The exact and near-exact distributions of the likelihood ratio statistic for the block sphericity test

Filipe J. Marques and Carlos A. Coelho*

Using a suitable decomposition of the null hypothesis of the sphericity test for several blocks of variables, into a sequence of conditionally independent null hypotheses, we show that it is possible to obtain the expressions for the likelihood ratio test statistic, for its h-th null moment and for the characteristic function of its logarithm. The exact distribution of the logarithm of the likelihood ratio test statistic is obtained in the form of a sum of a Generalized Integer Gamma distribution with the sum of a given number of independent Logbeta distributions, taking the form of a single Generalized Integer Gamma distribution when each set of variables has two variables. The development of near-exact distributions arises, from the previous decomposition of the null hypothesis and from the consequent induced factorization of the characteristic function, as a natural and practical way to approximate the exact distribution of the test statistic. A measure based on the exact and approximating characteristic functions, which gives an upper bound on the distance between the corresponding distribution functions, is used to assess the quality of the near-exact distributions proposed and to compare them with an asymptotic approximation based on Box's method.

Keywords: asymptotic distributions; sphericity test; block; Generalized Integer Gamma distribution; mixtures.

1. Introduction

The block sphericity test is a very interesting generalization of the standard sphericity test with application in several statistical studies mainly when it is necessary to test if in a set of p variables, from a multivariate Normal population, we have independence among all of the variables and if in each of the k groups of variables, the i-th of which with p_i variables $(i = 1, \ldots, k)$, all the p_i variables have the same variance, that is, the k groups of variables are independent and in each group the sphericity assumption holds. One example of application is in testing the error structure in Linear and Nonlinear Mixed Models which are becoming more and more popular tools with applications in areas like Economics, Psychology, Biology, Agronomy, Genetics, Epidemiology and Medicine.

The normality assumptions under which the l.r.t. (likelihood ratio test) statistic is derived may seem too restrictive. However we should bear in mind the results in [2], [3], [1], [9] and [10], which show that l.r.t. statistic is the the same when the underlying distributions for the sample is multivariate elliptically contoured and left orthogonal-invariant distributions.

Universidade Nova de Lisboa, Faculdade de Ciências e Tecnologia, Departamento de Matemática
* *Correspondence to: Faculdade de Ciências e Tecnologia, Departamento de Matemática, Quinta da Torre, 2829-516 Caparica, Portugal. E-mail: cmacfct.unl.pt*

Contract/grant sponsor: This work was partially supported by 'Financiamento Base 2010 ISFL-1-297 from FCT/MCTES/PT.

We strongly believe that it is now a good time for a fresh and new look into this test. By showing that the null hypothesis of the test may be decomposed into a sequence of null hypotheses, and that this decomposition

- allows the obtention of the l.r.t. statistic and the expression for the h-th null moment of the l.r.t. statistic,
- induces a factorization on the c.f. (characteristic function) of the logarithm of the l.r.t. statistic which may be used as a basis to develop very precise approximations to the exact distribution of this statistic (see [7]),

we therefore aim at developing very precise near-exact distributions for the l.r.t. statistic for the block sphericity test. These are distributions that

- are built using a new technique which combines a decomposition of the exact c.f., most often a factorization, with keeping most of it unchanged and approximating asymptotically only a smaller part of it, often using the well known matching moments technique (see Section 4),
- may be used for statistics which have highly complex distributions, being easy to use when decompositions of the null hypothesis are considered,
- have very good asymptotic properties not only in terms of the sample size but also when other parameters involved, such as the number of variables, increase their values,
- may be implemented using simple computational resources and therefore may be used in statistical studies for a large number of practical problems.

2. Another way of looking into the null hypothesis

In this test we consider a sample of size N from a $p-$variate normal population, $N_p(\underline{\mu}\,,\,\Sigma)$ and we want to test the null hypothesis

$$H_0:\ \Sigma = diag\left(\sigma_i^2 I_{p_i},\ i=1,\dots,k\right) \tag{1}$$

with $1 \le k \le p = p_1 + p_2 + \cdots + p_k$.

The null hypothesis in (1) may be decomposed into a sequence of two null hypotheses,

$$H_0 = H_{0b|0a} \circ H_{0a} \tag{2}$$

where

$$H_{0a} \quad : \quad \Sigma_{ij} = 0 \ \text{ for } \ i \neq j \quad (i,j = 1,\dots,k)\,, \tag{3}$$

is the null hypothesis for testing the independence of k groups of r.v.'s (random variables) and

$$\begin{aligned} H_{0b|0a} \quad : \quad & \Sigma_{ii} = \sigma_i^2\, I_{p_i}\ ,\ \ i=1,\dots,k, \qquad (k \text{ independent sphericity tests}) \\ & \text{assuming that } H_{0a} \text{ is true} \end{aligned} \tag{4}$$

is the null hypothesis for testing sphericity in each of the k groups of r.v.'s.

Using this decomposition we will be able to derive, in a very simple way, the l.r.t. statistic to test the null hypothesis in (1) as product of the l.r.t. statistics used to test the null hypotheses in (3) and (4). We have this way, for the modified l.r.t. statistic,

$$\lambda^* \quad = \quad \underbrace{\left(\frac{|A|^{n/2}}{\prod_{i=1}^{k}|A_{ii}|^{n/2}}\right)}_{\lambda_a^*} \times \underbrace{\left(\prod_{i=1}^{k}\frac{|A_{ii}|^{\frac{n}{2}}}{(tr\ A_{ii})^{\frac{np_i}{2}}}\,(p_i)^{np_i/2}\right)}_{\lambda_{b|a}^*} \tag{5}$$

$$= \quad \prod_{i=1}^{k}\frac{|A|^{n/(2k)}}{(tr\ A_{ii})^{\frac{np_i}{2}}}\,(p_i)^{np_i/2} \tag{6}$$

where λ_a^* and $\lambda_{b|a}^*$ are modified l.r.t. statistics used to test the null hypotheses in (3) and (4) (see [1]), A is the $p \times p$ matrix of corrected sums of squares and products formed from the sample or alternatively the sample variance-covariance matrix or the MLE of Σ, $n = N - 1$ is the number of degrees of freedom of the Wishart distribution of A and A_{ii} denotes the i-th diagonal block of order p_i of A.

Given the independence of the statistics λ_a^* and $\lambda_{b|a}^*$ in (5), under H_0, we may use the expression for the h-th null moment of these statistics (see [1]) to derive the expression for the h-th null moment of λ^* as

$$E\left[(\lambda^*)^h\right] = \underbrace{\frac{\Gamma_p(\frac{1}{2}n + \frac{1}{2}hn)}{\Gamma_p(\frac{1}{2}n)} \prod_{i=1}^{k} \frac{\Gamma_{p_i}(\frac{1}{2}n)}{\Gamma_{p_i}(\frac{1}{2}n + \frac{1}{2}hn)}}_{E[(\lambda_a^*)^h]} \times \underbrace{\prod_{i=1}^{k} p_i^{hp_i n/2} \frac{\Gamma\left(\frac{np_i}{2}\right)}{\Gamma\left(\frac{np_i}{2} + \frac{p_i hn}{2}\right)} \frac{\Gamma_{p_i}\left(\frac{n}{2} + \frac{nh}{2}\right)}{\Gamma_{p_i}\left(\frac{n}{2}\right)}}_{E\left[\left(\lambda_{b|a}^*\right)^h\right]} \tag{7}$$

$$= \frac{\Gamma_p(\frac{1}{2}n + \frac{1}{2}hn)}{\Gamma_p(\frac{1}{2}n)} \prod_{i=1}^{k} \left\{ p_i^{hp_i n/2} \frac{\Gamma\left(\frac{np_i}{2}\right)}{\Gamma\left(\frac{np_i}{2} + \frac{p_i hn}{2}\right)} \right\} \tag{8}$$

where

$$\Gamma_p(z) = \pi^{p(p-1)/4} \prod_{j=1}^{p} \Gamma\left(z - \frac{j-1}{2}\right)$$

is the multivariate Gamma function.

3. The exact distribution of $W = -\log \lambda^*$ and λ^*

In the complex case, the exact null distribution of the l.r.t. statistic was obtained by [16] as a Beta series or as a chi-square series and by [15] based on the Meijer's G-function which is almost impossible to use in practical terms and as a series of Psi and Zeta functions. However the distribution of the l.r.t. statistic in the complex case is indeed and surprisingly a bit simpler to handle. For the real case [14] obtained an asymptotic distribution for the multisample block sphericity l.r.t. statistic, however although giving the expressions for the h-th moment he does not present any expressions for the exact p.d.f. or c.d.f..

In this section we will show that the exact distribution λ^* in (6) is the same as the distribution of

$$\left\{ \prod_{j=1}^{p-1} e^{-Z_j} \right\} \left\{ \prod_{j=1}^{p+m^*} (Y_j)^{n/2} \right\} \tag{9}$$

where Z_j and Y_j are all independent, with

$$Z_j \sim \Gamma\left(r_j, \frac{n-j}{n}\right)$$

and Y_j with Beta distributions and where $p = p_1 + p_2 + \cdots + p_k$, r_j given in (21) and

$$m^* = \left\lfloor \frac{\ell}{2} \right\rfloor \tag{10}$$

where ℓ is the number of odd p_i, that is the number of sets with an odd number of variables. More precisely, the second product in (9) should indeed be writen as

$$\left\{ \prod_{j=1}^{m^*} \left(Y_j^*\right)^{n/2} \right\} \left\{ \prod_{i=1}^{k} \left\{ \prod_{j=1}^{p_i - k_i^*} \left(Y_{ij}^{**}\right)^{n/2} \right\} \left\{ \prod_{j=p_i-k_i^*+1}^{p_i} \left(Y_{ij}^{***}\right)^{n/2} \right\} \right\} \tag{11}$$

where Y_j^*, Y_{ij}^{**} and Y_{ij}^{***} are all independent with

$$\begin{aligned}
Y_j^* &\sim Beta\left(\tfrac{n-1}{2}, \tfrac{1}{2}\right), \quad j=1,\dots,m^* \\
Y_{ij}^{**} &\sim Beta\left(\tfrac{n}{2}, \tfrac{j-1}{p_i}\right), \quad i=1,\dots,k \;;\; j=1,\dots,p_i-k_i^* \\
Y_{ij}^{***} &\sim Beta\left(\tfrac{n+1}{2}, \tfrac{j-1}{p_i}-\tfrac{1}{2}\right), \quad i=1,\dots,k \;;\; j=p_i-k_i^*+1,\dots,p_i\,.
\end{aligned}$$

From (7) above we may obtain a factorization of the c.f. of the r.v. $W=-\log\lambda^*$,

$$\begin{aligned}
\Phi_W(t) &= E\left[e^{it(-\log\lambda^*)}\right] = E\left[(\lambda^*)^{-it}\right] \\
&= \underbrace{\frac{\Gamma_p\left(\frac{1}{2}n-\frac{1}{2}itn\right)}{\Gamma_p\left(\frac{1}{2}n\right)}\prod_{i=1}^{k}\frac{\Gamma_{p_i}\left(\frac{1}{2}n\right)}{\Gamma_{p_i}\left(\frac{1}{2}n-\frac{1}{2}itn\right)}}_{\Phi_{-\log\lambda_a^*}(t)} \times \underbrace{\prod_{i=1}^{k} p_i^{-itp_in/2}\frac{\Gamma\left(\frac{np_i}{2}\right)}{\Gamma\left(\frac{np_i}{2}-\frac{p_iitn}{2}\right)}\frac{\Gamma_{p_i}\left(\frac{n}{2}-\frac{nit}{2}\right)}{\Gamma_{p_i}\left(\frac{n}{2}\right)}}_{\Phi_{-\log\lambda_{b|a}^*}(t)}\,,
\end{aligned} \tag{12}$$

where $\Phi_{-\log\lambda_a^*}(t)$ is the c.f. of $-\log\lambda_a^*$ and $\Phi_{-\log\lambda_{b|a}^*}(t)$ is the c.f. of $-\log\lambda_{b|a}^*$. In the next Subsection we will introduce new factorizations of the above c.f.'s which will be fundamental for the development of near-exact distributions for the l.r.t. statistic.

3.1. The c.f.'s of $-\log\lambda_a^*$ and of $-\log\lambda_{b|a}^*$ – the general case

According to [6] and [13], the c.f. $\Phi_{-\log\lambda_a^*}(t)$ in (12) may be written as

$$\Phi_{-\log\lambda_a^*}(t) = \underbrace{\left\{\prod_{j=2}^{p-1}\left(\frac{n-j}{n}\right)^{z_j}\left(\frac{n-j}{n}-it\right)^{-z_j}\right\}}_{\Phi_{1,1}(t)} \times \underbrace{\left\{\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}-\frac{1}{2}-\frac{n}{2}it\right)}{\Gamma\left(\frac{n}{2}-\frac{n}{2}it\right)\Gamma\left(\frac{n}{2}-\frac{1}{2}\right)}\right\}^{m^*}}_{\Phi_{1,2}(t)} \tag{13}$$

with m^* given by (10) and

$$z_j = \begin{cases} 0 & j=2 \\ h_1-m^* & j=3 \\ h_2+m^* & j=4 \\ z_{j-2}+h_{j-2} & j=5,\dots,p \end{cases} \tag{14}$$

with

$$h_j = (\#\text{ of } p_i\ (i=1,\dots,k)\ge j)-1\,, \quad j=1,\dots,p-2\,. \tag{15}$$

The c.f. $\Phi_{1,1}(t)$ in (13) corresponds to the sum of $p-2$ independent r.v.'s with Gamma distribution and integer shape parameters z_j given by (14), that is a GIG (Generalized Integer Gamma) distribution of depth $p-2$ (see [5] and Appendix A), while $\Phi_{1,2}(t)$ is the c.f. of the sum of m^* independent r.v.'s with Logbeta distribution with parameters $(n-1)/2$ and $1/2$, multiplied by $n/2$.

Given the independence, under H_{0a} in (3), of the k sphericity tests in (4), the c.f. of $-\log\lambda_{b|a}^*$ for $\lambda_{b|a}^*$ in (5) may be obtained as the product of k c.f.'s of the negative logarithm of the l.r.t. statistics to test each of the k sphericity hypotheses in (4). Near-exact distributions for the l.r.t. statistic to test sphericity were developed by [12] and [8]. In [8], the authors show that for a group of p_i variables the c.f. of the negative logarithm of the l.r.t. statistic is given by

$$\begin{aligned}
\Phi_i^*(t) &= \underbrace{\prod_{j=1}^{p_i-1}\left(\frac{n-j}{n}\right)^{\left\lfloor\frac{p_i-j+1}{2}\right\rfloor}\left(\frac{n-j}{n}-it\right)^{-\left\lfloor\frac{p_i-j+1}{2}\right\rfloor}}_{\Phi_{2,1}^*(t)} \\
&\quad \times \underbrace{\prod_{j=p_i-k_i^*+1}^{p_i}\frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p_i}\right)\Gamma\left(\frac{n+1}{2}-\frac{n}{2}it\right)}{\Gamma\left(\frac{n+1}{2}\right)\Gamma\left(\frac{n}{2}+\frac{j-1}{p_i}-\frac{n}{2}it\right)}\prod_{j=1}^{p_i-k_i^*}\frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p_i}\right)\Gamma\left(\frac{n}{2}-\frac{n}{2}it\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}+\frac{j-1}{p_i}-\frac{n}{2}it\right)}}_{\Phi_{2,2}^*(t)}
\end{aligned} \tag{16}$$

where $k_i^*=\lfloor p_i/2\rfloor$. In (16) above, $\Phi_{2,1}^*(t)$ is the c.f. of a GIG distribution of depth p_i-1, with rate parameters $(n-j)/n$ and shape parameters $\left\lfloor\frac{p_i-j+1}{2}\right\rfloor$ $(j=1,\dots,p_i-1)$ and $\Phi^*_{2,2}(t)$ is the c.f. of the sum of p_i independent Logbeta r.v.'s, k_i^* of which with parameters $(n+1)/2$ and $(j-1)/p_i-1/2$ $(j=p_i-k_i^*+1,\dots,p_i)$ and the remaining $p_i-k_i^*$ with parameters $n/2$ and $(j-1)/p_i$ $(j=1,\dots,p_i-k_i^*)$. The c.f. of $-\log\lambda^*_{b|a}$ may thus be given by

$$\begin{aligned}
\Phi_{-\log\lambda^*_{b|a}}(t) &= \prod_{i=1}^{k}\Phi_i^*(t)\\
&= \left\{\prod_{i=1}^{k}\prod_{j=1}^{p_i-1}\left(\frac{n-j}{n}\right)^{\left\lfloor\frac{p_i-j+1}{2}\right\rfloor}\left(\frac{n-j}{n}-\mathrm{i}t\right)^{-\left\lfloor\frac{p_i-j+1}{2}\right\rfloor}\right\}\\
&\quad\times\left\{\prod_{i=1}^{k}\prod_{j=p_i-k_i^*+1}^{p_i}\frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p_i}\right)\Gamma\left(\frac{n+1}{2}-\frac{n}{2}\mathrm{i}t\right)}{\Gamma\left(\frac{n+1}{2}\right)\Gamma\left(\frac{n}{2}+\frac{j-1}{p_i}-\frac{n}{2}\mathrm{i}t\right)}\prod_{j=1}^{p_i-k_i^*}\frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p_i}\right)\Gamma\left(\frac{n}{2}-\frac{n}{2}\mathrm{i}t\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}+\frac{j-1}{p_i}-\frac{n}{2}\mathrm{i}t\right)}\right\}\\
&= \underbrace{\left\{\prod_{j=1}^{p-1}\left(\frac{n-j}{n}\right)^{r_j^*}\left(\frac{n-j}{n}-\mathrm{i}t\right)^{-r_j^*}\right\}}_{\Phi_{2,1}(t)}\\
&\quad\times\underbrace{\left\{\prod_{i=1}^{k}\prod_{j=p_i-k_i^*+1}^{p_i}\frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p_i}\right)\Gamma\left(\frac{n+1}{2}-\frac{n}{2}\mathrm{i}t\right)}{\Gamma\left(\frac{n+1}{2}\right)\Gamma\left(\frac{n}{2}+\frac{j-1}{p_i}-\frac{n}{2}\mathrm{i}t\right)}\prod_{j=1}^{p_i-k_i^*}\frac{\Gamma\left(\frac{n}{2}+\frac{j-1}{p_i}\right)\Gamma\left(\frac{n}{2}-\frac{n}{2}\mathrm{i}t\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n}{2}+\frac{j-1}{p_i}-\frac{n}{2}\mathrm{i}t\right)}\right\}}_{\Phi_{2,2}(t)}
\end{aligned}\tag{17}$$

with r_j^* given by

$$r_j^*=\begin{cases}\sum_{i=1}^{k} r_{i,j}^{**} & j=1\dots,p_{max}-1\\ 0 & j=p_{max},\dots,p-1\end{cases}\tag{18}$$

where, for $i=1,\dots,k$

$$r_{i,j}^{**}=\begin{cases}\left\lfloor\frac{p_i-j+1}{2}\right\rfloor & j=1,\dots,p_i\\ 0 & j=p_i+1,\dots,p_{\max}-1\end{cases}\tag{19}$$

with $p_{\max}=\max\{p_1,\dots,p_k\}$.

The c.f.'s in (13) and (17) show that both the exact distribution of $-\log\lambda^*_a$ and $-\log\lambda^*_{b|a}$ may be seen as the sum of independent Gamma r.v.'s with shape parameters integers with the sum of independent Logbeta r.v.'s multiplied by $n/2$.

We have thus the c.f. of $W=-\log\lambda^*=-(\log\lambda^*_a+\log\lambda^*_{b|a})$ given by the following theorem.

Theorem 3.1
The c.f. of $W=-\log\lambda^*$ may be written as

$$\Phi_W(t) = \underbrace{\left\{\prod_{j=1}^{p-1}\left(\frac{n-j}{n}\right)^{r_j}\left(\frac{n-j}{n}-\mathrm{i}t\right)^{-r_j}\right\}}_{\Phi_{W_1}(t)}\times\underbrace{\Phi_{1,2}(t)\ \times\ \Phi_{2,2}(t)}_{\Phi_{W_2}(t)}\tag{20}$$

with $\Phi_{1,2}(t)$ and $\Phi_{2,2}(t)$ given by (13) and (17) respectively, and where r_j are given by

$$r_j=\begin{cases}r_j^* & j=1\\ r_j^*+z_{p-j} & j=2,\dots,p-1\,,\end{cases}\tag{21}$$

with z_j given by (14) and r_j^* given by (18).

Proof
Since $W=-\log\lambda^*=-(\log\lambda^*_a+\log\lambda^*_{b|a})$ we have $\Phi_W(t)=\Phi_{-\log\lambda^*_a}(t)\times\Phi_{-\log\lambda^*_{b|a}}(t)$ with $\Phi_{-\log\lambda^*_a}(t)$ given by (13) and $\Phi_{-\log\lambda^*_{b|a}}(t)$ given by (17). $\square$

From Theorem 3.1 we may say that the exact distribution of λ^* may be seen as the distribution of the product of two independent r.v.'s, e^{-W_1} and e^{-W_2}, where W_1 has the distribution of the sum of $p-1$ independent Gamma r.v.'s with shape integer parameters and W_2 has the distribution of the sum of $p+m^*$ independent Logbeta r.v.'s multiplied by $n/2$ (see (9)).

3.2. Particular cases of interest

3.2.1. When at most one of the p_i is odd

In the case where at most one of the p_i is odd, m^* in (10) is equal to zero and so the c.f. W is given by

$$\Phi_W(t) \quad = \quad \underbrace{\left\{\prod_{j=1}^{p-1}\left(\frac{n-j}{n}\right)^{r_j}\left(\frac{n-j}{n}-\mathrm{i}t\right)^{-r_j}\right\}}_{\Phi_{W_1}(t)} \times \underbrace{\Phi_{2,2}(t)}_{\Phi_{W_2}(t)} \tag{22}$$

with $\Phi_{2,2}(t)$ given by (17).

3.2.2. When all $p_i = 2$

For $p_i = 2$, $i = 1, \ldots, k$, we have $\quad \Phi_{2,2}(t) = 1\,,\ \forall\, t \in \mathbb{R} \quad$ and also m^* in (10) equal to zero, so that $\Phi_{1,2}(t)$ in (13) vanishes and thus the c.f. of W is, in this case, given by

$$\Phi_W(t) = \prod_{j=1}^{2k-1}\left(\frac{n-j}{n}\right)^{r_j}\left(\frac{n-j}{n}-\mathrm{i}t\right)^{-r_j}\,,$$

with

$$r_j = \begin{cases} k & j=1 \\ k-1-\left\lfloor\frac{j-2}{2}\right\rfloor & j=2,\ldots,2k-1\,. \end{cases} \tag{23}$$

In this case, we have a single GIG distribution as the exact distribution of W and from the expression of the p.d.f (probability density function) and c.d.f. (cumulative distribution function) of $W = -\log\lambda^*$ we may obtain, by simple transformation, the p.d.f and c.d.f. of λ^*. Thus, a distribution with c.d.f. (using the notation in Appendix A)

$$1 - F^{GIG}\Big(-\log x|r_1,\ldots,r_{2k-1};\frac{n-1}{n},\ldots,\frac{n-(2k-1)}{n}\Big)\,,$$

and p.d.f

$$f^{GIG}\Big(-\log x|r_1,\ldots,r_{2k-1};\frac{n-1}{n},\ldots,\frac{n-(2k-1)}{n}\Big)\frac{1}{x}\,,$$

where the shape parameters r_j, $j = 1, \ldots, 2k-1$ are given by (23), and $0 < x < 1$ represents the running value of the statistic $\lambda^* = e^{-W}$, is the exact distribution for λ^*, when $p_i = 2$ for all $i = 1, \ldots, k$.

In the Figure 1 we present examples of p.d.f.'s and c.d.f.'s considering different sample sizes and for $p_i = 2$ for all $i = 1, 2, 3$.

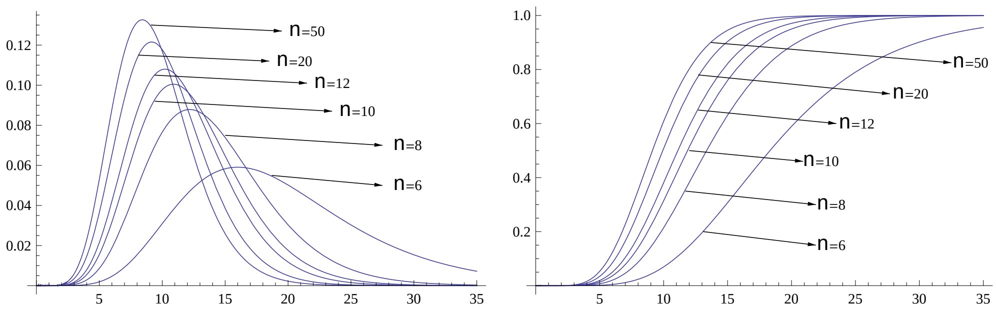

Figure 1. Plots of the p.d.f's and c.d.f's for GIG distributions when $p_i = 2$ for all $i = 1, 2, 3$

In the next Section we will obtain near-exact distributions for λ^* that have very good asymptotic properties not only in terms of sample size but also in terms of the number of variables and the number of sets of variables involved.

4. Near-exact distributions for W and λ^*

The near-exact distributions we will develop in this paper will have c.f.'s of the form

$$\Phi_{W^*}(t) = \underbrace{\Phi_{W_1}(t)}_{\text{GIG distribution}} \times \Phi_{W_2^*}(t) \tag{24}$$

where $\Phi_{W_1}(t)$ is the same as in (20) above and $\Phi_{W_2^*}(t)$ is the c.f. that we will use to approximate the c.f. $\Phi_{W_2}(t)$ in (20).

We propose that $\Phi_{W_2^*}(t)$ may be either the c.f. of a single Gamma distribution or of a mixture of two or three Gamma distributions, depending on the number of exact moments we want to match. The c.f. $\Phi_{W_2^*}(t)$ will indeed have, accordingly, the same 2, 4 or 6 first derivatives (with respect to t at $t = 0$) as the part of the exact c.f. of W that will be replaced, that is, $\Phi_{W_2}(t)$ in (20). In other words, we will have

$$\frac{d^j}{dt^j}\,\Phi_{W_2^*}(t)\bigg|_{t=0} = \frac{d^j}{dt^j}\,\Phi_{W_2}(t)\bigg|_{t=0}, \qquad j = 1, \ldots, h \tag{25}$$

for $h = 2, 4$ or 6, according to the case of $\Phi_{W_2^*}(t)$ being the c.f. of a single Gamma distribution, or the c.f. of a mixture of 2 or 3 Gamma distributions with the same rate parameter, that is,

$$\Phi_{W_2^*}(t) = \sum_{k=1}^{h/2} p_k\,\lambda^{s_k}\,(\lambda - \mathrm{i}t)^{-s_k}\,, \tag{26}$$

with weights $p_k > 0$ ($k = 1, \ldots, h/2$ with $h = 2, 4$ or 6) and $\sum_{k=1}^{h/2} p_k = 1$.

This way we will be able to write the near-exact c.f. of the negative logarithm of the l.r.t. statistic for the one sample block sphericity test in the form in (24) where $\Phi_{W_2^*}(t)$ is either the c.f. of a Gamma distribution or the c.f. of a mixture of 2 or 3 Gamma distributions, being thus the near-exact distributions obtained in this way, correspondingly a GNIG (Generalized Near-Integer Gamma) distribution (see [6] and Appendix A) of depth p, or a mixture of two or three GNIG distributions of the same depth, which have very manageable expressions, allowing this way for an easy computation of very accurate near-exact quantiles. From the near-exact distributions obtained for W, a GNIG distribution or a mixture of two or three GNIG distributions of depth p (for $h = 2, 4$ or 6), by a simple transformation we may obtain the expressions for the near-exact distributions for λ^* with p.d.f.'s and c.d.f.'s given in the next theorem (using the notation in Appendix A)

Theorem 4.1
The near-exact distributions for λ^* are either a exponential GNIG distribution or a mixture of two or three exponential GNIG distributions of depth p and for $h = 2, 4$ or 6 with p.d.f. for λ^* given by

$$\sum_{\nu=1}^{h/2} p_\nu\, f^{\,GNIG}\left(-\log w \middle| r_1, \ldots, r_{p-1}, s_\nu\,;\, \frac{n-1}{n}, \ldots, \frac{n-p+1}{n}, \lambda\,;\, p\right)\frac{1}{w}$$

and c.d.f. for λ^* given by

$$1 - \sum_{\nu=1}^{h/2} p_\nu\, F^{\,GNIG}\left(-\log w \middle| r_1, \ldots, r_{p-1}, s_\nu\,;\, \frac{n-1}{n}, \ldots, \frac{n-p+1}{n}, \lambda\,;\, p\right)$$

with r_j given by (21), and where for $h = 2$

$$\lambda = \frac{m_1}{m_2 - m_1^2} \qquad \text{and} \qquad s_1 = \frac{m_1^2}{m_2 - m_1^2} \tag{27}$$

with

$$m_j = \mathrm{i}^{-j} \frac{\partial^j}{\partial t^j} \Phi_{W_2}(t)\bigg|_{t=0}, \qquad j = 1, 2,$$

and for $h = 4$ or $h = 6$ (according to the case of $\Phi_{W_2^*}(t)$ being the c.f. of a mixture of 2 or 3 Gamma distributions with the same rate parameter) the values of p_ν, s_ν and λ are obtained from the numerical solution of the system of equations in (25), that is

$$\frac{d^j}{dt^j} \Phi_{W_2^*}(t)\bigg|_{t=0} = \frac{d^j}{dt^j} \Phi_{W_2}(t)\bigg|_{t=0}, \qquad j = 1, \ldots, h$$

with

$$p_{h/2} = 1 - \sum_{k=1}^{h/2-1} p_k \, .$$

Proof
See Appendix B. □

Some authors use different versions of this statistic. For example, instead of the l.r.t. statistic use could have used $(\lambda^*)^{N/n}$. However, we may note that we can easily obtain both the distribution and quantiles of different powers of λ^* from the ones for λ^*.

5. Numerical studies

In order to evaluate the quality of the near-exact distributions proposed in this work we will use a measure of proximity between c.f.'s which is also a measure of proximity between c.d.f.'s. This measure is,

$$\Delta = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \frac{\Phi_W(t) - \Phi_{W^*}(t)}{t} \right| \mathrm{d}t \, , \tag{28}$$

where $\Phi_W(t)$ represents the exact c.f. of the negative logarithm of the modified l.r.t. statistic and $\Phi_{W^*}(t)$ represents an approximate c.f. for the same statistic. We have that, taking S for the support of W,

$$\max_{y \in S} |F_W(y) - F_{W^*}(y)| \leq \Delta \, , \tag{29}$$

where $F_W(y)$ represents the exact c.d.f. of W and $F_{W^*}(y)$ represents the c.d.f. corresponding to W^*. For more details on the measure see for example [12].

We intend to assess the performance of the near-exact approximations developed in this paper by computing the values of the measure Δ between the exact distribution of $W = -\log \lambda^*$ and the three proposed near-exact approximations. In the calculations we use the exact c.f. in (12) and the near-exact c.f.'s corresponding to the near-exact distributions in Theorem 4.1 and given by (24) and (26) for $h = 2, 4$ and 6. We will denote respectively by $GNIG$, $M2GNIG$ and $M3GNIG$ the near-exact distributions corresponding to the c.f.'s in (24) and (26) with $h = 2, 4$ and 6. In order to better assess the quality of the new near-exact distributions proposed in this paper, we will also use the asymptotic approximation presented by [16] based on [4] method, that we will denote by Box.

We may observe from Tables 1, 2 and 3 that the near-exact approximations present in every case considered much better results than the approximation given in [16] at the same time the near-exact distributions reveal good asymptotic properties for increasing sample sizes (see Table 2) and also for increasing number of variables (see Tables 1 and 3).

Table 1. Values of Δ for the approximating distributions for $W = -\log\lambda^*$, for increasing number of variables (p), with $n = p + 2$

				Measure Δ			
p	p_i	k	n	$GNIG$	$M2GNIG$	$M3GNIG$	Box
10	{3,3,4}	3	12	7.9×10^{-7}	4.2×10^{-11}	4.4×10^{-13}	1.5×10^{-1}
12	{5,5,2}	3	14	3.9×10^{-7}	1.7×10^{-11}	1.5×10^{-14}	2.0×10^{-1}
14	{5,5,4}	3	16	2.2×10^{-7}	6.7×10^{-12}	2.2×10^{-15}	2.6×10^{-1}
20	{7,7,6}	3	22	5.3×10^{-8}	6.9×10^{-13}	5.4×10^{-17}	4.5×10^{-1}
50	{15,15,20}	3	52	1.0×10^{-9}	9.5×10^{-16}	2.4×10^{-21}	1.4×10^{0}

Table 2. Values of Δ for the approximating distributions for $W = -\log\lambda^*$, for increasing sample size (n)

				Measure Δ			
p	p_i	k	n	$GNIG$	$M2GNIG$	$M3GNIG$	Box
10	{3,3,4}	3	12	7.9×10^{-7}	4.2×10^{-11}	4.4×10^{-13}	1.5×10^{-1}
10	{3,3,4}	3	20	5.8×10^{-7}	4.5×10^{-11}	2.6×10^{-13}	4.2×10^{-2}
10	{3,3,4}	3	50	1.3×10^{-7}	6.4×10^{-12}	5.1×10^{-15}	1.3×10^{-2}
10	{3,3,4}	3	100	3.5×10^{-8}	9.7×10^{-13}	3.5×10^{-16}	6.1×10^{-3}

Table 3. Values of Δ for the approximating distributions for $W = -\log\lambda^*$, for increasing k and p

				Measure Δ			
p	p_i	k	n	$GNIG$	$M2GNIG$	$M3GNIG$	Box
9	{4,5}	2	11	1.3×10^{-6}	5.1×10^{-10}	9.1×10^{-14}	1.3×10^{-1}
10	{4,4,2}	3	12	6.8×10^{-7}	5.0×10^{-11}	1.2×10^{-14}	1.5×10^{-1}
15	{5,4,4,2}	4	17	1.1×10^{-7}	5.7×10^{-12}	7.9×10^{-16}	2.9×10^{-1}
20	{6,4,4,4,2}	5	22	3.4×10^{-8}	4.2×10^{-13}	3.8×10^{-17}	4.5×10^{-1}
50	{30,6,4,4,4,2}	6	52	1.5×10^{-10}	1.8×10^{-16}	5.7×10^{-21}	1.4×10^{0}

6. Conclusions

The decomposition of the null hypothesis for testing block sphericity in a sequence of null hypotheses, is a practical and effective tool not only to derive the l.r.t. statistic but also the expression of the h-th null moment. But, yet more important, the induced factorization on the c.f. of the logarithm of the test statistic can be used to obtain very well fit near-exact distributions. Numerical studies show that these near-exact approximations are very accurate and at the same time reveal very good asymptotic properties both for increasing sample sizes and number of variables.

7. Acknowledgment

This research was financially supported by the Portuguese Foundation for Science and Technology (FCT), through the Center for Mathematics and its Applications (CMA) from Universidade Nova de Lisboa, through grant 'Financiamento Base 2010 ISFL-1-297 from FCT/MCTES/PT.

A. Appendix

The GIG and GNIG distributions

We will use this Appendix to establish some notation concerning distributions used in the paper, as well as to give the expressions for the p.d.f.'s (probability density functions) and c.d.f.'s (cumulative distribution functions) of the GIG (Generalized Integer Gamma) and GNIG (Generalized Near-Integer Gamma) distributions.

We will say that the r.v. X has a Gamma distribution with rate parameter $\lambda > 0$ and shape parameter $r > 0$, if its p.d.f. may be written as

$$f_X(x) = \frac{\lambda^r}{\Gamma(r)}\, e^{-\lambda x}\, x^{r-1}\,, \qquad (x > 0)$$

and we will denote this fact by

$$X \sim \Gamma(r, \lambda)\,.$$

Let

$$X_j \;\sim\; \Gamma(r_j, \lambda_j) \qquad j = 1, \ldots, p$$

be p independent r.v.'s with Gamma distributions with shape parameters $r_j \in \mathbb{N}$ and rate parameters $\lambda_j > 0$, with $\lambda_j \neq \lambda_{j'}$, for all $j, j' \in \{1, \ldots, p\}$. We will say that then the r.v.

$$Y \;=\; \sum_{j=1}^{p} X_j$$

has a GIG distribution of depth p, with shape parameters r_j and rate parameters λ_j, ($j = 1, \ldots, p$), and we will denote this fact by

$$Y \;\sim\; GIG(r_j, \lambda_j;\, p)\,.$$

The p.d.f. and c.d.f. (cumulative distribution function) of Y are respectively given by (Coelho, 1998)

$$f^{GIG}(y|r_1, \ldots, r_p; \lambda_1, \ldots, \lambda_p; p) \;=\; K \sum_{j=1}^{p} P_j(y)\, e^{-\lambda_j y}\,, \qquad (y > 0) \tag{30}$$

and

$$F^{GIG}(y|r_1, \ldots, r_j; \lambda_1, \ldots, \lambda_p; p) \;=\; 1 - K \sum_{j=1}^{p} P_j^*(y)\, e^{-\lambda_j y}\,, \qquad (y > 0) \tag{31}$$

where

$$K = \prod_{j=1}^{p} \lambda_j^{r_j}\,, \qquad P_j(y) = \sum_{k=1}^{r_j} c_{j,k}\, y^{k-1} \tag{32}$$

and

$$P_j^*(y) = \sum_{k=1}^{r_j} c_{j,k}\, (k-1)! \sum_{i=0}^{k-1} \frac{y^i}{i!\, \lambda_j^{k-i}}$$

with

$$c_{j,r_j} = \frac{1}{(r_j - 1)!} \prod_{\substack{i=1\\ i\neq j}}^{p} (\lambda_i - \lambda_j)^{-r_i}\,, \qquad j = 1, \ldots, p\,, \tag{33}$$

and

$$c_{j,r_j-k} = \frac{1}{k} \sum_{i=1}^{k} \frac{(r_j - k + i - 1)!}{(r_j - k - 1)!}\, R(i, j, p)\, c_{j,r_j-(k-i)}\,, \qquad (k = 1, \ldots, r_j - 1;\; j = 1, \ldots, p) \tag{34}$$

where

$$R(i, j, p) = \sum_{\substack{k=1\\ k\neq j}}^{p} r_k\, (\lambda_j - \lambda_k)^{-i} \qquad (i = 1, \ldots, r_j - 1)\,. \tag{35}$$

The GNIG (Generalized Near-Integer Gamma) distribution of depth $p+1$ (Coelho, 2004) is the distribution of the r.v.

$$Z = Y_1 + Y_2$$

where Y_1 and Y_2 are independent, Y_1 having a GIG distribution of depth p and Y_2 with a Gamma distribution with a non-integer shape parameter r and a rate parameter $\lambda \neq \lambda_j$ $(j = 1, \ldots, p)$. The p.d.f. (probability density function) of Z is given by

$$\begin{aligned} & f^{GNIG}(z|r_1, \ldots, r_p, r;\ \lambda_1, \ldots, \lambda_p, \lambda; p+1) = \\ & \qquad K\lambda^r \sum_{j=1}^{p} e^{-\lambda_j z} \sum_{k=1}^{r_j} \left\{ c_{j,k} \frac{\Gamma(k)}{\Gamma(k+r)} z^{k+r-1} {}_1F_1(r, k+r, -(\lambda-\lambda_j)z) \right\}, \qquad (z>0) \end{aligned} \tag{36}$$

and the c.d.f. (cumulative distribution function) given by

$$\begin{aligned} F^{GNIG}(z|r_1, \ldots, r_p, r;\ \lambda_1, \ldots, \lambda_p, \lambda; p+1) = & \frac{\lambda^r z^r}{\Gamma(r+1)} {}_1F_1(r, r+1, -\lambda z) \\ & -K\lambda^r \sum_{j=1}^{p} e^{-\lambda_j z} \sum_{k=1}^{r_j} c^*_{j,k} \sum_{i=0}^{k-1} \frac{z^{r+i}\lambda_j^i}{\Gamma(r+1+i)} {}_1F_1(r, r+1+i, -(\lambda-\lambda_j)z) \qquad (z>0) \end{aligned} \tag{37}$$

where

$$c^*_{j,k} = \frac{c_{j,k}}{\lambda_j^k}\Gamma(k)$$

with $c_{j,k}$ given by (33) through (35) above. In the above expressions ${}_1F_1(a, b; z)$ is the Kummer confluent hypergeometric function. This function typically has very good convergence properties and is nowadays easily handled by a number of software packages.

B. Appendix

Proof of Theorem 4.1

Proof

In this proof we will consider only the case of $h = 6$, since the cases $h = 2$ and $h = 4$ are derived in a similar way.

If in the c.f. of W in (20) we replace $\Phi_{W_2}(t)$ by

$$\Phi^*_{W_2}(t) = \sum_{k=1}^{3} p_k\, \lambda^{s_k} (\lambda - \mathrm{i}t)^{-s_k},$$

we obtain

$$\begin{aligned} \Phi_W(t) & \approx \Phi_{W_1} \times \underbrace{\sum_{k=1}^{3} p_k\, \lambda^{s_k} (\lambda - \mathrm{i}t)^{-s_k}}_{\Phi^*_{W_3}(t)} \\ & \approx \sum_{k=1}^{3} p_k \underbrace{\underbrace{\Phi_{W_1}}_{\text{GIG distribution}} \times \underbrace{\lambda^{s_k} (\lambda - \mathrm{i}t)^{-s_k}}_{\text{Gamma distribution}}}_{\text{GNIG distribution}} \end{aligned}$$

that is the c.f. of the mixture of three GNIG distributions of depth p as a near-exact distribution for W with the p.d.f and c.d.f. for the GNIG distribution given in the Appendix A. Having in mind the general relation

$$F_{\lambda^*}(w) = 1 - F_W(-\log w)$$

we obtain the c.d.f. and also the p.d.f. of λ^* given in (27) and (27) respectively. The parameters p_ν, s_ν and λ are defined in such a way that the first six moments of W_2^* match the first six exact moments of W_2, that is, in such a way that

$$\frac{d^j}{dt^j}\,\Phi^*_{W_2}(t)\bigg|_{t=0} = \frac{d^j}{dt^j}\,\Phi_{W_2}(t)\bigg|_{t=0}, \qquad j=1,\ldots,6$$

what gives rise to the evaluation of these parameters as the numerical solution of the correspondent system of equations. □

References

1. Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis*, 3rd ed. New York: J. Wiley & Sons.
2. Anderson, T. W. and Fang, K. T. and Hsu, H. (1986). Maximum-likelihood estimates and likelihood-ratio criteria for multivariate elliptically contoured distributions. *Canadian Journal of Statistics*, 14, 55-59.
3. Anderson, T. W. and Fang, K. T. (1990). Inference in multivariate elliptically contoured distributions base on maimum likelihood. Satistical Inference in Elliptially Contoured and Related Distributions. *K. T. Fang and T. W. Anderson*, 201-216.
4. Box, G. E. P. (1949). A general distribution theory for a class of likelihood criteria. *Biometrika*, 36, 317-346.
5. Coelho, C. A. (1998). The Generalized Integer Gamma Distribution – A basis for distributions in Multivariate Statistics. *Journal of Multivariate Analysis*, 64, 86-102.
6. Coelho, C. A. (2004). 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. *Journal of Multivariate Analysis*, 89, 191-218.
7. Coelho, C. A. and Marques, F. J. (2009). The advantage of decomposing elaborate hypotheses on covariance matrices into conditionally independent hypotheses in building near-exact distributions for the test statistics. *Linear Algebra and its Applications*, 430, 2592-2606.
8. Coelho, C. A. and Marques, F. J. (2010). Near-exact distributions for the independence and sphericity likelihood ratio test statistics. *Journal of Multivariate Analysis*, 101, 583–593.
9. Jensen DR, Good IJ (1981) Invariant distributions associated with matrix laws under structural symmetry. *J. Roy. Statist. Soc. Ser. B*, 43, 327–332.
10. Kariya T (1981) Robustness of multivariate tests. *Ann. Statist.*, 9, 1267–1275.
11. Krishnaiah, P. R., Lee, J. C. and Chang, T. C. (1976). The distribution of the likelihood ratio statistics for tests of certain covariance structures of complex multivariate normal populations. *Biometrika*, 63, 513-519.
12. Marques, F. J. and Coelho, C. A. (2008). Near-exact distributions for the sphericity likelihood ratio test statistic. *Journal of Statistical Planning and Inference*, 138, 726-741.
13. Marques, F. J., Coelho, C. A. and Arnold, B. C. (2010). A general near-exact distribution theory for the most common likelihood ratio test statistics used in Multivariate Analysis. *Test*, (in print).
14. Moschopoulos, P. G. (1992). The hypothesis of multisample block sphericity. *Sankhyã; The Indian Journal of Statistics*, 54, 260-270.
15. Nagar, D. K., Jain, S. K. and Gupta, A. K. (1985). Distribution of lrt for testing sphericity of a complex multivariate Gaussian model. *Internat. J. Math. & Math. Sci.*, 8, 555-562.
16. Nagarsenker, B. N. and Nagarsenker, P. B. (1981). Distribution of the likelihood ratio statistic for testing sphericity structure for a complex normal covariance matrix. *Sankhya*, B43, 352-359.
17. Pillai, K. C. S. and Nagarsenker, B. N. (1971). On the distribution of the sphericity test criterion in classical and complex normal populations having unknown covariance matrices. *The Annals of Mathematics Statistics*, 42, 764-767.