The multi-sample block-scalar sphericity test – exact and near-exact distributions for its likelihood ratio test statistic

Carlos A. Coelho*, Filipe J. Marques

Departamento de Matemática and Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa

Abstract

In this paper the authors show how by adequately decomposing the null hypothesis of the multi-sample block sphericity test it is possible to easily obtain the expression for its likelihood ratio test statistic as well as a different look over the exact distribution of this statistic. This different view will enable the construction of very well-performing near-exact approximations for the distribution of this test statistic, whose exact distribution is quite elaborate and non-manageable. The near-exact distributions obtained are quite manageable and perform much better than the available asymptotic distributions, even for small sample sizes, and they show a good asymptotic behavior not only for increasing sample sizes as well as for increasing number of variables and/or samples involved.

Keywords: mixtures, Generalized Integer Gamma distribution, Generalized Near-Integer Gamma distribution.

1. Introduction

The multi-sample block-scalar sphericity test, hereafter simply designated by multi-sample block sphericity test, is a test which null hypothesis may be written as

$$H_0:\ \Sigma_1 = \cdots = \Sigma_m = diag(\Delta_1,\ldots,\Delta_k)\,, \qquad \text{with} \qquad \Delta_i = \sigma_i^2\, I_{p_i},\ (i=1,\ldots,k)\,, \tag{1}$$

where σ_i^2 $(i = 1,\ldots,k)$ is non-specified.

This test is an interesting and useful generalization of the standard sphericity test and is of interest in several studies and statistical applications where we need to carry out a bus test, to test:

- if a set of m multivariate normal or elliptically contoured populations all have the same variance-covariance matrix,
- if this matrix has a block-diagonal structure, that is, if the overall set of variables is indeed formed by a given number of independent sub-sets of variables, and
- if in each of these sets all variables are independent and have the same variance.

*Corresponding author: Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Departamento de Matemática, Quinta da Torre, 2829-516 Caparica; phone: +351 21 294 8388, Fax: +351 21 294 8391.

Email addresses: `cmac@fct.unl.pt` (Carlos A. Coelho), `fjm@fct.unl.pt` (Filipe J. Marques)

The multi-sample block sphericity test plays a key role in tests of homocedasticity in multivariate analysis and repeated measures designs, where the validity of several other tests rest on the assumption of sphericity. His role is the equivalent to the one that the multi-sample sphericity test has in univariate analysis.

The multi-sample block sphericity test has as particular cases a number of other interesting tests, namely:

- for $m = 1$, the one-sample block sphericity test $\left(H_0 \colon \ \Sigma = diag(\sigma_i^2 I_{p_i},\ i = 1, \ldots, k)\right)$
- for $k = 1$, the multi-sample sphericity test $\left(H_0 \colon \ \Sigma_1 = \ldots \Sigma_m = \sigma^2 I_p\right)$
- for $m = 1$ and $k = 1$, the usual sphericity test $\left(H_0 \colon \ \Sigma = \sigma^2 I_p\right)$
- for all $p_i = 1$ $(i = 1, \ldots, k)$, the multi-sample independence test $\left(H_0 \colon \ \Sigma_1 = \ldots \Sigma_m = diag(\sigma_j^2,\ j = 1, \ldots, k)\right)$
- for $m = 1$ and all $p_i = 1$ $(i = 1, \ldots, k)$, the usual (one-sample) independence test $\left(H_0 \colon \ \Sigma = diag(\sigma_j^2,\ j = 1, \ldots, k)\right)$.

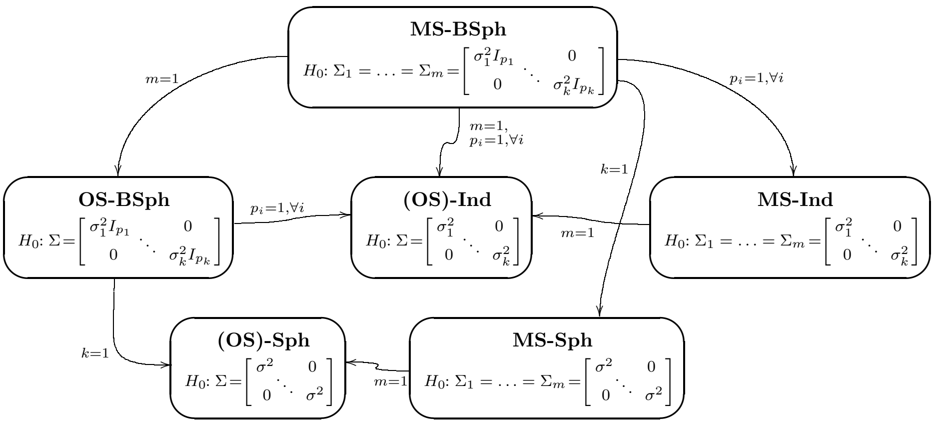

Figure 1 – The multi-sample block sphericity test and its particular cases: **MS-BSph** - multi-sample block sphericity, **OS-BSph** - one-sample block sphericity, **(OS)-Sph** - (one-sample) sphericity, **(OS)-Ind** (one-sample) independence, **MS-Ind** multi-sample independence, **MS-Sph** - multi-sample sphericity.

In this paper we will show how, following Coelho and Marques (2009), by using a suitable decomposition of the test null hypothesis into a set of nested, conditionally independent, hypotheses, we will be able to induce a much useful factorization of the characteristic function of the negative logarithm of the associated likelihood ratio test (l.r.t.) statistic, which may be used

- to easily find the structure of the exact distribution of this statistic, and
- to develop extremely sharp and well-performing near-exact approximations to the exact distribution of the l.r.t. statistic.

These near-exact distributions are asymptotic distributions built using a different concept of approximation:

– starting with an adequate factorization of the c.f. of the negative logarithm of the l.r.t. statistic,

– the major part of it is left unchanged, while the remaining smaller part is asymptotically approximated, in such a way that

– the product of the two resulting terms yields a known manageable distribution.

These distributions besides being asymptotic for increasing sample sizes, are also asymptotic for increasing number of variables or sets of variables, and increasing number of populations involved, as it is shown by the results of the numerical studies carried out in Section 4.

Other characteristics of these near-exact distributions are:

– they lie very close to the exact distribution and they match, by construction, some of the first exact moments;

– they perform very well even for very small sample sizes;

– they are far more manageable than the exact distributions and are quite easy to implement computationally, allowing for the easy computation of near-exact quantiles and p-values, which being so close to the exact ones, may, in practice, be used instead of these latter ones.

2. A decomposition of the null hypothesis which induces an adequate factorization of the characteristic function of the logarithm of the likelihood ratio test statistic and the exact distribution of the likelihood ratio test statistic to test H_0 in (1)

The null hypothesis of the multi-sample block sphericity test in (1) may be written as

$$H_0 \equiv H_{0c|(0b|0a)} \circ H_{0b|0a} \circ H_{0a} \,,$$

where

$$H_{0a} : \Sigma_1 = \cdots = \Sigma_m (= \Sigma) \,, \quad (\Sigma_{(p \times p)} \text{ unspecified}) \tag{2}$$

is the null hypothesis to test the equality of m covariance matrices, with $p = \sum_{i=1}^{k} p_i$ and

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1k} \\ \Sigma_{21} & \Sigma_{22} & \cdots & \Sigma_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{1k} & \Sigma_{2k} & \cdots & \Sigma_{kk} \end{bmatrix},$$

$$H_{0b|0a} : \begin{array}{l} \Sigma_{ij} = 0 \;\text{ for } i \neq j \quad (i, j \in \{1, \ldots, k\}) \\ \text{assuming } \Sigma_1 = \cdots = \Sigma_m (= \Sigma), \end{array} \tag{3}$$

is the null hypothesis to test the independence of k groups of r.v.'s, and

$$H_{0c|(0b|0a)} : \begin{array}{l} \Sigma_{ii} = \sigma_i^2 I_{p_i} \,, \; i = 1, \ldots, k \;\; (k \text{ independent sphericity tests}) \\ \text{assuming } \Sigma_{ij} = 0 \;\text{ for } i \neq j \quad (i, j \in \{1, \ldots, k\}) \end{array} \tag{4}$$

is the null hypothesis to test the sphericity of the k covariance matrices Σ_{ii} ($i = 1, \ldots, k$).

Now, let us suppose that we want to test the null hypothesis H_0 in (1), based on m independent samples, one from each of the m populations $N_p(\underline{\mu}_j, \Sigma_j)$ $(j = 1, \ldots, m)$. Let N_j be the size of the sample from the j-th population and let $N = \sum_{j=1}^m N_j$. Let also A_j be the maximum likelihood estimator of Σ_j and $A = \sum_{j=1}^m A_j$. Then the l.r.t. statistic to test H_{0a} is

$$\Lambda_a = \frac{N^{Np/2}}{\prod_{j=1}^m N_j^{pN_j/2}} \frac{\prod_{j=1}^m |A_j|^{N_j/2}}{|A|^{N/2}},$$

while the l.r.t. statistic to test $H_{0b|0a}$ is

$$\Lambda_{b|a} = \frac{|A|^{N/2}}{\prod_{i=1}^k |A_{ii}|^{N/2}},$$

where A_{ii} is the i-th $(i = 1, \ldots, k)$ diagonal block of A, and the l.r.t. statistic used to test $H_{0c|(0b|0a)}$ is

$$\Lambda_{c|(b|a)} = \prod_{i=1}^k \frac{|A_{ii}|^{N/2}}{(tr\,A_{ii})^{Np_i/2}}\, p_i^{Np_i/2}.$$

Let then Λ be the l.r.t. statistic used to test H_0 in (1). We have (see Lemma 10.3.1 in Anderson (2003))

$$\Lambda = \Lambda_a\, \Lambda_{b|a}\, \Lambda_{c|(b|a)} = \frac{\prod_{i=1}^k (p_i N)^{p_i N/2}}{\prod_{j=1}^m N_j^{pN_j/2}} \frac{\prod_{j=1}^m |A_j|^{N_j/2}}{\prod_{i=1}^k (tr\,A_{ii})^{Np_i/2}}, \tag{5}$$

and, given the independence, under H_0 in (1) of the l.r.t. statistics to test H_{0a}, $H_{0b|0a}$ and $H_{0c|(0b|0a)}$ (see Appendix A),

$$E\left(\Lambda^h\right) = E\left(\Lambda_a^h\right) E\left(\Lambda_{b|a}^h\right) E\left(\Lambda_{c|(b|a)}^h\right),$$

where $E\left(\Lambda_a^h\right)$, $E\left(\Lambda_{b|a}^h\right)$ and $E\left(\Lambda_{c|(b|a)}^h\right)$ may be obtained from (Anderson, 2003, Chap. 9,10) or (Muirhead, 2005, Chap. 8,11). However, the final expression obtained in this way for $E\left(\Lambda^h\right)$ is not quite useful in order to better understand and work through the fine details of the distribution of Λ. We rather need a more indirect approach.

Let then $W = -\log \Lambda$ and also

$$W_a = -\log \Lambda_a\,, \quad W_{b|a} = -\log \Lambda_{b|a} \quad \text{and} \quad W_{c|(b|a)} = -\log \Lambda_{c|(b|a)}\,.$$

Then we have

$$\begin{aligned} \Phi_W(t) &= E\left(e^{\mathrm{i}tW}\right) = E\left(e^{-\mathrm{i}t\left(W_a + W_{b|a} + W_{c|(b|a)}\right)}\right) = E\left(e^{\mathrm{i}tW_a}\right) E\left(e^{\mathrm{i}tW_{b|a}}\right) E\left(e^{\mathrm{i}tW_{c|(b|a)}}\right) \\ &= \Phi_{W_a}(t) \times \Phi_{W_{b|a}}(t) \times \Phi_{W_{c|(b|a)}}(t)\,, \end{aligned} \tag{6}$$

where, for $N_j = n$ $(j = 1, \ldots, m)$, $p = \sum_{i=1}^k p_i$ and $q_i = p_{i+1} + \cdots + p_k$ $(i = 1, \ldots, k-1)$, $\Phi_{W_a}(t)$, $\Phi_{W_{b|a}}(t)$ and $\Phi_{W_{c|(b|a)}}(t)$, after some algebraic manipulation may be written as (see Marques et al. (2011)),

$$\Phi_{W_a}(t) = \prod_{\ell=1}^p \prod_{j=1}^m \frac{\Gamma\left(\frac{n-1}{2} - \frac{\ell-1}{2m} + \frac{j-1}{m}\right) \Gamma\left(\frac{n-\ell}{2} - \frac{n}{2}\mathrm{i}t\right)}{\Gamma\left(\frac{n-1}{2} - \frac{\ell-1}{2m} + \frac{j-1}{m} - \frac{n}{2}\mathrm{i}t\right) \Gamma\left(\frac{n-\ell}{2}\right)}, \tag{7}$$

$$\Phi_{W_{b|a}}(t) = \prod_{i=1}^{k-1}\prod_{\ell=1}^{p_i} \frac{\Gamma\left(\frac{nm-\ell}{2}\right)\Gamma\left(\frac{nm-q_i-\ell}{2} - \frac{nm}{2}\text{i}t\right)}{\Gamma\left(\frac{nm-q_i-\ell}{2}\right)\Gamma\left(\frac{nm-\ell}{2} - \frac{nm}{2}\text{i}t\right)}\,, \tag{8}$$

and

$$\Phi_{W_{c|(b|a)}}(t) = \prod_{i=1}^{k}\prod_{\ell=2}^{p_i} \frac{\Gamma\left(\frac{nm-1}{2} + \frac{\ell-1}{p_i}\right)\Gamma\left(\frac{nm-\ell}{2} - \frac{nm}{2}\text{i}t\right)}{\Gamma\left(\frac{nm-1}{2} + \frac{\ell-1}{p_i} - \frac{nm}{2}\text{i}t\right)\Gamma\left(\frac{nm-\ell}{2}\right)}\,. \tag{9}$$

These expressions show that the exact distribution of Λ is indeed the same as the distribution of the product of $p(m+2) - p_k - k - 1$ independent Beta r.v.'s, raised to the power $n/2$, since from (6), and (7)-(9) above we may write

$$\Lambda \stackrel{st}{\sim} \left\{ \underset{\text{except for } \ell=j=1}{\prod_{\ell=1}^{p}\prod_{j=1}^{m}} Y_{\ell j} \right\}^{n/2} \left\{\prod_{i=1}^{k-1}\prod_{\ell=1}^{p_i} Y^*_{i\ell}\right\}^{n/2} \left\{\prod_{i=1}^{k}\prod_{\ell=2}^{p_i} Y^{**}_{i\ell}\right\}^{n/2}\,, \tag{10}$$

where $\stackrel{st}{\sim}$ is to be read 'is stochastically equivalent to' and

$$\underset{\substack{\ell=1,\dots,p;\ j=1,\dots,m\\ (\text{except for } \ell=j=1)}}{Y_{\ell j} \sim Beta\left(\frac{n-\ell}{2}, \frac{j-1}{m} + \frac{\ell-1}{2} - \frac{\ell-1}{2m}\right)} \qquad \underset{i=1,\dots,k-1;\ \ell=1,\dots,p_i}{Y^*_{i\ell} \sim Beta\left(\frac{nm-q_i-\ell}{2}, \frac{q_i}{2}\right)} \qquad \underset{i=1,\dots,k;\ \ell=2,\dots,p_i}{Y^{**}_{i\ell} \sim Beta\left(\frac{nm-\ell}{2}, \frac{\ell-1}{2} + \frac{\ell-1}{p_i}\right)}\,.$$

However, another more useful way to look at the exact distribution of Λ may be obtained from the fact that, after some rather long manipulations, we may write, for

$$p \perp\!\!\!\perp 2 = \begin{cases} 0 & \text{if} \quad p \text{ even} \\ 1 & \text{if} \quad p \text{ odd}\,, \end{cases}$$

$\Phi_{W_a}(t)$ as (see Marques et al. (2011) and Coelho and Marques (2011))

$$\begin{aligned} \Phi_{W_a}(t) \;=\; & \underbrace{\left\{\prod_{\ell=2}^{p}\left(\frac{n-\ell}{n}\right)^{r_\ell}\left(\frac{n-\ell}{n} - \text{i}t\right)^{-r_\ell}\right\}}_{\Phi_{1,W_a}(t)} \\ & \underbrace{\times\left\{\prod_{\ell=1}^{\lfloor p/2\rfloor}\prod_{j=1}^{m} \frac{\Gamma(a_\ell + b_{\ell j})\,\Gamma(a_\ell + b^*_{\ell j} - n\text{i}t)}{\Gamma(a_\ell + b^*_{\ell j})\,\Gamma(a_\ell + b_{\ell j} - n\text{i}t)}\right\}\times\left\{\prod_{j=1}^{m}\frac{\Gamma(a_p + b_{pj})\,\Gamma(a_p + b^*_{pj} - \frac{n}{2}\text{i}t)}{\Gamma(a_p + b^*_{pj})\,\Gamma(a_p + b_{pj} - \frac{n}{2}\text{i}t)}\right\}^{p \perp\!\!\!\perp 2}}_{\Phi_{2,W_a}(t)} \end{aligned} \tag{11}$$

where

$$a_\ell = n - 2\ell\,, \qquad b_{\ell j} = 2\ell - 1 + \frac{j-2\ell}{m}\,, \qquad b^*_{\ell j} = \lfloor b_{\ell j} \rfloor\,, \tag{12}$$

$$a_p = \frac{n-p}{2}\,, \qquad b_{pj} = \frac{pm - m - p + 2j - 1}{2m}\,, \qquad b^*_{pj} = \lfloor b_{pj} \rfloor\,, \tag{13}$$

$$r_\ell = \begin{cases} r^*_{\ell-1} & \ell = 2, \ldots, p, \\ & \text{except for } \ell = p - 2\alpha_1, \\ r^*_{\ell-1} + (p \perp\!\!\!\perp 2)(\alpha_2 - \alpha_1)\left(m - \frac{p-1}{2} + m\left\lfloor \frac{p}{2m} \right\rfloor\right) & \ell = p - 2\alpha_1\,, \end{cases} \tag{14}$$

with

$$r^*_\ell = \begin{cases} \gamma_\ell & \ell = 1, \ldots, \alpha + 1, \\ m\left(\left\lfloor \frac{p}{2} \right\rfloor - \left\lfloor \frac{\ell}{2} \right\rfloor\right) & \ell = \alpha + 2, \ldots, \min(p - 2\alpha_1, p - 1), \\ & \text{and } \ell = 2 + p - 2\alpha_1, \ldots, 2\left\lfloor \frac{p}{2} \right\rfloor - 1, \text{ by steps of 2} \\ m\left(\left\lfloor \frac{p+1}{2} \right\rfloor - \left\lfloor \frac{\ell}{2} \right\rfloor\right) & \ell = 1 + p - 2\alpha_1, \ldots, p - 1, \text{ by steps of 2,} \end{cases} \tag{15}$$

and

$$\alpha = \left\lfloor \frac{p-1}{m} \right\rfloor, \quad \alpha_1 = \left\lfloor \frac{m-1}{m}\,\frac{p-1}{2} \right\rfloor, \quad \alpha_2 = \left\lfloor \frac{m-1}{m}\,\frac{p+1}{2} \right\rfloor, \tag{16}$$

where, for $\ell = 1, \ldots, \alpha$,

$$\gamma_\ell = \left\lfloor \frac{m}{2} \right\rfloor \left((\ell - 1)m - 2\left((m+1) \perp\!\!\!\perp 2\right) \left\lfloor \frac{\ell}{2} \right\rfloor \right) + \left\lfloor \frac{m}{2} \right\rfloor \left\lfloor \frac{m + \ell \perp\!\!\!\perp 2}{2} \right\rfloor \tag{17}$$

and

$$\gamma_{\alpha+1} = -\left(\left\lfloor \frac{p}{2} \right\rfloor - \alpha \left\lfloor \frac{m}{2} \right\rfloor \right)^2 + m\left(\left\lfloor \frac{p}{2} \right\rfloor - \left\lfloor \frac{\alpha+1}{2} \right\rfloor \right) + (m \perp\!\!\!\perp 2)\left(\alpha \left\lfloor \frac{p}{2} \right\rfloor + \frac{\alpha \perp\!\!\!\perp 2}{4} - \frac{\alpha^2}{4} - \alpha^2 \left\lfloor \frac{m}{2} \right\rfloor \right), \tag{18}$$

which shows that W_a has the distribution of the sum of a r.v. with a GIG (Generalized Integer Gamma) distribution of depth $p - 1$, with an independent sum of $(m-1)\left\lfloor \frac{p+1}{2} \right\rfloor$ independent Logbeta r.v.'s, some of them multiplied by n and the other multiplied by $n/2$. Note that in each product in j in $\Phi_{2,W_a}(t)$ the Logbeta distribution vanishes for one of the values of j. The GIG distribution is the distribution of the sum of independent Gamma r.v.'s with integer shape parameters. For details see Appendix B and Coelho (1998, 1999).

We may also write (see Marques et al. (2011) and Coelho (2004))

$$\Phi_{W_{b|a}}(t) = \underbrace{\left\{ \prod_{\ell=3}^{p} \left(\frac{n - \frac{\ell}{m}}{n} \right)^{s_\ell} \left(\frac{n - \frac{\ell}{m}}{n} - \mathrm{i}t \right)^{-s_\ell} \right\}}_{\Phi_{1,W_{b|a}}(t)} \underbrace{\left(\frac{\Gamma\left(\frac{nm-1}{2}\right)\Gamma\left(\frac{nm-1}{2} - \frac{1}{2} - \frac{nm}{2}\mathrm{i}t\right)}{\Gamma\left(\frac{nm-1}{2} - \frac{1}{2}\right)\Gamma\left(\frac{nm-1}{2} - \frac{nm}{2}\mathrm{i}t\right)} \right)^{k^*}}_{\Phi_{2,W_{b|a}}(t)} \tag{19}$$

where $k^* = \left\lfloor \frac{q}{2} \right\rfloor$, with q denoting the number of odd p_i's, and where

$$s_\ell = \begin{cases} h_{\ell-2} + (-1)^\ell k^* & \ell = 3, 4 \\ s_{\ell-2} + h_{\ell-2} & \ell = 5, \ldots, p \end{cases} \tag{20}$$

with

$$h_\ell = (\#\text{ of } p_i \ (i = 1, \ldots, k) \geq \ell) - 1\,, \quad \ell = 1, \ldots, p-2\,, \tag{21}$$

which shows that the distribution of $W_{b|a}$ is the same as that of the sum a r.v. with a GIG distribution of depth $p - 2$, with an independent sum of k^* independent and identically distributed Logbeta r.v.'s multiplied by $nm/2$.

And yet

$$\Phi_{W_{c|(b|a)}}(t) = \underbrace{\left\{\prod_{i=1}^{k}\prod_{\ell=2}^{p_i}\left(\frac{n-\frac{\ell}{m}}{n}\right)^{z_{\ell,i}}\left(\frac{n-\frac{\ell}{m}}{n}-\mathrm{i}t\right)^{-z_{\ell,i}}\right\}}_{\Phi_{1,W_{c|(b|a)}}(t)}$$

$$\times\underbrace{\left\{\prod_{i=1}^{k}\left\{\prod_{\ell=2}^{p_i-k_i^*}\frac{\Gamma\left(\frac{nm-1}{2}+\frac{\ell-1}{p_i}\right)\Gamma\left(\frac{nm-1}{2}-\frac{nm}{2}\mathrm{i}t\right)}{\Gamma\left(\frac{nm-1}{2}\right)\Gamma\left(\frac{nm-1}{2}+\frac{\ell-1}{p_i}-\frac{nm}{2}\mathrm{i}t\right)}\right\}\left\{\prod_{\ell=p_i-k_i^*+1}^{p_i}\frac{\Gamma\left(\frac{nm-1}{2}+\frac{\ell-1}{p_i}\right)\Gamma\left(\frac{nm}{2}-\frac{nm}{2}\mathrm{i}t\right)}{\Gamma\left(\frac{nm}{2}\right)\Gamma\left(\frac{nm-1}{2}+\frac{\ell-1}{p_i}-\frac{nm}{2}\mathrm{i}t\right)}\right\}\right\}}_{\Phi_{2,W_{c|(b|a)}}(t)} \quad (22)$$

where $k_i^* = \lfloor p_i/2 \rfloor$ and

$$z_{\ell,i} = \left\lfloor \frac{p_i-\ell+2}{2} \right\rfloor, \quad \ell = 2,\dots,p_i\,,\; i=1,\dots.k\,, \quad (23)$$

which shows that the distribution of $W_{c|(b|a)}$ is the same as that of the sum a r.v. with a GIG distribution of depth $p-k$, with an independent sum of $p-k$ independent Logbeta r.v.'s, multiplied by $nm/2$.

As such, it will be possible to express the distribution of W as that of the sum of a GIG distributed r.v. with an independent sum of independently distributed Logbeta r.v.'s. This actually amounts to being able to write the c.f. of $W = -\log \Lambda$ as in the following Theorem.

Theorem 1. *The c.f. of $W = -\log \Lambda$, where Λ is the l.r.t. statistic in* (5)*, may be written as*

$$\Phi_W(t) = \underbrace{\prod_{\ell=1}^{p}\left(\frac{n-\ell}{n}\right)^{r_\ell^+}\left(\frac{n-\ell}{n}-\mathrm{i}t\right)^{-r_\ell^+}}_{\Phi_1(t)}\underbrace{\prod_{\substack{\ell=2\\ \ell\neq m,2m,\dots,\alpha m}}^{p}\left(\frac{n-\frac{\ell}{m}}{n}\right)^{r_\ell^{++}}\left(\frac{n-\frac{\ell}{m}}{n}-\mathrm{i}t\right)^{-r_\ell^{++}}}_{\Phi_2(t)} \quad (24)$$

$$\times\,\Phi_{2,W_a}(t)\,\Phi_{2,W_{b|a}}(t)\,\Phi_{2,W_{c|(b|a)}}(t)\,,$$

where $\Phi_{2,W_a}(t)$, $\Phi_{2,W_{b|a}}(t)$ and $\Phi_{2,W_{c|(b|a)}}(t)$ are given respectively in (11), (19) *and* (22)*, and where m is the same as in* (1)*, $\alpha = \left\lfloor \frac{p-1}{q} \right\rfloor$,*

$$r_\ell^+ = \begin{cases} r_m^{++} & \ell = 1 \\ r_\ell + r_{m\ell}^{++} & \ell = 2,\dots,\alpha \\ r_\ell & \ell = \alpha+1,\dots,p \end{cases} \quad (25)$$

and

$$r_\ell^{++} = \begin{cases} z_\ell^{**} & \ell = 2 \\ z_\ell^{**} + s_\ell & \ell = 3,\dots,p \end{cases} \quad (26)$$

with r_ℓ and s_ℓ respectively given by (14)-(18) *and* (20)-(21) *and with*

$$z_\ell^{**} = \begin{cases} \sum_{i=1}^{k} z_{\ell,i}^* & \ell = 2,\dots,p_{max} \\ 0 & \ell = p_{max}+1,\dots,p \end{cases}$$

where

$$z^*_{\ell,i} = \begin{cases} z_{\ell,i} & \ell = 2, \dots, p_i \\ 0 & \ell = p_i + 1, \dots, p_{max} \end{cases}$$

for $p_{max} = \max\{p_1, \dots, p_k\}$ *and* $z_{\ell,i}$ *given by* (23).

Proof. We only have to write $\Phi_W(t)$ as

$$\begin{aligned} \Phi_W(t) &= \underbrace{\Phi_{1,W_a}(t)\,\Phi_{2,W_a}(t)}_{\Phi_{W_a}(t)} \times \underbrace{\Phi_{1,W_{b|a}}(t)\,\Phi_{2,W_{b|a}}(t)}_{\Phi_{W_{b|a}}(t)} \times \underbrace{\Phi_{1,W_{c|(b|a)}}(t)\,\Phi_{2,W_{c|(b|a)}}(t)}_{\Phi_{W_{c|(b|a)}}(t)} \\ &= \underbrace{\Phi_{1,W_a}(t)\,\Phi_{1,W_{b|a}}(t)\,\Phi_{1,W_{c|(b|a)}}(t)}_{\text{c.f. of a GIG distribution}} \times \underbrace{\Phi_{2,W_a}(t)\,\Phi_{2,W_{b|a}}(t)\,\Phi_{2,W_{c|(b|a)}}(t)}_{\text{c.f. of a sum of indep. Logbeta r.v.'s}} \end{aligned}$$

an then group together the rate parameters in $\Phi_{1,W_a}(t)\,\Phi_{1,W_{b|a}}(t)\,\Phi_{1,W_{c|(b|a)}}(t)$ and adequately add the corresponding shape parameters, in order to yield $\Phi_1(t)\Phi_2(t)$, which is the c.f. of a GIG distribution of depth $2p - 1 - \alpha$. □

It is interesting to note that in (24):

– when all or all but one of the p_i's ($i = 1, \dots, k$) are even, then $\Phi_{2,W_{b|a}}$ vanishes, or equals 1, and

– when all p_i ($i = 1, \dots, k$) equal 2, $\Phi_{2,W_{c|(b|a)}}$ vanishes.

Also, from (24) we may see that the exact distribution of Λ in (5) is the same as the distribution of the product of the exponential of the negative of $2p - 1 - \alpha$ independent Gamma r.v.'s, times the product of $(m-1)\left\lfloor \frac{p+1}{2} \right\rfloor + \left\lfloor \frac{q}{2} \right\rfloor + p - k$ independent Beta r.v.'s raised to some powers (where q is the number of odd p_i's). More precisely, from (24) we may write

$$\Lambda \overset{st}{\sim} \left\{ \prod_{\ell=1}^{p} e^{-Z_\ell} \right\} \left\{ \prod_{\substack{\ell=2 \\ \ell \neq m,\dots,\alpha m}}^{p} e^{-Z^*_\ell} \right\} \left\{ \prod_{\ell=1}^{\lfloor p/2 \rfloor} \prod_{j=1}^{m} \left(Y_{\ell j}\right)^n \right\} \left\{ \prod_{j=1}^{m(p \perp\!\!\!\perp 2)} \left(Y^*_j\right)^{n/2} \right\} \left\{ \prod_{j=1}^{\lfloor q/2 \rfloor} \left(Y^{**}_j\right)^{nm/2} \right\}$$
$$\left\{ \prod_{i=1}^{k} \prod_{\ell=2}^{p_i - k^*_i} \left(Y^{***}_{i\ell}\right)^{nm/2} \right\} \left\{ \prod_{i=1}^{k} \prod_{\ell=p_i-k^*_i+1}^{p_i} \left(Y^{****}_{i\ell}\right)^{nm/2} \right\}$$

where all the r.v.'s involved are independent and where q represents the number of odd p_i, $k^*_i = \left\lfloor \frac{p_i}{2} \right\rfloor$,

$$Z_\ell \sim \Gamma\left(r^+_\ell, \frac{n-\ell}{n}\right) \qquad \ell = 1, \dots, p$$

$$Z^*_\ell \sim \Gamma\left(r^{++}_\ell, \frac{n-\ell/m}{n}\right) \qquad \ell = 2, \dots, p;\ \ell \neq m, 2m, \dots, \alpha m$$

and

$$
\begin{aligned}
&Y_{\ell j} \sim Beta\left(a_\ell + b^*_{\ell j}, b_{\ell j} - b^*_{\ell j}\right) \qquad \ell = 1, \ldots, \lfloor p/2 \rfloor;\;\; j = 1, \ldots, m \;(j \neq 2\ell \perp\!\!\!\perp^* m)\\
&Y^*_j \sim Beta\left(a_p + b^*_{pj}, b_{pj} - b^*_{pj}\right) \qquad j = 1, \ldots, m(p \perp\!\!\!\perp 2) \;\left(j \neq \frac{p+1}{2} \perp\!\!\!\perp^* m\right)\\
&Y^{**}_j \sim Beta\left(\frac{nm-2}{2}, \frac{1}{2}\right) \qquad j = 1, \ldots, \lfloor q/2 \rfloor\\
&Y^{***}_{i\ell} \sim Beta\left(\frac{nm-1}{2}, \frac{\ell-1}{p_i}\right) \qquad i = 1, \ldots, k;\;\; \ell = 2, \ldots, p_i - \lfloor p_i/2 \rfloor\\
&Y^{****}_{i\ell} \sim Beta\left(\frac{nm}{2}, \frac{\ell-1}{p_i} - \frac{1}{2}\right) \qquad i = 1, \ldots, k;\;\; \ell = p_i - \lfloor p_i/2 \rfloor + 1, \ldots, p_i\,,
\end{aligned} \tag{27}
$$

where q is the number of odd p_i, r^+_ℓ and r^{++}_ℓ are given by (25) and (26), and a_ℓ, $b_{\ell j}$, a_p and b_{pj} are given by (12)-(13) and where

$$
a \perp\!\!\!\perp^* b = \begin{cases} a \perp\!\!\!\perp b\,, & a \perp\!\!\!\perp b \neq 0 \\ m & a \perp\!\!\!\perp b = 0 \end{cases}
$$

where $a \perp\!\!\!\perp b$ represents the remainder of the integer ratio of a by b.

Although this representation of the distribution looks decidedly far more elaborate than the one in (10) and as such it may seem quite useless, it happens that it will enable us to develop very well-fitting near-exact distributions, which bear an extreme closeness to the exact distribution of Λ.

3. Near-exact distributions

3.1. The case of equal sample sizes

In this subsection we will address the case of equal sample sizes, which is indeed the case treated in the previous section and referred to in Theorem 1. In the next subsection we will address the case of unequal sample sizes, based on the approach followed for the equal sample sizes case.

The distributions of W and Λ are quite elaborate and it is not easy to obtain a manageable form for their exact distribution. However, the way the distribution of W is shown in (24) in Theorem 1 enables us to obtain very well-fitting near-exact approximations. These will be obtained by leaving $\Phi_1(t)\Phi_2(t)$ in (24) unchanged and approximating asymptotically the term $\Phi_{2,W_a}(t)\,\Phi_{2,W_{b|a}}(t)\,\Phi_{2,W_{c|(b|a)}}(t)$ by the c.f. of a finite mixture of Gamma distributions, all with the same rate parameter.

This asymptotic replacement is quite easy to justify, since from the results in Tricomi and Erdélyi (1951) we may infer that any *Logbeta*(a, b) distribution may be asymptotically approximated by an infinite mixture of $\Gamma(b + j, a)$ $(j = 0, 1, \ldots)$ distributions. As such, we may replace any sum of any number of independent Logbeta r.v.'s by an infinite mixture of sums of that same number of independent Gamma r.v.'s. The problem is that these Gamma distributions coming out of the Logbeta distributions in $\Phi_{2,W_a}(t)\,\Phi_{2,W_{b|a}}(t)\,\Phi_{2,W_{c|(b|a)}}(t)$, as may be seen from the exposition at the end of last section, do not have the same rate parameter, thus rendering difficult the obtention of a manageable expression for the whole mixture, since then each component of this infinite mixture would be itself an infinite mixture.

But then we may argue that a good approximation for each component of that infinite mixture would be a Gamma distribution with a rate parameter which would be the average of the different rate parameters and a shape parameter which would be the sum of the shape parameters of the different independent Gamma distributions. Although this would indeed work quite well, there is indeed another choice for this rate parameter which works even much better. We will state it shortly. Another problem which is left is the problem of computing the weights of that mixture, which, using the expressions in Tricomi and Erdélyi (1951) may turn into a quite hard task.

A good solution for these two problems may be obtained in the following way. Let us take for near-exact c.f. of W the c.f.

$$\Phi^*(t) = \Phi_1(t)\Phi_2(t)\sum_{\nu=0}^{m^*} \pi_\nu \lambda^{r+\nu}(\lambda - \mathrm{i}t)^{-(r+\nu)} \tag{28}$$

where λ is the common rate parameter in a mixture of two Gamma distributions which matches the first four exact moments of W, that is, λ is the rate parameter in

$$\Psi(t) = \pi^* \lambda^{u_1}(\lambda - \mathrm{i}t)^{-u_1} + (1-\pi^*)\lambda^{u_2}(\lambda - \mathrm{i}t)^{-u_2} \tag{29}$$

where π^*, u_1, u_2 and λ are determined so that

$$\left.\frac{\partial^h}{\partial t^h}\Psi(t)\right|_{t=0} = \left.\frac{\partial^h}{\partial t^h}\Phi_W(t)\right|_{t=0} \quad \text{for} \quad h = 1, \ldots, 4\,.$$

Then, we will take in (28) r as the sum of all the second parameters in the Beta distributions in (27), that is,

$$r = \frac{m-1}{2}\left\lfloor \frac{p+1}{2} \right\rfloor + \frac{1}{2}\left\lfloor \frac{q}{2} \right\rfloor + \sum_{i=1}^{k} \frac{1}{2}\left(\left\lceil \frac{p_i}{2} \right\rceil - 1\right) \tag{30}$$

where q is the number of odd p_i, and we will compute the weights π_ν ($\nu = 0, \ldots, m^* - 1$) in such a way that

$$\left.\frac{\partial^h}{\partial t^h}\Phi^*(t)\right|_{t=0} = \left.\frac{\partial^h}{\partial t^h}\Phi_W(t)\right|_{t=0} \quad \text{for} \quad h = 1, \ldots, m^*\,,$$

taking then $\pi_{m^*} = 1 - \sum_{\nu=0}^{m^*-1} \pi_\nu$.

It is not hard to see that (30) yields for r values which are either integer values or one half of an odd integer.

Using the notation in Appendix B, in case r in (30) is an integer, the near-exact c.f. in (28) yields for W near-exact distributions which are mixtures of $m^* + 1$ GIG distributions of depth $2p - \alpha$ (for α given by (16)), with p.d.f.

$$\sum_{\nu=0}^{m^*} \pi_\nu f^{GIG}\Bigg(w \,|\, \underbrace{r_1^+, \ldots, r_p^+}_{p \text{ shape param.}}, \underbrace{r_2^{++}, \ldots, r_p^{++}}_{p-1-\alpha \text{ shape param.}},\, r+\nu\,;\, \underbrace{\frac{n-1}{n}, \ldots, \frac{n-p}{n}}_{p \text{ rate param.}}, \underbrace{\frac{n-2/m}{n}, \ldots, \frac{n-p/m}{n}}_{p-1-\alpha \text{ rate param.}},\, \lambda\Bigg)$$

and c.d.f.

$$\sum_{\nu=0}^{m^*} \pi_\nu F^{GIG}\left(w \,|\, r_1^+, \ldots, r_p^+, r_2^{++}, \ldots, r_p^{++},\, r+\nu\,;\, \frac{n-1}{n}, \ldots, \frac{n-p}{n}, \frac{n-2/m}{n}, \ldots, \frac{n-p/m}{n},\, \lambda\right),$$

while if r in (30) is one half of an odd integer, the same near-exact c.f. yields for W near-exact distributions which are mixtures of $m^* + 1$ GNIG (Generalized Near-Integer Gamma) distributions of depth $2p - \alpha$ (see Appendix B for the GNIG distribution), with p.d.f.

$$\sum_{\nu=0}^{m^*} \pi_\nu f^{GNIG}\left(w \,|\, r_1^+, \ldots, r_p^+, r_2^{++}, \ldots, r_p^{++},\, r+\nu\,;\, \frac{n-1}{n}, \ldots, \frac{n-p}{n}, \frac{n-2/m}{n}, \ldots, \frac{n-p/m}{n},\, \lambda\right) \tag{31}$$

and c.d.f.

$$\sum_{\nu=0}^{m^*} \pi_\nu F^{GNIG}\left(w \,|\, r_1^+,\ldots,r_p^+, r_2^{++},\ldots,r_p^{++},\, r+\nu\,;\, \frac{n-1}{n},\ldots,\frac{n-p}{n},\frac{n-2/m}{n},\ldots,\frac{n-p/m}{n},\,\lambda\right), \tag{32}$$

where r_ℓ^+ ($\ell = 1,\ldots,p$) are given by (25) and r_ℓ^{++} ($\ell = 2,\ldots,p; \ell \neq m, 2m,\ldots,\alpha m$) are given by (26), with α given by (16) and n represents the common sample size of the m independent samples.

From these p.d.f.'s and c.d.f.'s is then easy to obtain the corresponding near-exact p.d.f. and c.d.f. of $\Lambda = e^{-W}$ by simple transformation.

The numerical studies carried out in the next section show that near-exact distributions built in this way, will display a very good performance, laying very close to the exact distribution and yielding a marked asymptotic behavior not only for increasing sample sizes but also for increasing values of m, k and $p = \sum_{i=1}^k p_i$. Their performance is in all cases much better than any available asymptotic distribution, with very good performances even for small sample sizes.

3.2. The unequal sample sizes case

When not all the samples have the same size, with the sample from the j-th population having size N_j ($j = 1,\ldots,m$), the problem of addressing the exact distribution of either Λ or W and getting good near-exact approximations for their distributions becomes much harder to tackle.

The problem is that in this case we cannot any more address the distribution of $W_a = -\log \Lambda_a$ in the same way that is done in (11). But, we may anyway write

$$\Phi_{W_a}(t) = \Phi^*_{1,W_a}(t)\,\frac{\Phi_{W_a}(t)}{\Phi^*_{1,W_a}(t)}\,,$$

where $\Phi^*_{1,W_a}(t)$ is $\Phi_{1,W_a}(t)$ in (11) with n replaced by N/m, for $N = \sum_{j=1}^m N_j$. Although this way to write $\Phi_{W_a}(t)$ may seem at first sight rather useless, it will indeed become much useful. It happens that, interestingly enough, when handling the different sample size case, we end up obtaining for $\Phi_{W_{b|a}}(t)$ and $\Phi_{W_{c|(b|a)}}(t)$ expressions which are in all similar to the ones in (19) and (22), with n replaced by N/m.

This way, the exact c.f. of W may, in this case of unequal sample sizes, be written in a similar way to the one in (24), with n replaced by N/m in $\Phi_1(t)$ and $\Phi_2(t)$ and $\Phi_{2,W_a}(t)$ replaced by $\Phi^*_{2,W_a}(t) = \frac{\Phi_{W_a}(t)}{\Phi^*_{1,W_a}(t)}$. Then, we will once again keep $\Phi_1(t)\Phi_2(t)$ unchanged and replace $\Phi^*_{2,W_a}(t)\,\Phi_{2,W_{b|a}}(t)\,\Phi_{2,W_{c|(b|a)}}(t)$ by the c.f. of a finite mixture of Gamma distributions, all with the same rate parameter,

$$\sum_{\nu=0}^{m^*} \pi_\nu \lambda^{r+\nu}(\lambda - \mathrm{i}t)^{-(r+\nu)}$$

as in (28), now with r equal to u_1 in (29), obtaining this way a near-exact c.f. for W, in all similar to the one in (28), with r replaced by u_1 in (29). Since in general u_1 will not be an integer, this near-exact c.f. will yield for W near-exact distributions that will correspond to mixtures of m^*+1 GNIG distributions of depth $2p-\alpha$, with p.d.f.'s and c.d.f.'s in all similar to the ones in (31) and (32), with n replaced by N/m and r replaced by u_1.

These near-exact distributions that, by construction, will match the first m^* exact moments of W, will show, as it happened in the equal sample sizes case, very good asymptotic behavior not only for increasing sample sizes but also for increasing values of p, m and even k, although exhibiting in this unequal sample case a lesser closeness to the exact distribution, but anyway still with much better performances than any available asymptotic distribution.

As we may see from the results in the next section, even better performances may be actually obtained, for larger sample sizes, if also for this unequal sample sizes case, we take for r once again the value given by (30).

4. Numerical studies

In this section we will use the measure

$$\Delta = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left|\frac{\Phi_W(t)-\Phi^*(t)}{t}\right|dt$$

where $\Phi_W(t)$ is the exact c.f. of W and $\Phi^*(t)$ represents any approximate, that is, asymptotic or near-exact, c.f. of W, to evaluate the performance of the near-exact distributions proposed in the previous section and to compare them with the asymptotic distribution proposed by Moschopoulos (1992).

This measure may be seen as based on the Berry-Esseen bound (Berry, 1941; Esseen , 1945; Loève, 1977, chap. VI; Hwang, 1998) and it satisfies the relations

$$\max_{w\in S_W}\left|F_W(w)-F^*_W(w)\right| \le \Delta \qquad \text{and} \qquad \max_{z\in S_\Lambda}\left|F_\Lambda(z)-F^*_\Lambda(z)\right| \le \Delta$$

where w and z represent respectively the running values of the r.v.'s W and Λ, S_W and S_Λ the supports of these two r.v.'s and $F_W(w)$ and $F_\Lambda(z)$ the exact c.d.f.'s of W and Λ, and $F^*_W(w)$ is the c.d.f. corresponding to $\Phi^*(t)$ and $F^*_\Lambda(z)$ the corresponding c.d.f. of Λ.

For comparison we used the asymptotic distribution in Moschopoulos (1992), which was developed for the modified l.r.t. statistic. If we adapt it to the non-modified statistic, which we use in this paper, we obtain an asymptotic c.f. for W which may be written as

$$\Phi_{Mos}(t) = \left(1-\frac{\gamma}{(m^{**})^2}\right)\left(\frac{m^{**}}{N^*}\right)^{f/2}\left(\frac{m^{**}}{N^*}-\mathrm{i}t\right)^{-f/2} + \frac{\gamma}{(m^{**})^2}\left(\frac{m^{**}}{N^*}\right)^{2+f/2}\left(\frac{m^{**}}{N^*}-\mathrm{i}t\right)^{-2-f/2}$$

where $N^* = N - m = \sum_{j=1}^{m}(N_j-1)$,

$$f = mp\frac{p+1}{2} - k\,, \qquad m^{**} = N^* - m - 2*\beta\,,$$

and

$$\gamma = -\frac{2}{3}\left(\sum_{j=1}^{m}\sum_{\ell=1}^{p}\frac{B_3\left(\beta\frac{N_j-1}{N^*}-\frac{\ell-1}{2}\right)}{((N_j-1)/N^*)^2} - \sum_{i=1}^{k}\frac{B_3(\beta p_i)}{p_i^2}\right),$$

with

$$B_3(x) = x^3 - \frac{3}{2}x^2 + \frac{x}{2} \qquad \text{and} \qquad \beta = \frac{1}{6f}\left(p(2p^2+p+1)\sum_{j=1}^{m}\frac{1}{4((N_j-1)/N^*)} - \sum_{i=1}^{k}\frac{1}{p_i}\right).$$

In the sequel we denote the corresponding asymptotic distribution by Mos.

4.1. The equal sample sizes case

In this case, as we may see from Table 1, the near-exact distributions show a quite clear asymptotic behavior for increasing values of $p = \sum_{i=1}^{k} p_i$, even for very small sample sizes, while the asymptotic distribution behaves the other way around. This asymptotic behavior is even more accentuated for the near-exact distributions that equate more exact moments and in all cases the near-exact distributions show a much better performance than the asymptotic distribution.

Table 1 – Values of the measure Δ for increasing values of p (equal sample sizes case)

							near-exact distributions number of exact moments matched		
p	p_i	k	m	n	r	Mos	4	6	10
10	{5,5}	2	2	12	5	1.65×10^{-1}	7.43×10^{-8}	8.61×10^{-10}	2.97×10^{-13}
12	{5,7}	2	2	14	6	2.70×10^{-1}	3.26×10^{-8}	2.65×10^{-10}	4.54×10^{-14}
16	{7,9}	2	2	18	8	5.15×10^{-1}	8.69×10^{-9}	3.97×10^{-11}	2.07×10^{-15}
20	{9,11}	2	2	22	10	7.62×10^{-1}	3.09×10^{-9}	9.04×10^{-12}	1.81×10^{-16}
50	{21,29}	2	2	52	25	2.15×10^{0}	3.35×10^{-11}	2.17×10^{-14}	1.05×10^{-20}

From Table 2 we may see how, opposite to the asymptotic distribution, the near-exact distributions also show a quite clear asymptotic behavior for increasing values of m, once again with the near-exact distributions that equate more moments exhibiting a more marked asymptotic response and also once again with the near-exact distribution showing in all cases a clearly better performance than the near-exact distribution.

Table 2 – Values of the measure Δ for increasing values of m (equal sample sizes case)

							near-exact distributions number of exact moments matched		
p	p_i	k	m	n	r	Mos	4	6	10
12	{5,7}	2	2	14	6	2.70×10^{-1}	3.26×10^{-8}	2.65×10^{-10}	4.54×10^{-14}
			5	14	15	5.90×10^{-1}	2.20×10^{-11}	7.10×10^{-15}	1.78×10^{-21}
			7	14	21	7.57×10^{-1}	2.08×10^{-12}	8.97×10^{-17}	1.31×10^{-23}
			10	14	30	9.65×10^{-1}	1.49×10^{-13}	4.23×10^{-17}	9.82×10^{-25}

Comparing the values of Δ in Table 3 with the values in Table 1 we may see how the near-exact distributions for the same overall value of p show a more clear asymptotic behavior than the asymptotic distribution, keeping in every case a much better performance. Actually for the larger value of p, that is, for $p = 50$, the asymptotic distribution is not any more a genuine distribution, as we may see from the value of Δ, which for genuine distributions should always be smaller than one.

Table 3 – Values of the measure Δ for increasing values of k (equal sample sizes case)

							near-exact distributions number of exact moments matched		
p	p_i	k	m	n	r	Mos	4	6	10
10	{5,3,2}	3	2	12	9/2	1.63×10^{-1}	5.91×10^{-8}	6.36×10^{-10}	1.84×10^{-13}
12	{5,5,2}	3	2	14	11/2	2.69×10^{-1}	2.74×10^{-8}	2.11×10^{-10}	3.18×10^{-14}
16	{4,3,5,4}	4	2	18	7	5.12×10^{-1}	6.73×10^{-9}	2.83×10^{-11}	1.23×10^{-15}
20	{4,5,3,3,5}	5	2	22	19/2	7.58×10^{-1}	2.81×10^{-9}	7.91×10^{-12}	1.48×10^{-16}
50	{12,9,9,7,7,6}	6	2	52	24	2.14×10^{0}	3.21×10^{-11}	2.00×10^{-14}	9.21×10^{-21}

From Tables 4 and 5 we may observe the good asymptotic behavior of both the asymptotic as well as of the near-exact distributions for increasing sample sizes, with the near-exact distributions displaying always much better, that is, much lower, values of Δ, namely for the smaller sample sizes.

Table 4 – Values of the measure Δ for increasing values of n (equal sample sizes case)

							near-exact distributions		
							number of exact moments matched		
p	p_i	k	m	n	r	Mos	4	6	10
10	{5,5}	2	2	12	5	1.65×10^{-1}	7.43×10^{-8}	8.61×10^{-10}	2.97×10^{-13}
				50	5	2.40×10^{-3}	4.88×10^{-11}	2.52×10^{-14}	1.39×10^{-20}
				100	5	7.08×10^{-5}	1.16×10^{-12}	1.42×10^{-16}	4.14×10^{-24}
				150	5	1.90×10^{-4}	1.38×10^{-13}	7.36×10^{-18}	4.03×10^{-26}

Table 5 – Values of the measure Δ for increasing values of n (equal sample sizes case)

							near-exact distributions		
							number of exact moments matched		
p	p_i	k	m	n	r	Mos	4	6	10
16	{5,5,4,2}	4	5	17	19	1.40×10^{0}	1.12×10^{-12}	8.45×10^{-17}	6.38×10^{-25}
				50	19	1.00×10^{-2}	1.03×10^{-14}	2.45×10^{-18}	2.34×10^{-26}
				100	19	7.24×10^{-3}	2.84×10^{-15}	6.29×10^{-20}	2.61×10^{-29}
				150	19	5.50×10^{-3}	5.48×10^{-16}	4.77×10^{-21}	3.34×10^{-31}

As an overall observation we would point out the good asymptotic characteristics of the near-exact distributions, concerning all the parameters in the distribution of the l.r.t. statistic being considered, with extremely good performances even for very small sample sizes. Of course these properties being extensive to the near-exact distributions relating to any particular case of the test being considered.

4.2. The unequal sample sizes case

As we may see from the observation of the values in Tables 6-10, in this case the near-exact distributions show a bit less good performance than for the case of equal sample sizes. Anyway, in all cases they still show a much better performance than the asymptotic distribution, still with very good performances even for the smaller sample sizes and with quite clear asymptotic behavior for all the parameters considered.

Table 6 – Values of the measure Δ for increasing values of p (unequal sample sizes case)

						near-exact distributions		
						number of exact moments matched		
p	p_i	k	m	N_j	Mos	4	6	10
10	{5,5}	2	2	{12,19}	9.47×10^{-2}	3.89×10^{-4}	7.25×10^{-5}	2.70×10^{-6}
12	{5,7}	2	2	{14,21}	1.52×10^{-1}	3.51×10^{-4}	6.18×10^{-5}	2.06×10^{-6}
16	{7,9}	2	2	{18,25}	2.97×10^{-1}	3.03×10^{-4}	4.91×10^{-5}	1.52×10^{-6}
20	{9,11}	2	2	{22,29}	4.71×10^{-1}	2.73×10^{-4}	4.18×10^{-5}	1.30×10^{-6}
50	{21,29}	2	2	{52,59}	1.76×10^{0}	1.89×10^{-4}	2.48×10^{-5}	8.96×10^{-7}

Table 7 – Values of the measure Δ for increasing values of *m* (equal sample sizes case)

p	p_i	k	m	N_j	Mos	near-exact distributions number of exact moments matched 4	6	10
12	{5,7}	2	2	{14,21}	1.52×10^{-1}	3.51×10^{-4}	6.18×10^{-5}	2.06×10^{-6}
			5	{14,21,28,35,42}	2.26×10^{-1}	1.52×10^{-4}	2.01×10^{-5}	8.14×10^{-7}
			7	{14,21,28,35,42,49,56}	2.49×10^{-1}	1.27×10^{-4}	1.69×10^{-5}	7.80×10^{-7}
			10	{14,21,28,35,42,49,46,63,70,77}	2.66×10^{-1}	1.00×10^{-4}	1.29×10^{-5}	5.69×10^{-7}

Table 8 – Values of the measure Δ for increasing values of *k* (unequal sample sizes case)

p	p_i	k	m	N_j	Mos	near-exact distributions number of exact moments matched 4	6	10
10	{5,3,2}	3	2	{12,19}	9.42×10^{-2}	3.32×10^{-4}	6.65×10^{-5}	2.88×10^{-6}
12	{5,5,2}	3	2	{14,21}	1.51×10^{-1}	3.07×10^{-4}	5.64×10^{-5}	1.95×10^{-6}
16	{4,3,5,4}	4	2	{18,25}	2.96×10^{-1}	2.45×10^{-4}	4.14×10^{-5}	1.20×10^{-6}
20	{4,5,3,3,5}	5	2	{22,29}	4.68×10^{-1}	2.53×10^{-4}	3.86×10^{-5}	1.08×10^{-6}
50	{12,9,9,7,7,6}	6	2	{52,59}	1.76×10^{0}	1.79×10^{-4}	2.30×10^{-5}	7.61×10^{-7}

Table 9 – Values of the measure Δ for increasing values of *n* (unequal sample sizes case)

p	p_i	k	m	N_j	Mos	near-exact distributions number of exact moments matched 4	6	10
10	{5,5}	2	2	{12,19}	9.47×10^{-2}	3.89×10^{-4}	7.25×10^{-5}	2.70×10^{-6}
				{50,60}	3.66×10^{-3}	1.62×10^{-8}	8.37×10^{-10}	8.01×10^{-12}
				{100,120}	8.81×10^{-4}	1.48×10^{-9}	8.63×10^{-11}	8.82×10^{-13}
				{150,170}	2.14×10^{-4}	2.34×10^{-10}	1.40×10^{-11}	1.46×10^{-13}

Table 10 – Values of the measure Δ for increasing values of *n* (unequal sample sizes case)

p	p_i	k	m	N_j	Mos	near-exact distributions number of exact moments matched 4	6	10
16	{5,5,4,2}	4	5	{17,24,31,38,45}	5.31×10^{-1}	2.01×10^{-3}	9.01×10^{-4}	4.36×10^{-4}
				{50,55,60,65,70}	8.90×10^{-3}	1.28×10^{-8}	7.69×10^{-11}	1.69×10^{-14}
				{100,110,120,130,140}	5.99×10^{-3}	7.14×10^{-10}	6.27×10^{-12}	1.91×10^{-15}
				{150,170,190,210,230}	4.32×10^{-3}	2.57×10^{-10}	2.49×10^{-12}	8.02×10^{-16}

Tables 6-10 report the values of Δ for the near-exact distributions with r equal to u_1 in (29). Tables 11-12 refer to near-exact distributions for which the value of r is computed from (30). By comparing the values in Tables 9-10 with the values in Tables 11-12 we may see how the near-exact distributions with r computed from (30) show a much better performance than the ones with r equal to u_1 in (29) for larger sample sizes, but not for the smaller sample sizes.

Table 11 – Values of the measure Δ for increasing values of n, with r computed from (30) (unequal sample sizes case)

						near-exact distributions		
						number of exact moments matched		
p	p_i	k	m	N_j	r	4	6	10
10	{5,5}	2	2	{12,19}	5	2.32×10^{-3}	4.60×10^{-4}	2.37×10^{-5}
				{50,60}	5	1.82×10^{-9}	5.18×10^{-12}	1.01×10^{-16}
				{100,120}	5	3.41×10^{-11}	2.01×10^{-14}	1.66×10^{-20}
				{150,170}	5	2.00×10^{-12}	4.77×10^{-16}	6.75×10^{-23}

Table 12 – Values of the measure Δ for increasing values of n, with r computed from (30) (unequal sample sizes case)

						near-exact distributions		
						number of exact moments matched		
p	p_i	k	m	N_j	r	4	6	10
16	{5,5,4,2}	4	5	{17,24,31,38,45}	19	2.67×10^{-2}	8.45×10^{-3}	8.67×10^{-4}
				{50,55,60,65,70}	19	5.31×10^{-9}	1.28×10^{-11}	1.46×10^{-16}
				{100,110,120,130,140}	19	8.33×10^{-11}	3.75×10^{-14}	1.50×10^{-20}
				{150,170,190,210,230}	19	1.33×10^{-11}	2.65×10^{-15}	1.84×10^{-22}

5. Conclusions

In this paper the authors have shown how by considering an adequate decomposition of the null hypothesis of the overall test it becomes easy to obtain the expression for the corresponding l.r.t. statistic and its moments. Even more important than this, this decomposition induces a factorization of the c.f. of the negative logarithm of the l.r.t. statistic which enables us to get a different look over the exact distribution of the test statistic, which itself then enables us to figure out a practical way to build extremely well-fitting, but yet manageable, near-exact approximations to the exact distribution of this l.r.t. statistic. These near-exact distributions are not only much better performing than any available asymptotic distribution, namely for small sample sizes and large numbers of variables, but also have quite good asymptotic behaviors for increasing numbers of variables, sets of variables and samples involved.

These near-exact distributions may be readily applied to any particular case of this test (see the Introduction section of this paper for these particular cases). Namely for $m = 1$ we have the test in Marques and Coelho (2011), and in this case, the results in this paper yield results quite similar to the ones in that reference, with the small difference that in that reference the authors used the modified l.r.t. statistic.

Also, as it is remarked in the Introduction of this paper, for $m = 1$ and $k = 1$ one obtains the usual sphericity test and as such, from the near-exact distributions in Section 3, the corresponding near-exact distributions for the sphericity l.r.t. statistic. These near-exact distributions resulting as particular cases of the ones developed in this paper, for this particular case, may be used in alternative to the ones developed in Marques and Coelho (2008) and Coelho and Marques (2010), with some advantages, which are mainly related to the fact that opposite to the near-exact distributions developed in those references, the ones developed in this paper may be virtually taken as far as one wants and the available computing power is able to handle, in order to obtain approximations which still remaining manageable become almost indistinguishable from the exact distribution.

Acknowledgments

This research was partially financially supported by CMA/FCT/UNL, under project PEst-OE/MAT/UI0297/2011.

Appendix A. Independence of statistics

The independence of the three statistics Λ_a, $\Lambda_{b|a}$ and $\Lambda_{c|(b|a)}$ in (5), under H_0 in (1) is easy to establish. We only have to note that, by Lemma 10.4.1 in (Anderson, 2003, Sec. 10.4) and the note right after expression (13) in Section 10.4 of the same reference, Λ_a is independent of

$$A = A_1 + \cdots + A_m = \sum_{j=1}^{m} A_j\,.$$

This way Λ_a is independent of both $\Lambda_{b|a}$ and $\Lambda_{c|(b|a)}$ since both these statistics are built only from A.

It remains to show that $\Lambda_{b|a}$ and $\Lambda_{c|(b|a)}$ are independent. This may be easily shown by showing that $\Lambda_{b|a}$ is independent of A_{ii} ($i = 1, \ldots, k$), the diagonal blocks of A. This fact is possible to prove through what may be seen as an extended version of Lemma 10.4.1 of (Anderson, 2003, Section 10.4) or the results in Section 8.2 of Kshirsagar (1972).

In fact we may write

$$\Lambda_{b|a} = \prod_{i=1}^{k-1} \Lambda_{b|a(i)}\,,$$

where

$$\Lambda_{b|a(i)} = \frac{|\widetilde{A}_i|^{N/2}}{|A_{ii}|^{N/2}|\widetilde{A}_{i+1}|^{N/2}} \qquad \text{with} \qquad \widetilde{A}_i = \begin{bmatrix} A_{ii} & A_{i,i+1} & \ldots & A_{ik} \\ A_{i+1,i} & A_{i+1,i+1} & \ldots & A_{i+1,k} \\ \vdots & \vdots & \vdots & \vdots \\ A_{ki} & A_{k,i+1} & \ldots & A_{kk} \end{bmatrix},$$

is the l.r.t. statistic to test the null hypothesis

$$H_{0b|0a(i)} : \bigwedge_{i'=i+1}^{k} \Sigma_{ii'} = 0\,,$$

which is the null hypothesis of independence between the i-th set of variables and the super-set formed by joining the sets $i + 1$ through k.

But then, since

$$|\widetilde{A}_i| = |A_{ii}|\,\left|\widetilde{A}_{(i+1).i}\right| = \left|\widetilde{A}_{i+1}\right|\,\left|\widetilde{A}_{i.(i+1)}\right|$$

where

$$\widetilde{A}_{(i+1).i} = \widetilde{A}_{i+1} - \widetilde{A}_{i+1,i}A_{ii}^{-1}\widetilde{A}_{i,i+1} \qquad \text{and} \qquad \widetilde{A}_{i.(i+1)} = A_{ii} - \widetilde{A}_{i,i+1}\widetilde{A}_{i+1}^{-1}\widetilde{A}_{i+1,i}$$

with

$$\widetilde{A}_{i,i+1} = \left[\, A_{i,i+1} \,|\, \cdots \,|\, A_{ik} \right] \qquad \text{and} \qquad \widetilde{A}_{i+1,i} = \widetilde{A}'_{i,i+1} \;\; \text{(where the prime denotes transpose).}$$

so that we may write

$$\Lambda_{b|a(i)} = \frac{\left|\widetilde{A}_{(i+1).i}\right|^{N/2}}{\left|\widetilde{A}_{i+1}\right|^{N/2}} = \frac{\left|\widetilde{A}_{i.(i+1)}\right|}{|A_{ii}|}\,.$$

Then, by applying to these expressions for $\Lambda_{b|a(i)}$ Lemma 10.4.1 in Anderson (2003) or the results in Section 8.2 of Kshirsagar (1972) we may see that $\Lambda_{b|a(i)}$ is independent fo both A_{ii} and $\widetilde{A}_{i+1}$, and as such, also independent of $A_{i+1,i+1}, \ldots, A_{kk}$ ($i = 1, \ldots, k-1$). This shows that not only are the statistics $\Lambda_{b|a(i)}$ ($i = 1, \ldots, k-1$) independent but also that $\Lambda_{b|a}$ is independent of $A_{11}, \ldots, A_{kk}$.

Then, since $\Lambda_{c|(b|a)}$ is only function of $A_{11}, \ldots, A_{kk}$, this statistic is independent of $\Lambda_{b|a}$.

Appendix B. The Gamma, GIG (Generalized Integer Gamma) and GNIG (Genealized Near-Integer Gamma) distributions

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

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

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 random variables with Gamma distributions with shape parameters $r_j \in \mathbf{N}$ and rate parameters $\lambda_j > 0$, with $\lambda_j \neq \lambda_{j'}$, for all $j \neq j' \in \{1, \ldots, p\}$. 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. of Y are respectively given by (see Coelho (1998))

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

and

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

where

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

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

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

and

$$c_{j,r_j-k}\;=\;\frac{1}{k}\sum_{i=1}^{k}\frac{(r_j-k+i-1)!}{(r_j-k-1)!}\,R(i,j,p)\,c_{j,r_j-(k-i)}\,,\qquad (k=1,\ldots,r_j-1;\;j=1,\ldots,p) \tag{B.2}$$

where

$$R(i,j,p)=\sum_{\substack{k=1\\k\neq j}}^{p}r_k\left(\lambda_j-\lambda_k\right)^{-i}\quad (i=1,\ldots,r_j-1)\,. \tag{B.3}$$

The GNIG distribution of depth $p+1$ (see 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. of Z is given by

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

and the c.d.f. given by

$$F^{GNIG}(z|r_1,\ldots,r_p,r;\;\lambda_1,\ldots,\lambda_p,\lambda)=\frac{\lambda^r z^r}{\Gamma(r+1)}\,{}_1F_1(r,r+1,-\lambda z) \\ -K\lambda^r\sum_{j=1}^{p}e^{-\lambda_j z}\sum_{k=1}^{r_j}c_{j,k}^*\sum_{i=0}^{k-1}\frac{z^{r+i}\lambda_j^i}{\Gamma(r+1+i)}\,{}_1F_1(r,r+1+i,-(\lambda-\lambda_j)z)\quad (z>0)$$

where

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

with $c_{j,k}$ given by (B.1)–(B.3) above. In the above expressions ${}_1F_1(a,b;z)$ is the Kummer confluent hypergeometric function. This function has usually very good convergence properties and is nowadays easily handled by a number of software packages.

References

Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis*, 3rd ed.. New York: J. Wiley & Sons.

Berry, A. (1941). The accuracy of the Gaussian approximation to the sum of independent variates. *Trans. Amer. Math. Soc.* 49:122-136.

Coelho, C. A. (1998). The Generalized Integer Gamma distribution – a basis for distributions in Multivariate Statistics. *J. Multivariate Anal.* 64:86-102.

Coelho, C. A. (1999). Addendum to the paper 'The Generalized Integer Gamma distribution – a basis for distributions in Multivariate Analysis'. *J. Multivariate Anal.* 69:281-285.

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

Coelho, C. A., 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 Appl.* 430:2592-2606.

Coelho, C. A., Marques, F. J. (2010). Near-exact distributions for the independence and sphericity likelihood ratio test statistics. *J. Multivariate Anal.* 101:583-593.

Coelho, C. A., Marques, F. J. (2011). Near-exact distributions for the likelihood ratio test statistic to test equality of several variance-covariance matrices. *Comput. Statist.* (in print, DOI: 10.1007/s00180-011-0281-1).

Esseen, C.-G. (1945). Fourier analysis of distribution functions – a mathematical study of the Laplace-Gaussian law. *Acta Math.* 77:1-125.

Hwang, H.-K. (1998). On convergence rates in the central limit theorems for com- binatorial structures. *European J. Combin.* 19:329-343.

Kshirsagar, A. M. (1972). *Multivariate Analysis*. New York: Marcel Dekker.

Loève, M. (1977). *Probability Theory*, Vol. I, 4th ed.. New York: Springer-Verlag.

Marques, F. J., Coelho, C. A. (2008). Near-exact distributions for the sphericity likelihood ratio test statistic. *J. Statist. Plann. Inference* 138:726-741.

Marques, F. J., Coelho, C. A. (2011). The block sphericity test exact and near-exact distributions for the likelihood ratio statistic. *Math. Methods Appl. Sci.* (in print, DOI: 10.1002/mma.1497).

Marques, F. J., Coelho, C. A., Arnold, B. C. (2011). A general near-exact distribution theory for the most common likelihood ratio test statistics used in Multivariate Analysis. *TEST* 20:180-203.

Moschopoulos, P.-G. (1992). The hypothesis of multisample block sphericity. *Sankhya* 54:260-270.

Muirhead, R. J. (2005). *Aspects of Multivariate Statistical Theory*, 2nd ed.. New York: J. Wiley & Sons.

Tricomi, F. G., Erdélyi, A. (1951). The asymptotic expansion of a ratio of Gamma functions. *Pacific J. Math.* 1:133-142.