation in* (42) *in Appendix A)*

$$\sum_{k=1}^{3} \theta_k^* \, F(w | r_1, \ldots, r_{p-2}, s_k^*; \lambda_1, \ldots, \lambda_{p-2}, \nu) \, , \qquad (24)$$

with r_j and λ_j $(j = 1, \dots, p-2)$ given by (20) above and θ_1^*, θ_2^*, ν, s_1^*, s_2^* and s_3^* obtained from the numerical solution of the system of six equations

$$\sum_{k=1}^{3} \theta_j^* \, \frac{\Gamma(r_k^* + h)}{\Gamma(r_k^*)} \, \nu^{-h} \; = \; \mathrm{i}^{-h} \, \frac{\partial^h}{\partial t^h} \, \Phi_1^*(t) \bigg|_{t=0} \qquad (h = 1, \dots, 6) \tag{25}$$

for these parameters, with $\theta_3^* = 1 - \theta_1^* - \theta_2^*$.

Proof: If in the characteristic function of W_1 in (13) we replace $\Phi_1^*(t)$ by $\lambda^s(\lambda - \mathrm{i}t)^{-s}$ we obtain

$$\Phi_{W_1}(t) \;\approx\; \lambda^s(\lambda - \mathrm{i}t)^{-s}$$

$$\times \underbrace{\prod_{k=1}^{p-2} \left(\frac{n-1-k}{2}\right)^{k^* - \left\lfloor \frac{k+1-(p \perp\!\!\!\perp 2)}{2} \right\rfloor} \left(\frac{n-1-k}{2} - it\right)^{-k^* + \left\lfloor \frac{k+1-(p \perp\!\!\!\perp 2)}{2} \right\rfloor}}_{\Phi_2^*(t)} ,$$

that is the characteristic function of the sum of $p-1$ independent Gamma random variables, $p-2$ of which with integer shape parameters r_j and rate parameters λ_j given by (20), and a further Gamma random variable with rate parameter $s > 0$ and shape parameter λ. This characteristic function is thus the c.f. of the GNIG distribution of depth $p-1$ with distribution function given in (19). The parameters s and λ are determined in such a way that (16) holds. This compels s and λ to be given by (21) and makes the two first moments of this near-exact distribution for W_1 to be the same as the two first exact moments of W_1.

If in the characteristic function of W_1 in (13) we replace $\Phi_1^*(t)$ by $\sum\limits_{k=1}^{2} \theta_k \, \mu^{r_k}(\mu - \mathrm{i}t)^{-r_k}$ we obtain

$$\Phi_{W_1}(t) \approx \sum_{k=1}^{2} \theta_k \, \mu^{r_k}(\mu - \mathrm{i}t)^{-r_k}$$

$$\times \underbrace{\prod_{k=1}^{p-2} \left(\frac{n-1-k}{2}\right)^{k^* - \left\lfloor \frac{k+1-(p \perp\!\!\!\perp 2)}{2} \right\rfloor} \left(\frac{n-1-k}{2} - it\right)^{-k^* + \left\lfloor \frac{k+1-(p \perp\!\!\!\perp 2)}{2} \right\rfloor}}_{\Phi_2^*(t)} ,$$

that is the characteristic function of the mixture of two GNIG distributions of depth $p-1$ with density function given in (22). The parameters θ_1, μ, r_1 and r_2 are defined in such a way that (17) holds, giving rise to the evaluation of these parameters as the numerical solution of the system of equations in (23) and to a near-exact distribution that matches the first four exact moments of W_1.

If in the characteristic function of W_1 in (13) we replace $\Phi_1^*(t)$ by $\sum_{k=1}^{3} \theta_k^* \, \nu^{r_k^*} (\nu - \mathrm{i}t)^{-r_k^*}$ we obtain

$$\Phi_{W_1}(t) \approx \sum_{k=1}^{3} \theta_k^* \, \nu^{r_k^*} (\nu - \mathrm{i}t)^{-r_k^*} \underbrace{\times \prod_{k=1}^{p-2} \left(\frac{n-1-k}{2}\right)^{k^* - \left\lfloor \frac{k+1-(p \perp\!\!\!\perp 2)}{2} \right\rfloor} \left(\frac{n-1-k}{2} - it\right)^{-k^* + \left\lfloor \frac{k+1-(p \perp\!\!\!\perp 2)}{2} \right\rfloor}}_{\Phi_2^*(t)},$$

that is the characteristic function of the mixture of three GNIG distributions of depth $p-1$ with density function given in (22). The parameters θ_1^*, θ_2^*, ν, r_1^*, r_2^* and r_3^* are defined in such a way that (18) holds, what gives rise to the evaluation of these parameters as the numerical solution of the system of equations in 25, giving rise to a near-exact distribution that matches the first six exact moments of W_1. $\square$

Corollary 3 *Distributions with cdf's given by*

i) $1 - F(-\log z | r_1, \ldots, r_{p-2}, s; \lambda_1, \ldots, \lambda_{p-2}, \lambda)\,,$

ii) $1 - \sum_{k=1}^{2} \theta_k \, F(-\log z | r_1, \ldots, r_{p-2}, s_k; \lambda_1, \ldots, \lambda_{p-2}, \mu)\,,$ *or*

iii) $1 - \sum_{k=1}^{3} \theta_k^* \, F(-\log z | r_1, \ldots, r_{p-2}, s_k^*; \lambda_1, \ldots, \lambda_{p-2}, \nu)\,,$

where the parameters are the same as in Theorem 2, and $0 < z < 1$ *represents the running value of the statistic* $\Lambda_1 = e^{-W_1}$*, may be used as near-exact distributions for this statistic.*

Proof: Since the near-exact distributions developed in Theorem 2 were for the random variable $W_1 = -\log \Lambda_1$ we only need to mind the relation

$$F_{\Lambda_1}(z) = 1 - F_{W_1}(-\log z)$$

where $F_{\Lambda}(\cdot)$ is the cumulative distribution function of Λ_1 and $F_{W_1}(\cdot)$ is the cumulative distribution function of W_1, in order to obtain the corresponding near-exact distributions for Λ_1. $\square$

Indeed in order to obtain near-exact α-quantiles for Λ_1 we do not even need the near-exact distributions for Λ_1, since if we consider the relation

$$\Lambda_1(\alpha) = e^{-W_1(1-\alpha)}\,,$$

where $\Lambda_1(\alpha)$ is the α-quantile of Λ_1 and $W_1(1-\alpha)$ is the $(1-\alpha)$-quantile of W_1 we may easily obtain the near-exact α-quantiles of Λ_1 from the corresponding $(1-\alpha)$-quantiles of W_1.

3 Near-exact distributions for the likelihood ratio test statistic of sphericity

Lemma 4 *The c.f. of* $W = -\log\Lambda$*, where* Λ *is the test statistic in* (5) *may be writen as*

$$\Phi_W(t) = \underbrace{\prod_{j=p-k^*+1}^{p} \frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p}\right)\Gamma\left(\frac{n+1}{2}-it\right)}{\Gamma\left(\frac{n+1}{2}\right)\Gamma\left(\frac{n}{2}+\frac{j-1}{p}-it\right)} \prod_{j=1}^{p-k^*} \frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p}\right)\Gamma\left(\frac{n}{2}-it\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}+\frac{j-1}{p}-it\right)}}_{\Phi_1^{**}(t)} \times$$
$$\underbrace{\prod_{k=1}^{p-1}\left(\frac{n-k}{2}\right)^{k^*-\left\lfloor\frac{k-(p \perp\!\!\!\perp 2)}{2}\right\rfloor}\left(\frac{n-k}{2}-it\right)^{-k^*+\left\lfloor\frac{k-(p \perp\!\!\!\perp 2)}{2}\right\rfloor}}_{\Phi_2^{**}(t)}\,, \qquad (26)$$

where $k^* = \lfloor p/2 \rfloor$*. In* (26) *above,* $\Phi_1^{**}(t)$ *is the c.f. of the sum of* p *independent Logbeta r.v.'s,* k^* *of which with parameters* $(n+1)/2$ *and* $(j-1)/p - 1/2$ $(j = p-k^*+1,\ldots,p)$ *and the remaining* $p-k^*$ *with parameters* $n/2$ *and* $(j-1)/p$ $(j = 1,\ldots,p-k^*)$ *and* $\Phi_2^{**}(t)$ *is the c.f. of a GIG distribution of depth* $p-1$*, with rate parameters* $(n-k)/2$ *and shape parameters* $k^*-\left\lfloor\frac{k-(p \perp\!\!\!\perp 2)}{2}\right\rfloor$ $(k = 1,\ldots,p-1)$*.*

Proof: From (7) we have

$$E\left(\Lambda_2^h\right) \,=\, p^{ph}\,\frac{\Gamma\left(\frac{np}{2}\right)}{\Gamma\left(\frac{np}{2}+hp\right)} \prod_{j=1}^{p} \frac{\Gamma\left(\frac{n}{2}+h\right)}{\Gamma\left(\frac{n}{2}\right)}$$

so that taking $W_2 = -\log\Lambda_2$ and using the multiplication formula for the Gamma function

$$\Gamma(kz) \,=\, (2\pi)^{-(k-1)/2}\, k^{kz-1/2} \prod_{i=1}^{k} \Gamma\left(z+\frac{i-1}{k}\right)\,,$$

we have

$$
\begin{aligned}
\Phi_{W_2}(t) &= E\left(e^{itW_2}\right) = E\left(e^{-it\log\Lambda_2}\right) = E\left(\Lambda_2^{-it}\right) \\
&= p^{-itp}\,\frac{\Gamma\left(\frac{np}{2}\right)}{\Gamma\left(\frac{np}{2}-itp\right)}\prod_{j=1}^{p}\frac{\Gamma\left(\frac{n}{2}-it\right)}{\Gamma\left(\frac{n}{2}\right)} \\
&= \prod_{j=1}^{p}\frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p}\right)\Gamma\left(\frac{n}{2}-it\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}+\frac{j-1}{p}-it\right)}\,.
\end{aligned}
\tag{27}
$$

Thus, given the definition of k^* in the previous section, which may indeed be written as

$$
k^* = \left\lfloor \frac{p}{2} \right\rfloor = \begin{cases} \frac{p}{2} & \text{even } p \\ \frac{p-1}{2} & \text{odd } p\,, \end{cases}
$$

and given that then for $p-k^*+1 \le j \le p$,

$$
\frac{j-1}{p}+\frac{1}{2} \ge 1
$$

since for $p-k^*+1 \le j \le p$ we have

$$
\frac{j-1}{p}+\frac{1}{2} \ge \frac{p-k^*+1-1}{p}+\frac{1}{2} = \begin{cases} \frac{p-p/2}{p}+\frac{1}{2}=1 & \text{even } p \\ \frac{p-(p-1)/2}{p}+\frac{1}{2}=1+\frac{1}{2p} & \text{odd p}\,, \end{cases}
$$

and given that under H_0 in (2), Λ_1 and Λ_2 are indepenent, we may write, from (13) and (27), the c.f. of $W=-\log\Lambda$, where Λ is the statistic in (5), as

$$
\begin{aligned}
\Phi_W(t) &= \Phi_{W_1}(t)\,\Phi_{W_2}(t) \\
&= \Phi_2^*(t)\,\left(\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}-\frac{1}{2}-it\right)}{\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)\Gamma\left(\frac{n}{2}-it\right)}\right)^{k^*}\prod_{j=1}^{p}\frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p}\right)\Gamma\left(\frac{n}{2}-it\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}+\frac{j-1}{p}-it\right)} \\
&= \Phi_2^*(t)\prod_{j=1}^{p-k^*}\frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p}\right)\Gamma\left(\frac{n}{2}-it\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}+\frac{j-1}{p}-it\right)}\prod_{j=p-k^*+1}^{p}\frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p}\right)\Gamma\left(\frac{n-1}{2}-it\right)}{\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n}{2}+\frac{j-1}{p}-it\right)} \\
&= \Phi_2^*(t)\,\times \\
&\quad \prod_{j=1}^{p-k^*}\frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p}\right)\Gamma\left(\frac{n}{2}-it\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}+\frac{j-1}{p}-it\right)}\prod_{j=p-k^*+1}^{p}\frac{\Gamma\left(\frac{n-1}{2}+\frac{1}{2}+\frac{j-1}{p}\right)\Gamma\left(\frac{n-1}{2}-it\right)}{\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-1}{2}+\frac{1}{2}+\frac{j-1}{p}-it\right)}
\end{aligned}
$$

$$= \prod_{j=p-k^*+1}^{p} \frac{\Gamma\left(\frac{n-1}{2}+1\right)\,\Gamma\left(\frac{n-1}{2}-it\right)}{\Gamma\left(\frac{n-1}{2}\right)\,\Gamma\left(\frac{n-1}{2}+1-it\right)}\,\frac{\Gamma\left(\frac{n-1}{2}+\frac{1}{2}+\frac{j-1}{p}\right)\,\Gamma\left(\frac{n-1}{2}+1-it\right)}{\Gamma\left(\frac{n-1}{2}+1\right)\,\Gamma\left(\frac{n-1}{2}+\frac{1}{2}+\frac{j-1}{p}-it\right)} \times$$

$$\prod_{j=1}^{p-k^*} \frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p}\right)\,\Gamma\left(\frac{n}{2}-it\right)}{\Gamma\left(\frac{n}{2}\right)\,\Gamma\left(\frac{n}{2}+\frac{j-1}{p}-it\right)} \times \Phi_2^*(t)$$

$$= \underbrace{\left(\left(\frac{n-1}{2}\right)\left(\frac{n-1}{2}-it\right)^{-1}\right)^{k^*} \times \Phi_2^*(t)}_{\Phi_2^{**}(t)}$$

$$\underbrace{\prod_{j=p-k^*+1}^{p} \frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p}\right)\,\Gamma\left(\frac{n+1}{2}-it\right)}{\Gamma\left(\frac{n+1}{2}\right)\,\Gamma\left(\frac{n}{2}+\frac{j-1}{p}-it\right)} \prod_{j=1}^{p-k^*} \frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p}\right)\,\Gamma\left(\frac{n}{2}-it\right)}{\Gamma\left(\frac{n}{2}\right)\,\Gamma\left(\frac{n}{2}+\frac{j-1}{p}-it\right)}}_{\Phi_1^{**}(t)} . \ \square$$

Then, by replacing $\Phi_1^{**}(t)$ in (26) by the c.f. of a Gamma distribution or the c.f. of a mixture of Gamma distributions, we will get near-exact distributions for $W = -\log \Lambda$ under the form of a GNIG distribution or mixtures of GNIG distributions.

Surprisingly enough, as we will see in the next section, these near-exact distributions have an even better performance than the already well-fit ones in [19].

Theorem 5 *Using for $\Phi_1^{**}(t)$ in (26), the approximations:*

- $\lambda^s(\lambda - \mathrm{i}t)^{-s}$ *with $s, \lambda > 0$, such that*

$$\frac{\partial^h}{\partial t^h}\,\lambda^s(\lambda-\mathrm{i}t)^{-s}\bigg|_{t=0} = \frac{\partial^h}{\partial t^h}\,\Phi_1^{**}(t)\bigg|_{t=0} \qquad \textit{for} \quad h=1,2\,; \tag{28}$$

- $\displaystyle\sum_{k=1}^{2} \theta_k\,\mu^{s_k}(\mu-\mathrm{i}t)^{-s_k}$, *where $\theta_2 = 1-\theta_1$ with $\theta_k,\ s_k,\ \mu > 0$, such that*

$$\frac{\partial^h}{\partial t^h}\,\sum_{k=1}^{2} \theta_k\,\mu^{s_k}(\mu-\mathrm{i}t)^{-s_k}\bigg|_{t=0} = \frac{\partial^h}{\partial t^h}\,\Phi_1^{**}(t)\bigg|_{t=0} \qquad \textit{for} \quad h=1,\dots,4\,; \tag{29}$$

- $\displaystyle\sum_{k=1}^{3} \theta_k^*\,\nu^{s_k^*}(\nu-\mathrm{i}t)^{-s_k^*}$, *where $\theta_3^* = 1-\theta_1^*-\theta_2^*$ with $\theta_k^*,\ s_k^*,\ \nu > 0$, such that*

$$\frac{\partial^h}{\partial t^h}\,\sum_{k=1}^{3} \theta_k^*\,\nu^{s_k^*}(\nu-\mathrm{i}t)^{-s_k^*}\bigg|_{t=0} = \frac{\partial^h}{\partial t^h}\,\Phi_1^{**}(t)\bigg|_{t=0} \qquad \textit{for} \quad h=1,\dots,6\,; \tag{30}$$

we obtain as near-exact distributions for W, respectively,

i) a GNIG distribution of depth p with cdf (using the notation in (42) *in Appendix A)*

$$F(w|r_1,\ldots,r_{p-1},s;\lambda_1,\ldots,\lambda_{p-1},\lambda)\,, \tag{31}$$

where

$$r_j=\left\lfloor\frac{p}{2}\right\rfloor-\left\lfloor\frac{j-(p\perp\!\!\!\perp 2)}{2}\right\rfloor\,,\quad \lambda_j=\frac{n-j}{2}\,,\quad (j=1,\ldots,p-1), \tag{32}$$

and

$$\lambda=\frac{m_1}{m_2-m_1^2}\qquad\text{and}\qquad s=\frac{m_1^2}{m_2-m_1^2} \tag{33}$$

with

$$m_h=\mathrm{i}^{-h}\,\left.\frac{\partial^h}{\partial t^h}\,\Phi_1^{**}(t)\right|_{t=0}\,,\qquad h=1,2\,;$$

ii) a mixture of two GNIG distributions of depth p, with cdf (using the notation in (42) *in Appendix A)*

$$\sum_{k=1}^{2}\theta_k\,F(w|r_1,\ldots,r_{p-1},s_k;\lambda_1,\ldots,\lambda_{p-1},\mu)\,, \tag{34}$$

where r_j and λ_j ($j=1,\ldots,p-1$) are given by (32) *above and θ_1, μ, r_1 and r_2 are obtained from the numerical solution of the system of four equations*

$$\sum_{k=1}^{2}\theta_k\,\frac{\Gamma(r_k+h)}{\Gamma(r_k)}\,\mu^{-h}\;=\;\mathrm{i}^{-h}\,\left.\frac{\partial^h}{\partial t^h}\,\Phi_1^{**}(t)\right|_{t=0}\qquad(h=1,\ldots,4) \tag{35}$$

for these parameters, with $\theta_2=1-\theta_1$;

iii) or a mixture of three GNIG distributions of depth p, with cdf (using the notation in (42) *in Appendix A)*

$$\sum_{k=1}^{3}\theta_k^*\,F(w|r_1,\ldots,r_{p-1},s_k^*;\lambda_1,\ldots,\lambda_{p-1},\nu)\,, \tag{36}$$

with r_j and λ_j ($j=1,\ldots,p-1$) given by (32) *above and θ_1^*, θ_2^*, ν, s_1^*, s_2^* and s_3^* obtained from the numerical solution of the system of six equations*

$$\sum_{k=1}^{3}\theta_j^*\,\frac{\Gamma(r_k^*+h)}{\Gamma(r_k^*)}\,\nu^{-h}\;=\;\mathrm{i}^{-h}\,\left.\frac{\partial^h}{\partial t^h}\,\Phi_1^{**}(t)\right|_{t=0}\qquad(h=1,\ldots,6) \tag{37}$$

for these parameters, with $\theta_3^*=1-\theta_1^*-\theta_2^*$.

Proof: The proof of this Theorem is in all similar to the proof of Theorem 2, more precisely, if in the characteristic function of W in (26) we replace $\Phi_1^{**}(t)$ by $\lambda^s(\lambda-\mathrm{i}t)^{-s}$ we obtain

$$\Phi_W(t)\;\approx\;\lambda^s(\lambda-\mathrm{i}t)^{-s}\;\underbrace{\prod_{k=1}^{p-1}\left(\frac{n-k}{2}\right)^{k^*-\left\lfloor\frac{k-(p\perp\!\!\!\perp 2)}{2}\right\rfloor}\left(\frac{n-k}{2}-it\right)^{-k^*+\left\lfloor\frac{k-(p\perp\!\!\!\perp 2)}{2}\right\rfloor}}_{\Phi_2^{**}(t)}\,,$$

that is the characteristic function of the sum of p independent Gamma random variables, $p-1$ of which with integer shape parameters r_j and rate parameters λ_j given by (32), and a further Gamma random variable with rate parameter $s>0$ and shape parameter λ. This characteristic function is thus the c.f. of the GNIG distribution of depth p with distribution function given in (31). The parameters s and λ are determined in such a way that (28) holds. This compels s and λ to be given by (33) and makes the two first moments of this near-exact distribution for W to be the same as the two first exact moments of W.

If in the characteristic function of W in (26) we replace $\Phi_1^{**}(t)$ by $\sum_{k=1}^{2}\theta_k\,\mu^{r_k}(\mu-\mathrm{i}t)^{-r_k}$ we obtain

$$\Phi_W(t)\approx\sum_{k=1}^{2}\theta_k\,\mu^{r_k}(\mu-\mathrm{i}t)^{-r_k}\;\underbrace{\prod_{k=1}^{p-1}\left(\frac{n-k}{2}\right)^{k^*-\left\lfloor\frac{k-(p\perp\!\!\!\perp 2)}{2}\right\rfloor}\left(\frac{n-k}{2}-it\right)^{-k^*+\left\lfloor\frac{k-(p\perp\!\!\!\perp 2)}{2}\right\rfloor}}_{\Phi_2^{**}(t)}\,,$$

that is the characteristic function of the mixture of two GNIG distributions of depth p with density function given in (34). The parameters θ_1, μ, r_1 and r_2 are defined in such a way that (29) holds, giving rise to the evaluation of these parameters as the numerical solution of the system of equations in (35) and to a near-exact distribution that matches the first four exact moments of W.

If in the characteristic function of W in (26) we replace $\Phi_1^{**}(t)$ by $\sum_{k=1}^{3}\theta_k^*\,\nu^{r_k^*}(\nu-\mathrm{i}t)^{-r_k^*}$ we obtain

$$\Phi_W(t)\;\approx\;\sum_{k=1}^{3}\theta_k^*\,\nu^{r_k^*}(\nu-\mathrm{i}t)^{-r_k^*}\;\times$$
$$\underbrace{\prod_{k=1}^{p-1}\left(\frac{n-k}{2}\right)^{k^*-\left\lfloor\frac{k-(p\perp\!\!\!\perp 2)}{2}\right\rfloor}\left(\frac{n-k}{2}-it\right)^{-k^*+\left\lfloor\frac{k-(p\perp\!\!\!\perp 2)}{2}\right\rfloor}}_{\Phi_2^{**}(t)}\,,$$

that is the characteristic function of the mixture of three GNIG distributions of depth $p-1$ with density function given in (34). The parameters θ_1^*, θ_2^*, ν, r_1^*, r_2^* and r_3^* are defined in such a way that (30) holds, what gives rise to the evaluation of these parameters as the numerical solution of the system of equations in (37), giving rise to a near-exact distribution that matches the first six exact moments of W. □

Corollary 6 *Distributions with cdf's given by*

i) $1-F(-\log z|r_1,\dots,r_{p-1},s;\lambda_1,\dots,\lambda_{p-1},\lambda)\,,$

ii) $1-\sum_{k=1}^{2}\theta_k\,F(-\log z|r_1,\dots,r_{p-1},s_k;\lambda_1,\dots,\lambda_{p-1},\mu)\,,$ *or*

iii) $1-\sum_{k=1}^{3}\theta_k^*\,F(-\log z|r_1,\dots,r_{p-1},s_k^*;\lambda_1,\dots,\lambda_{p-1},\nu)\,,$

where the parameters are the same as in Theorem 5, and $0<z<1$ *represents the running value of the statistic* $\Lambda=e^{-W}$*, may be used as near-exact distributions for this statistic.*

The proof of this Corollary is in all similar to the proof of Corollary 3 and also similar considerations to the ones right after Corollary 3, concerning the computation of near-exact quantiles of the statistics W_1 and Λ_1, apply here to the computation of near-exact quantiles of the statistics W and Λ.

4 Numerical and comparative studies

In order to evaluate the quality of the near-exact approximations developed for the likelihood ratio test statistics for testing independence in a set of variables and for the sphericity test we use, whenever the c.f.'s are available, two measures of proximity,

$$\Delta_1=\int_{-\infty}^{\infty}|\phi_Y(t)-\phi_n(t)|\,\mathrm{d}t \quad\text{and}\quad \Delta_2=\frac{1}{2\pi}\int_{-\infty}^{\infty}\left|\frac{\phi_Y(t)-\phi_n(t)}{t}\right|\,\mathrm{d}t\,, \quad (38)$$

with

$$\max_{y\in S}|f_Y(y)-f_n(y)|\le\frac{1}{2\pi}\Delta_1 \quad\text{and}\quad \max_{y\in S}|F_Y(y)-F_n(y)|\;\le\;\Delta_2\,, \quad (39)$$

where Y represents a continuous random variable defined on S with distribution function $F_Y(y)$, density function $f_Y(y)$ and characteristic function $\phi_Y(t)$, and $\phi_n(t)$, $F_n(y)$ and $f_n(y)$ represent respectively the characteristic, distribution and density function of a random variable X_n.

These two measures may be derived directly from inversion formulas, and Δ_2 may be seen as based on the Berry-Esseen upper bound on $|F_Y(y) - F_n(y)|$ (see [5], [13], [16], Chap. VI, sec. 21 in [18]).

We should note that for continuous random variables,

$$\lim_{n\to\infty} \Delta_1 = 0 \iff \lim_{n\to\infty} \Delta_2 = 0 \qquad (40)$$

and either one of the equalities above imply that

$$X_n \xrightarrow{d} Y\,. \qquad (41)$$

For further details on these measures see [19], where they are used to study the quality of near-exact distributions for the sphericity test statistic.

4.1 Studies for the independence test statistic

Mudholkar et al. in [20] developed a Normal approximation to the distribution of the likelihood ratio test statistic used for testing H_{01} in (2). These authors presented numerical studies comparing their Normal approximation with the approximations due to Box and Bartlett ([6], [4]).

Since the asymptotic Normal approximation from Mudholkar et al. in [20] yields indeed for $\log \Lambda_1$ a non-central generalized Gamma distribution, whose c.f. is not manageable, in order to compare the performance of the near-exact distribution developed with this Normal asymptotic approximation, instead of using measures Δ_1 and Δ_2, we decided to use a similar method to the one used in [20] to assess the performance of their Normal asymptotic approximation.

We used the exact quantiles for Λ_1 computed directly from the numerical inversion of the c.f. of $\log \Lambda_1$ by using the Gil-Pelaez inversion formulas (see [14]) what gives us a precision at least equal to the one used in [20] in terms of exact quantiles, which in turn give for the Normal asymptotic approximation of [20] exactly the same results obtained by these authors.

Table 1 – Values of the tail probability error = $(approx.\ prob - \alpha) \times 10^5$ for the near-exact distributions

			α					
p	n		0.005	0.01	0.05	0.10	0.20	0.50
3	6	GNIG	1.29×10^{0}	2.72×10^{0}	-1.23×10^{0}	-7.92×10^{0}	-1.65×10^{1}	-5.75×10^{0}
		M2GNIG	-9.91×10^{-1}	-9.77×10^{-2}	-4.20×10^{-1}	-1.82×10^{-1}	5.42×10^{-1}	-6.78×10^{-1}
	8	GNIG	1.37×10^{0}	1.10×10^{0}	-5.45×10^{-1}	-5.36×10^{0}	-9.40×10^{0}	-2.21×10^{0}
		M2GNIG	5.51×10^{-2}	-3.37×10^{-1}	3.26×10^{-1}	-6.20×10^{-1}	4.26×10^{-2}	-3.71×10^{-1}
	13	GNIG	1.26×10^{0}	3.47×10^{-1}	-2.61×10^{-1}	-1.76×10^{0}	-3.30×10^{0}	-5.33×10^{-1}
		M2GNIG	7.79×10^{-1}	-1.85×10^{-1}	1.75×10^{-1}	4.88×10^{-2}	4.69×10^{-2}	-1.13×10^{-1}
4	7	GNIG	1.93×10^{0}	1.55×10^{0}	1.17×10^{0}	-1.57×10^{0}	-6.80×10^{0}	-7.90×10^{0}
		M2GNIG	7.73×10^{-1}	-6.58×10^{-2}	3.44×10^{-3}	1.24×10^{-2}	8.95×10^{-2}	-1.17×10^{-2}
	9	GNIG	8.44×10^{-1}	1.10×10^{0}	5.08×10^{-1}	-1.38×10^{0}	-4.60×10^{0}	-4.48×10^{0}
		M2GNIG	2.58×10^{-2}	1.16×10^{-2}	-5.76×10^{-3}	1.80×10^{-2}	6.38×10^{-2}	-3.87×10^{-2}
	14	GNIG	3.94×10^{-1}	4.63×10^{-1}	1.80×10^{-1}	-7.26×10^{-1}	-1.99×10^{0}	-1.67×10^{0}
		M2GNIG	2.48×10^{-2}	-4.79×10^{-3}	6.41×10^{-2}	1.90×10^{-3}	4.56×10^{-2}	-3.17×10^{-2}
5	8	GNIG	1.08×10^{0}	5.16×10^{-1}	6.09×10^{-1}	-8.25×10^{-2}	-1.69×10^{0}	-3.05×10^{0}
		M2GNIG	7.16×10^{-1}	-1.33×10^{-2}	3.49×10^{-3}	-6.88×10^{-3}	7.96×10^{-3}	1.40×10^{-2}
	10	GNIG	3.14×10^{-1}	3.81×10^{-1}	3.73×10^{-1}	-1.88×10^{-1}	-1.35×10^{0}	-2.01×10^{0}
		M2GNIG	1.92×10^{-2}	-3.04×10^{-2}	-3.85×10^{-3}	1.07×10^{-3}	8.56×10^{-3}	5.78×10^{-3}
	15	GNIG	1.55×10^{-1}	2.54×10^{-1}	1.79×10^{-1}	-1.35×10^{-1}	-6.74×10^{-1}	-8.84×10^{-1}
		M2GNIG	-2.47×10^{-3}	4.37×10^{-2}	3.36×10^{-2}	2.17×10^{-2}	3.85×10^{-2}	4.01×10^{-4}
6	9	GNIG	2.05×10^{-1}	2.97×10^{-1}	4.25×10^{-1}	9.65×10^{-2}	-7.81×10^{-1}	-1.83×10^{0}
		M2GNIG	1.58×10^{-3}	-8.79×10^{-3}	1.85×10^{-3}	-3.24×10^{-3}	-3.26×10^{-3}	7.92×10^{-3}
	11	GNIG	1.33×10^{-1}	3.20×10^{-1}	3.02×10^{-1}	1.10×10^{-3}	-6.97×10^{-1}	-1.35×10^{0}
		M2GNIG	-5.09×10^{-2}	5.59×10^{-2}	-2.03×10^{-3}	-1.02×10^{-3}	6.30×10^{-3}	4.94×10^{-3}
	16	GNIG	1.15×10^{-1}	1.59×10^{-1}	1.43×10^{-1}	4.31×10^{-2}	4.27×10^{-1}	-6.82×10^{-1}
		M2GNIG	2.80×10^{-3}	4.49×10^{-3}	-1.08×10^{-3}	4.80×10^{-4}	2.21×10^{-3}	1.56×10^{-3}
7	10	GNIG	9.08×10^{-2}	1.30×10^{-1}	2.02×10^{-1}	8.12×10^{-2}	-2.81×10^{-1}	-8.07×10^{-1}
		M2GNIG	2.31×10^{-3}	-4.00×10^{-3}	-9.51×10^{-4}	-7.33×10^{-4}	4.77×10^{-4}	2.44×10^{-3}
	12	GNIG	8.85×10^{-2}	1.25×10^{-1}	1.61×10^{-1}	3.58×10^{-2}	-2.81×10^{-1}	-6.47×10^{-1}
		M2GNIG	2.55×10^{-3}	-1.68×10^{-4}	-1.05×10^{-3}	-1.62×10^{-3}	-1.31×10^{-3}	1.91×10^{-3}
	17	GNIG	7.50×10^{-2}	8.12×10^{-2}	8.80×10^{-2}	-6.53×10^{-2}	-1.93×10^{-1}	-3.65×10^{-1}
		M2GNIG	1.68×10^{-2}	-1.78×10^{-4}	-4.01×10^{-4}	-6.63×10^{-2}	5.78×10^{-4}	8.09×10^{-4}
10	13	GNIG	2.12×10^{-2}	3.25×10^{-2}	5.45×10^{-2}	3.26×10^{-2}	-4.52×10^{-2}	-1.87×10^{-1}
		M2GNIG	7.02×10^{-6}	-2.87×10^{-7}	-7.41×10^{-5}	-9.29×10^{-5}	-5.28×10^{-6}	2.01×10^{-4}
	15	GNIG	2.35×10^{-2}	3.49×10^{-2}	5.19×10^{-2}	2.53×10^{-2}	-5.39×10^{-2}	-1.75×10^{-1}
		M2GNIG	1.09×10^{-5}	-4.74×10^{-6}	-8.76×10^{-5}	-8.87×10^{-5}	2.41×10^{-5}	2.09×10^{-4}
	20	GNIG	1.97×10^{-2}	2.82×10^{-2}	3.70×10^{-2}	1.32×10^{-2}	-4.80×10^{-2}	-1.25×10^{-1}
		M2GNIG	1.50×10^{-5}	-2.32×10^{-5}	-5.39×10^{-5}	-6.23×10^{-5}	2.71×10^{-5}	1.48×10^{-4}

However, given that the exact quantiles computed in this way have a precision that does not go beyond 12 digits and given that this precision is not enough for making comparisons with the near-exact distribution M3GNIG, which requires a higher precision, we have used in Table 1 only the near-exact distributions GNIG and M2GNIG.

In Table 1 the errors displayed are evaluated using the exact same method used by Mudholkar et al. in [20], the difference between the approximate and the exact tail probabilities multiplied by 10^5. The values considered for p and n correspond to the same cases considered by Mudholkar et al. in [20]. We can observe that the errors obtained when using the near-exact distributions are always much smaller than the ones given in Table 1 of Mudholkar et al. in [20] for their Normal approximation, mainly for larger values of p.

Table 2– Values of the measures Δ_1 and Δ_2 for the near-exact distributions

		Δ_1			Δ_2		
p	n	*GNIG*	*M2GNIG*	*M3GNIG*	*GNIG*	*M2GNIG*	*M3GNIG*
3	6	5.8×10^{-2}	5.7×10^{-3}	2.4×10^{-3}	5.1×10^{-4}	2.1×10^{-5}	4.6×10^{-6}
	8	4.4×10^{-2}	4.2×10^{-3}	5.3×10^{-4}	2.7×10^{-4}	1.1×10^{-5}	8.9×10^{-7}
	13	2.7×10^{-2}	1.9×10^{-3}	2.7×10^{-5}	9.4×10^{-5}	3.1×10^{-6}	3.3×10^{-8}
4	7	4.7×10^{-3}	1.8×10^{-4}	1.4×10^{-5}	1.7×10^{-4}	3.7×10^{-6}	2.0×10^{-7}
	9	4.0×10^{-3}	1.4×10^{-4}	5.4×10^{-6}	1.1×10^{-4}	2.3×10^{-6}	6.1×10^{-8}
	14	2.7×10^{-3}	7.8×10^{-5}	3.7×10^{-7}	4.4×10^{-5}	7.4×10^{-7}	3.0×10^{-9}
5	8	8.1×10^{-4}	1.4×10^{-5}	4.4×10^{-7}	4.7×10^{-5}	5.2×10^{-7}	1.2×10^{-8}
	10	7.5×10^{-4}	1.3×10^{-5}	2.1×10^{-7}	3.2×10^{-5}	3.6×10^{-7}	4.4×10^{-9}
	15	5.8×10^{-4}	8.1×10^{-6}	1.8×10^{-8}	1.5×10^{-5}	1.4×10^{-7}	2.6×10^{-10}
6	9	3.3×10^{-4}	3.5×10^{-6}	3.6×10^{-8}	2.4×10^{-5}	1.7×10^{-7}	1.3×10^{-9}
	11	3.3×10^{-4}	3.4×10^{-6}	1.9×10^{-8}	1.8×10^{-5}	1.3×10^{-7}	5.5×10^{-10}
	16	2.8×10^{-4}	2.4×10^{-6}	2.1×10^{-9}	9.5×10^{-6}	5.6×10^{-8}	3.5×10^{-11}
7	10	1.2×10^{-4}	7.3×10^{-7}	3.7×10^{-9}	9.7×10^{-6}	4.2×10^{-8}	1.7×10^{-10}
	12	1.2×10^{-4}	7.7×10^{-7}	2.1×10^{-9}	7.9×10^{-6}	3.4×10^{-8}	7.4×10^{-11}
	17	1.1×10^{-4}	6.1×10^{-7}	4.0×10^{-10}	4.6×10^{-6}	1.8×10^{-8}	8.3×10^{-12}
10	13	1.9×10^{-5}	4.2×10^{-8}	6.2×10^{-12}	2.0×10^{-6}	3.2×10^{-9}	3.5×10^{-13}
	15	2.2×10^{-5}	5.1×10^{-8}	2.2×10^{-11}	1.9×10^{-6}	3.1×10^{-9}	1.1×10^{-12}
	20	2.4×10^{-5}	5.1×10^{-8}	4.6×10^{-11}	1.4×10^{-6}	2.1×10^{-9}	1.5×10^{-12}
20	23	5.2×10^{-7}	1.4×10^{-10}	3.0×10^{-14}	7.8×10^{-8}	1.5×10^{-11}	2.7×10^{-15}
	50	1.1×10^{-6}	2.8×10^{-10}	7.4×10^{-14}	4.9×10^{-8}	9.7×10^{-12}	2.1×10^{-15}
	100	7.4×10^{-7}	1.2×10^{-10}	2.5×10^{-14}	1.5×10^{-8}	1.8×10^{-12}	3.2×10^{-16}
50	53	5.8×10^{-9}	1.0×10^{-13}	1.5×10^{-19}	1.1×10^{-9}	1.5×10^{-14}	1.9×10^{-20}
	100	1.1×10^{-6}	2.8×10^{-10}	1.9×10^{-17}	4.9×10^{-8}	9.7×10^{-12}	7.5×10^{-19}
	150	2.3×10^{-8}	4.6×10^{-13}	4.3×10^{-19}	8.5×10^{-10}	1.3×10^{-14}	1.0×10^{-20}
	200	2.0×10^{-8}	3.2×10^{-13}	3.3×10^{-19}	5.3×10^{-10}	6.5×10^{-15}	5.7×10^{-21}
	500	9.8×10^{-9}	7.1×10^{-14}	5.0×10^{-20}	1.0×10^{-12}	5.4×10^{-16}	3.2×10^{-22}

In Table 2 we use measures Δ_1 and Δ_2 to better assess the relative performance of the three near-exact distributions GNIG, M2GNIG and M3GNIG as approximating distributions for the independence test statistic.

From Table 2 we may easily see that the near-exact distribution M3GNIG has always a better performance than the other two near-exact distributions and we can also see that the near-exact distribution M2GNIG always outperforms the GNIG near-exact distribution. The values exhibited for the M3GNIG distribution for both measures, mainly for the measure Δ_2, which represents an upper bound for the absolute value of the difference between its c.d.f. and the exact c.d.f., would lead us to recommend its use as a replacement for the exact distribution, mainly for larger values of p. The three near-exact distributions display a marked asymptotic behavior both for increasing sample sizes and increasing number of variables, although for larger values of p we need large enough sample sizes in order to be able to observe their asymptotic behavior in terms of increasing values of n.

The tables in this subsection present the values of the measures Δ_1 and Δ_2 given in (38) for the new near-exact distributions developed in this paper for the likelihood ratio test statistic used for testing sphericity. In this subsection our purpose is to assess the quality of the new near-exact distributions comparing them with the ones already developed, using a different method, in [19]. In order to achieve our purpose we have considered the exact same values for n and p already considered in the numerical studies presented in that reference. We will denote the new near-exact distributions corresponding to the GNIG distribution, the mixture of two GNIG distribution and the mixture of three GNIG distributions respectively by GNIGnew, M2GNIGnew and M3GNIGnew (leaving the names GNIG, M2GNIG and M3GNIG, used in Table 7, for the corresponding near-exact distributions developed in [19]).

Table 3 – Values of Δ_1 and Δ_2 for the near-exact distributions for $W = -\log\Lambda$, for $p=4$, $n=6$ and $p=5$, $n=7$

	$p=4, n=6$		$p=5, n=7$	
	Δ_1	Δ_2	Δ_1	Δ_2
GNIGnew	3.815×10^{-5}	2.408×10^{-6}	2.802×10^{-5}	2.300×10^{-6}
M2GNIGnew	5.617×10^{-7}	2.198×10^{-8}	2.580×10^{-7}	1.378×10^{-8}
M3GNIGnew	3.875×10^{-9}	1.180×10^{-10}	4.936×10^{-9}	1.992×10^{-10}

Table 4 – Values of Δ_1 and Δ_2 for the near-exact distributions for $W = -\log\Lambda$, for $p=7$, $n=9$ and $p=10$, $n=12$

	$p=7, n=9$		$p=10, n=12$	
	Δ_1	Δ_2	Δ_1	Δ_2
GNIGnew	3.953×10^{-6}	4.234×10^{-7}	3.366×10^{-7}	4.402×10^{-8}
M2GNIGnew	1.363×10^{-8}	1.007×10^{-9}	2.863×10^{-10}	2.667×10^{-11}
M3GNIGnew	5.617×10^{-11}	3.241×10^{-12}	8.847×10^{-14}	6.518×10^{-15}

Comparing Tables 3 and 4 with Tables 1 and 2 in [19] we can observe that the values for the new approximations are always better with the exception of M3GNIGnew for $p=5$, $n=7$ and $p=7$, $n=9$.

Table 5 – Values of Δ_1 and Δ_2 for the near-exact distributions for $W = -\log\Lambda$, for $p=4,5,7$ and $n=50$

	$p=4, n=50$		$p=5, n=50$		$p=7, n=50$	
	Δ_1	Δ_2	Δ_1	Δ_2	Δ_1	Δ_2
GNIGnew	9.702×10^{-6}	6.005×10^{-8}	9.490×10^{-6}	8.493×10^{-8}	2.376×10^{-6}	3.323×10^{-8}
M2GNIGnew	2.682×10^{-8}	1.105×10^{-10}	2.936×10^{-8}	1.835×10^{-10}	3.574×10^{-9}	3.618×10^{-11}
M3GNIGnew	1.233×10^{-10}	3.818×10^{-13}	4.932×10^{-11}	2.389×10^{-13}	2.760×10^{-12}	2.243×10^{-14}

Table 6 – Values of Δ_1 and Δ_2 for the near-exact distributions for $W = -\log\Lambda$, for $p = 10, 20, 30$ and $n = 50$

	$p=10, n=50$		$p=20, n=50$		$p=30, n=50$	
	Δ_1	Δ_2	Δ_1	Δ_2	Δ_1	Δ_2
GNIGnew	3.831×10^{-7}	8.153×10^{-9}	3.186×10^{-8}	1.513×10^{-9}	5.781×10^{-9}	4.646×10^{-10}
M2GNIGnew	1.713×10^{-10}	2.686×10^{-12}	3.622×10^{-12}	1.283×10^{-13}	2.609×10^{-13}	1.567×10^{-14}
M3GNIGnew	6.236×10^{-14}	8.006×10^{-16}	3.602×10^{-16}	1.057×10^{-17}	1.121×10^{-17}	5.588×10^{-19}

Comparing Tables 5 and 6 with Tables 3 and 4 in [19] we may verify that in almost all cases we have for the new near-exact approximations smaller values for the measures Δ_1 and Δ_2. The only case where this fact does not happen is when $p = 4$ and $n = 50$ for the measures of the M2GNIGnew and M3GNIGnew distributions. These new near-exact approximations also exhibit the good asymptotic properties revealed by the near-exact approximations in [19].

We may say as a general conclusion that the new near-exact distributions have a better performance than the ones developed in [19] for large values of n with p large enough ($p \geq 4$). Moreover in the next Table we may see that for large values of p and values of n close to p we also have better values of both measures for the near-exact distributions developed in this paper.

Table 7 – Values of Δ_1 and Δ_2 for the near-exact distributions for $W = -\log\Lambda$, for $p = 10, 20, 30$ and $n = 12, 22, 32$

	$p=10, n=12$		$p=20, n=22$		$p=30, n=32$	
	Δ_1	Δ_2	Δ_1	Δ_2	Δ_1	Δ_2
GNIGnew	1.058×10^{-6}	1.383×10^{-7}	3.526×10^{-8}	6.034×10^{-9}	5.043×10^{-9}	9.685×10^{-10}
GNIG	8.940×10^{-6}	1.171×10^{-6}	3.634×10^{-7}	6.221×10^{-8}	5.178×10^{-8}	9.945×10^{-9}
M2GNIGnew	8.994×10^{-10}	8.380×10^{-11}	3.664×10^{-12}	4.587×10^{-13}	1.543×10^{-13}	2.185×10^{-14}
M2GNIG	3.394×10^{-9}	3.189×10^{-10}	1.173×10^{-11}	1.470×10^{-12}	4.525×10^{-13}	6.408×10^{-14}
M3GNIGnew	2.779×10^{-13}	2.048×10^{-14}	2.956×10^{-16}	3.014×10^{-17}	4.413×10^{-18}	5.122×10^{-19}
M3GNIG	3.601×10^{-12}	2.706×10^{-13}	1.104×10^{-15}	1.128×10^{-16}	1.766×10^{-17}	2.051×10^{-18}

5 Conclusions

The process used to factorize the characteristic functions involved allowed us to obtain near-exact distributions almost simultaneously for the independence and the sphericity test statistics and also to obtain simple expressions for the shape parameters of $\Phi_2^*(t)$ in (13) and of $\Phi_2^{**}(t)$ in (26), with the shape parameters for the near-exact distributions for the sphericity test statistic having much simpler expressions than the ones for the near-exact distributions in [19].

The near-exact distributions developed for the independence test statistic show a much better precision than one that is obtained with the Normal approximation of [20] while the new near-exact distributions developed for the sphericity test statistic are more accurate than the ones developed in [19] for larger values of p ($p \geq 10$) or even for smaller values of p as long as the sample size is large enough.

Acknowledgment

This research was financially supported by the Portuguese Foundation for Science and Technology, through the Center for Mathematics and its Applications (CMA) from the New University of Lisbon.

Appendix A

Cumulative distribution function for the GNIG distribution

The density and distribution functions for the GNIG distribution are given in [10]. Let

$$Z = Z_1 + Z_2$$

where $Z_2 \sim \Gamma(r, \lambda)$, with $\lambda > 0$ and r a positive non-integer and

$$Z_1 = \sum_{i=1}^{g} X_i\,, \qquad \text{with} \quad X_i \sim \Gamma(r_i, \lambda_i)\,, \quad \text{independent,}$$

where $r_1, \ldots, r_g$ are positive integers and $\lambda_1, \ldots, \lambda_g > 0$ are all different. The distribution of Z_1 is a GIG distribution of depth g ([8]), while the distribution of Z, if Z_1 and Z_2 are assumed independent, is a GNIG distribution of depth $g+1$. We will denote this by

$$Z \sim GNIG(r_1, \ldots, r_g, r;\, \lambda_1, \ldots, \lambda_g, \lambda)\,.$$

The cumulative distribution function of Z is given by

$$\begin{aligned} F_Z(z|r_1, \ldots, r_g, r;\, \lambda_1, \ldots, \lambda_g, \lambda) &= \lambda^r \frac{z^r}{\Gamma(r+1)}\, {}_1F_1(r, r+1, -\lambda z) \\ &-K\lambda^r \sum_{j=1}^{g} \mathrm{e}^{-\lambda_j z} \sum_{k=1}^{r_j} c^*_{j,k} \sum_{i=0}^{k-1} \frac{z^{r+i}\lambda_j^i}{\Gamma(r+1+i)}\, {}_1F_1(r, r+1+i, -(\lambda-\lambda_j)z) \\ &\qquad (z > 0) \end{aligned} \tag{42}$$

where

$$K = \prod_{j=1}^{g} \lambda_j^{r_j} \qquad \text{and} \qquad c^*_{j,k} = \frac{c_{j,k}}{\lambda_j^k}\Gamma(k)$$

with $c_{j,k}$ given by (11) through (13) in [8]. In the above expression ${}_1F_1(a, b; z)$ is the Kummer confluent hypergeometric function (see [1]). This function has usually very good convergence properties and is nowadays easily handled by a number of software packages.

References

[1] M. Abramowitz, I. A. Stegun, I.A., Handbook of Mathematical Functions, 9th print, Dover, New York, 1974.

[2] R. P. Alberto, C. A. Coelho, C. A., 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 (2007) 1612-1626.

[3] T. W. Anderson, An Introduction to Multivariate Statistical Analysis, 3rd ed., J. Wiley & Sons, New York, 2003.

[4] M. S. Bartlett, A note on multiplying factors for various χ^2 approximations, Journal of the Royal Statistical Society, Ser. B, 16 (1954) 296-298.

[5] A. Berry, The accuracy of the Gaussian approximation to the sum of independent variates, Transactions of the American Mathematical Society, 49 (1941) 122-136.

[6] G. E. P. Box, A general distribution theory for a class of likelihood criteria, Biometrika, 36 (1949) 317-346.

[7] C. A. Coelho, Generalized Canonical Analysis, Ph.D. Thesis, The University of Michigan, Ann Arbor, MI, 1992.

[8] C. A. Coelho, The Generalized Integer Gamma distribution - a basis for distributions in Multivariate Statistics, J. Multivariate analysis 64 (1998) 86-102.

[9] C. A. Coelho, The Generalized Integer Gamma distribution as an asymptotic replacement for the logBeta random variable - Applications, American Journal of Mathematical and Management Sciences, 23 (2003) 383-399.

[10] C. A. Coelho, The Generalized Near-Integer Gamma distribution: a basis for 'near-exact' approximations to the distributions of statistics which are the product of an odd number of independent Beta random variables, J. Multivariate Analysis, 89 (2004) 191-218.

[11] C. A. Coelho, The wrapped Gamma distribution and wrapped sums and linear combinations of independent Gamma and Laplace distributions, Journal of Statistical Theory and Practice, 1 (2007) 1-29.

[12] C. A. Coelho, R. P. Alberto, L. M. Grilo, A mixture of Generalized Integer Gamma distributions as the exact distribution of the product of an odd number of independent Beta random variables, Journal of Interdisciplinary Mathematics, 9 (2006) 229-248.

[13] C.-G. Esseen, Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law, Acta Mathematica, 77 (1945) 1-125.

[14] J. Gil-Pelaez, Note on the inversion theorem, Biometrika, 38 (1951) 481-482.

[15] L. M. Grilo, C. A. Coelho, Development and 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 (2007) 1560-1575.

[16] H.-K. Hwang, On convergence rates in the Central Limit Theorems for combinatorial structures, European Journal of Combinatorics, 19 (1998) 329-343.

[17] A. M. Kshirsagar, Multivariate Analysis, Marcel Dekker, Inc., New York, 1972.

[18] M. Loève, Probability Theory, Vol. I, 4th ed., Springer-Verlag, New York, 1977

[19] F. J. Marques, C. A. Coelho, Near-exact distributions for the sphericity likelihood ratio test statistic, Journal of Statistical Planning and Inference (2008) (in print).

[20] G. S. Mudholkar, M. C. Trivedi, C. T. Lin, An approximation to the distribution of the likelihood ratio statistic for testing the complete independence, Technometrics, 24 (1982) 139-143.

[21] R. J. Muirhead, Aspects of Multivariate Statistical Theory, J. Wiley & Sons, New York, 1986.