NEAR-EXACT DISTRIBUTIONS FOR THE LIKELIHOOD RATIO TEST STATISTIC FOR TESTING EQUALITY OF SEVERAL VARIANCE-COVARIANCE MATRICES (REVISITED)*

By Carlos A. Coelho and Filipe J. Marques

The New University of Lisbon - Faculty for Science and Technology

The exact distribution of the l.r.t. (likelihood ratio test) statistic to test the equality of several variance-covariance matrices, as it also happens with several other l.r.t. statistics used in Multivariate Statistics, has a non-manageable form. This renders the computation of exact p-values and quantiles almost impossible, even for small numbers of variables. On the other hand, the existing asymptotic approximations do not exhibit the necessary precision for small sample sizes and their precision for larger samples is believed to be possible to be improved by using a different approach. For these reasons, the development of near-exact approximations to the distribution of this statistic, arising from an whole different method of approximating distributions, emerges as a desirable goal. From a factorization of the exact c.f. of the statistic where we adequately replace some of the factors, we obtain a near-exact c.f. which determines the near-exact distribution. This distribution, while being manageable lies much closer to the exact distribution than the available asymptotic distributions and opposite to these, is also asymptotic for increasing number of variables and matrices involved in the test. The evaluation of the performance of the distributions developed is done through the use of two measures based on the c.f.'s. Modules programmed in Mathematica are provided to compute p-values as well as the p.d.f., c.d.f. and c.f. of the near-exact distributions proposed.

1. Introduction. The test of equality of variance-covariance matrices among several populations is of interest in a large number of procedures used in multivariate analysis which have an underlying assumption of equal covariance matrices (see for example Anderson, 2003; Muirhead, 1982). Indeed, for example, methods to test the equality of mean vectors and discriminant methods, as Fisher linear discriminant function and the methods introduced by Rigby (1997) and Srivastava (1967), assume equal variance-covariance matrices for the populations . Examples of application may be

*This research was supported by the Portuguese Foundation for Science and Technology (FCT).

Keywords and phrases: asymptotic approximations, sum of Gamma distributions, product of Beta distributions, Generalized Near-Integer Gamma distribution, mixtures

large variety of fields such as genetic, psychometric, anthropometric, educational and biometrical studies (see for example Jamshidiana and Schott, 2007; Zhang and Boos, 1992). Furthermore, the test of equality of covariance matrices is also part of the test of equality of several multivariate Normal distributions.

While we strongly believe that it is not possible to improve significantly upon the existing results by using the available techniques to develop asymptotic distributions, we believe it is possible to do so by using a different approach. Furthermore, we also believe that it is possible to develop a method which, opposite to the usual asymptotic methods, besides yielding distributions that are asymptotic only in terms of sample size, may produce approximating distributions which, being manageable and much closer to the exact distribution than the usual asymptotic distributions, also display an asymptotic behavior for increasing values of other parameters (in the distribution being approximated), as the number of variables and the number of matrices involved.

It is in this context that the near-exact distributions arise, through the use of a different concept in approximating distributions.

In this paper we propose the development of near-exact distributions for the l.r.t. (likelihood ratio test) statistic to test the equality of several variance-covariance matrices and show how we may overcome the difficulties that arise in adequately factorizing the exact c.f. (characteristic function) of the logarithm of the l.r.t. statistic in this case.

The approach followed is based on factorizing adequately the exact c.f. of the logarithm of the l.r.t. statistic in a first step and then, while leaving unchanged the larger number of those factors, replace the remaining by an asymptotic function. More precisely, let us assume that the exact c.f. of $W = -\log \lambda$, where λ is the statistic of interest, may be factorized as

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

where both $\Phi_1(t)$ and $\Phi_2(t)$ are recognized as c.f.'s. Let us build $\Phi_1(t)$ in such a way that it collects the larger number of factors in $\Phi_W(t)$, while at the same time corresponds to a known manageable distribution and let $\Phi_2(t)$ correspond to a non-manageable distribution. We will replace $\Phi_2(t)$ by $\Phi_2^*(t)$ in such a way that $\Phi_2^*(t)$ and $\Phi_2(t)$ are in the same Pearson family of distributions and also in such a way that if we take

$$\Phi_W^*(n,p,q;t) = \Phi_1(t)\,\Phi_2^*(t)\,,$$

where $\Phi_W^*(n,p,q;t)$ represents the so-called near-exact c.f. of W, while function of the sample size n, the number of variables p and the number of matrices involved q, this c.f. corresponds to a near-exact distribution which is known and manageable in such a way that the computation of near-exact p-values and quantiles is rendered easy.

We should note that we also want to define $\Phi_2^*(t)$ in such a way that

$$\lim_{n\to\infty} \left| \frac{\Phi_W(t) - \Phi_W^*(n,p,q;t)}{t} \right| dt = 0\,, \tag{1.1}$$

$$\lim_{p\to\infty} \left| \frac{\Phi_W(t) - \Phi_W^*(n,p,q;t)}{t} \right| dt = 0 \qquad \text{and} \tag{1.2}$$

$$\lim_{q\to\infty} \left| \frac{\Phi_W(t) - \Phi_W^*(n,p,q;t)}{t} \right| dt = 0\,. \tag{1.3}$$

If we take $\Phi_W^*(n,p,q;t)$ as the c.f. of the r.v. W^*, (1.1), (1.2) and (1.3) above are indeed equivalent to have, respectively,

$$W^* \xrightarrow[n\to\infty]{d} W\,, \qquad W^* \xrightarrow[p\to\infty]{d} W \quad \text{and} \quad W^* \xrightarrow[q\to\infty]{d} W\,.$$

The reason why we usually factorize the c.f. of the negative logarithm of the l.r.t. statistic and not the c.f. of the l.r.t. statistic is related with the fact that usually the c.f. of $W = -\log\lambda$ is much easier to obtain and to handle than the c.f. of λ itself, upon which the factorizations we intend to carry out are not even possible to perform.

Anyway, once obtained a near-exact distribution for W, it will be very simple to obtain the corresponding near-exact distribution for $\lambda = e^{-W}$ by simple transformation.

The results in this paper, together with the results in Coelho (2004), Alberto and Coelho (2007) and Grilo and Coelho (2007) may be used to develop very accurate near-exact distributions for the l.r.t. statistic to test the equality of several multivariate Normal distributions, while on the other hand, also together with the results in Marques and Coelho (2008) may be used to establish an integrated approach, in terms of near-exact distributions, for the family of l.r.t. statistics most commonly used in multivariate statistics.

We will consider in this paper the l.r.t. statistic to test the equality of several variance-covariance matrices, under the underlying assumption of multivariate normality. We should however note that the null distribution of the l.r.t. statistic remains the same under the classes of elliptically contoured and left orthogonal-invariant distributions (Anderson, 2003; Anderson et al., 1986; Anderson and Fang, 1990; Jensen and Good, 1981; Kariya, 1981). We will show how we can factorize the c.f. of the logarithm of this statistic in two factors, one that is the c.f. of a Generalized Integer Gamma (GIG) distribution (Coelho, 1998) and the other the c.f. of a sum of independent r.v.'s (random variables) whose exponentials have Beta distributions. From this decomposition we will be able to build near-exact distributions for the logarithm of the l.r.t. statistic as well as for the statistic itself. The closeness of these approximate distributions to the exact distribution will be assessed and measured through the use of two measures, presented in Section 5, derived from inversion formulas, one of them with a link to the Berry-Esseen bound.

Let us suppose that we have q independent samples from q multivariate Normal distributions $N_p(\mu_j, \Sigma_j)$ $(j = 1, \ldots, q)$, the j-th sample having size $n + 1$, and that we want to test the null hypothesis

$$H_0 : \Sigma_1 = \Sigma_2 = \ldots = \Sigma_q \ (= \Sigma) \qquad (\Sigma \text{ unspecified}). \tag{1.4}$$

As Bartlett (1937) and Muirhead (1982, sec. 8.2) refer, by using the modified l.r.t. statistic we obtain an unbiased test. The modified l.r.t. statistic (where the sample sizes are replaced by the number of degrees of freedom of the Wishart distributions) may be written as (Anderson, 2003; Bartlett, 1937; Muirhead, 1982)

$$\lambda^* = \frac{(nq)^{npq/2}}{\prod_{j=1}^{q} n^{pn/2}} \, \frac{\prod_{j=1}^{q} |A_j|^{n/2}}{|A|^{nq/2}}\,, \tag{1.5}$$

where A_j is the matrix of corrected sums of squares and products formed from the j-th sample and $A = A_1 + \cdots + A_q$.

The h-th moment of λ^* in (1.5) is (Muirhead, 1982)

$$\begin{aligned} E\left[(\lambda^*)^h\right] &= q^{npqh/2} \prod_{j=1}^{p} \frac{\Gamma\left(\frac{nq+1-j}{2}\right)}{\Gamma\left(\frac{nq+1-j}{2} + \frac{nq}{2}h\right)} \\ &\quad \times \prod_{j=1}^{p}\prod_{k=1}^{q} \frac{\Gamma\left(\frac{n}{2}+\frac{1}{2}-\frac{j}{2}+\frac{n}{2}h\right)}{\Gamma\left(\frac{n}{2}+\frac{1}{2}-\frac{j}{2}\right)} \qquad \left(h > \tfrac{p-n-1}{n}\right). \end{aligned} \tag{1.6}$$

From (1.6) we may write the c.f. for the r.v. $W = -\log(\lambda^*)$ as

$$\begin{aligned} \Phi_W(t) &= E\left[e^{\mathrm{i}tW}\right] = E\left[(\lambda^*)^{-\mathrm{i}t}\right] \\ &= q^{-npq\mathrm{i}t/2} \prod_{j=1}^{p} \frac{\Gamma\left(\frac{nq+1-j}{2}\right)}{\Gamma\left(\frac{nq}{2+1-j} - \frac{nq}{2}\mathrm{i}t\right)} \\ &\qquad \times \prod_{j=1}^{p}\prod_{k=1}^{q} \frac{\Gamma\left(\frac{n+1-j}{2} - \frac{n}{2}\mathrm{i}t\right)}{\Gamma\left(\frac{n+1-j}{2}\right)}\,. \end{aligned} \tag{1.7}$$

It will be based on this expression that we will obtain, in Section 3, decompositions of the c.f. of W that will be used to build near-exact distributions for W and λ^*.

2. Asymptotic distributions. Box (1949) proposes for the statistic $W = -\log(\lambda^*)$ an asymptotic distribution based on an expansion of the form

$$P(2\rho W \leq z) = (1-\omega)P(\chi^2_g \leq z) + \omega P(\chi^2_{g+4} \leq z) + O\left((nq)^{-3}\right)$$

where

$$g = \frac{1}{2}(q-1)p(p+1)\,, \qquad \rho = 1 - \frac{q+1}{nq}\,\frac{2p^2+3p-1}{6(p+1)}$$

and

$$\omega = \frac{1}{48\rho^2}p(p+1)\left\{(p-1)(p-2)\frac{q^3-1}{n^2q^2} - 6(q-1)(1-\rho)^2\right\}$$

and where $P(\chi^2_g \leq z)$ stands for the value of the c.d.f. (cumulative distribution function) of a chi-square r.v. with g degrees of freedom evaluated at $z(>0)$.

However, taking into account that

$$X \sim \chi^2_g \equiv \Gamma\left(\frac{g}{2},\frac{1}{2}\right) \qquad \Longrightarrow \qquad \frac{X}{2\rho} \sim \Gamma\left(\frac{g}{2},\rho\right)\,,$$

where we use the notation

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

to denote the fact that the r.v. X has p.d.f. (probability density function)

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

we may write

$$P(W \leq z) \;\approx\; (1-\omega)P\left(\Gamma\left(\frac{g}{2},\rho\right) \leq z\right) + \omega\, P\left(\Gamma\left(\frac{g}{2}+2,\rho\right) \leq z\right),$$

where $P\left(\Gamma\left(\nu,\rho\right) \leq z\right)$ stands for the value of the c.d.f. of a $\Gamma(\nu,\rho)$ distributed r.v. evaluated at $z(>0)$.

We may thus write, for the c.f. of W,

$$\text{(2.1)} \quad \Phi_W(t) \approx \Phi_{Box}(t) = (1-\omega)\rho^{g/2}(\rho-\mathrm{i}t)^{-g/2} + \omega\,\rho^{2+g/2}(\rho-\mathrm{i}t)^{-2-g/2}\,.$$

We should note that actually, in many cases where n is close to p, Box's asymptotic distribution does not even correspond to a true distribution (see Appendix B).

Somehow inspired on this asymptotic approximation due to Box, which ultimately approximates the exact c.f. of W by the c.f. of a mixture of Gamma distributions and also on Box's introduction of his 1949 paper (Box, 1949), where he states that "Although in many cases the exact distributions cannot be obtained in a form which is of practical use, it is usually possible to obtain the moments, and these may be used to obtain approximations. In some cases, for instance, a suitable power of the likelihood statistic has been found to be distributed approximately in the type I form, and good approximations have been obtained by equating the moments of the likelihood statistic to this curve.", we propose two other mixtures of Gamma distributions, all with the same rate parameter, which match the first four or six exact moments to approximate asymptotically the c.f. of $W = -\log(\lambda^*)$.

These distributions are: the mixture of two Gamma distributions ($M2G$), both with the same rate parameter, with c.f.

$$\Phi_{M2G}(t) = \sum_{j=1}^{2} p_{2,j}\, \lambda_2^{r_{2,j}} (\lambda_2 - \mathrm{i}t)^{-r_{2,j}}\,, \tag{2.2}$$

where $p_{2,2} = 1 - p_{2,1}$ with $p_{2,j}, r_{2,j}, \lambda_2 > 0$, and the mixture of three Gamma distributions ($M3G$), all with the same rate parameter, with c.f.

$$\Phi_{M3G}(t) = \sum_{j=1}^{3} p_{3,j}\, \lambda_3^{r_{3,j}} (\lambda_3 - \mathrm{i}t)^{-r_{3,j}}\,, \tag{2.3}$$

where $p_{3,3} = 1 - p_{3,1} - p_{3,2}$, with $p_{3,j}, r_{3,j}, \lambda_3 > 0$.

The parameters in (2.2) and (2.3) are respectively obtained by solving the systems of equations

$$\left.\frac{\partial^h \Phi_{MkG}(t)}{\partial t^h}\right|_{t=0} = \mathrm{i}^h \sum_{j=1}^{k} p_{k,j} \frac{\Gamma(r_{k,j}+h)}{\Gamma(r_{k,j})}\, \lambda_k^{-h} = \left.\frac{\partial^h \Phi_W(t)}{\partial t^h}\right|_{t=0}, \tag{2.4}$$

for $h = 1, \ldots, 2k$, with $k = 2$ for the parameters in (2.2) and $k = 3$ for the parameters in (2.3).

Actually, since, as shown in Lemma 1 in Appendix A, $\Phi_W(t)$ is the c.f. of a sum of independent log Beta r.v.'s, the approximation of the distribution of W by a mixture of Gamma distributions is a well justified procedure, since as Coelho et al. (2006) proved, a log Beta distribution may be represented as an infinite mixture of Exponential distributions, and as such a sum of independent log Beta r.v.'s may be represented as an infinite mixture of sums of independent Exponential distributions, which are particular Generalized Integer Gamma distributions (Coelho, 1998). Thus, the use of a finite mixture of Gamma distributions to replace a log Beta distribution seems to be a much adequate simplification.

We will also use as asymptotic distributions the Gamma distribution from Jensen (1991, 1995) which matches the two first moments of W. This asymptotic distribution will indeed always be outperformed by the $M2G$ and $M3G$ asymptotic distributions since it may be seen as the single component member of the Gamma mixture family that comprises $M2G$ and $M3G$. We will also use as a reference the asymptotic saddle-point approximation from Jensen (1991, 1995).

3. The characteristic function of $W = -\log(\lambda^*)$. In this section we will present two results that will enable us to obtain a factorization of the c.f. of W which will be used to build near-exact distributions for W and λ^*.

In a first step we will show how the c.f. of $W = -\log(\lambda^*)$ may be factorized in two factors, one of them being the c.f. of the sum of independent Logbeta r.v.'s (multiplied by n or $n/2$) and the other the c.f. of the sum of independent Exponential r.v.'s.

In a second step we will identify the different Exponential distributions involved and devise a method to count them and to obtain analytic expressions for the number of times each different Exponential distribution appears in order to convert this part of the c.f. of W into the c.f. of a sum of independent Gamma distributions with integer shape parameters, that is, the c.f. of a GIG distribution, identifying the depth of this distribution.

Corresponding to the first step above, we have the next Theorem.

THEOREM 3.1. *The characteristic function of $W = -\log\lambda^*$ may be written as*

$$\Phi_W(t) = E\left(e^{\mathrm{i}tW}\right) = \Phi_1(t)\,\Phi_2(t)\,,$$

where

$$\Phi_1(t) = \left\{ \prod_{j=1}^{\lfloor p/2\rfloor} \prod_{k=1}^{q} \prod_{l=1-\lfloor\frac{k-2j}{q}\rfloor}^{2j-1} \frac{(n-l)/n}{(n-l)/n - \mathrm{i}t} \right\}$$

$$\times \left\{ \prod_{k=1}^{q} \prod_{l=1}^{b^*_{pk}} \frac{2(a_p+l-1)/n}{2(a_p+l-1)/n - \mathrm{i}t} \right\}^{p\perp 2}$$

and

$$\Phi_2(t) = \left\{ \prod_{j=1}^{\lfloor p/2 \rfloor} \prod_{k=1}^{q} \frac{\Gamma(a_j+b_{jk})}{\Gamma(a_j+b^*_{jk})} \frac{\Gamma(a_j+b^*_{jk}-n\mathrm{i}t)}{\Gamma(a_j+b_{jk}-n\mathrm{i}t)} \right\}$$

$$\times \left\{ \prod_{k=1}^{q} \frac{\Gamma(a_p+b_{pk})}{\Gamma(a_p+b^*_{pk})} \frac{\Gamma(a_p+b^*_{pk}-\frac{n}{2}\mathrm{i}t)}{\Gamma\left(a_p+b_{pk}-\frac{n}{2}\mathrm{i}t\right)} \right\}^{p \perp 2}$$

with

$$a_j = n+1-2j\,, \qquad b_{jk} = 2j-1+\tfrac{k-2j}{q}\,, \qquad b^*_{jk} = \lfloor b_{jk} \rfloor\,, \tag{3.1}$$

$$a_p = \frac{n+1-p}{2}\,, \qquad b_{pk} = \tfrac{p-1}{2}+\tfrac{k}{q}-\tfrac{p+1}{2q}\,, \qquad b^*_{pk} = \lfloor b_{pk} \rfloor\,, \tag{3.2}$$

for $j = 1, \ldots, \lfloor \frac{p}{2} \rfloor$ *and* $k = 1, \ldots, q$, *and where*

$$p \perp\!\!\!\perp 2 = p - 2\left\lfloor \frac{p}{2} \right\rfloor = Mod(p,2)\,.$$

PROOF. See Appendix A. □

Corresponding to the second step described above, we have the following Theorem.

THEOREM 3.2. *We may write for* $\Phi_1(t)$ *in Theorem 3.1,*

$$\Phi_1(t) = \prod_{k=1}^{p-1} \left(\frac{(n-k)/n}{(n-k)/n - \mathrm{i}t} \right)^{r_k}\,,$$

where for $\alpha = \left\lfloor \frac{p-1}{q} \right\rfloor$ *and*

$$\alpha_1 = \left\lfloor \frac{q-1}{q}\,\frac{p-1}{2} \right\rfloor, \qquad \alpha_2 = \left\lfloor \frac{q-1}{q}\,\frac{p+1}{2} \right\rfloor,$$

$$r_k = \begin{cases} r^*_k & \text{for } k = 1, \ldots, p-1, \\ & \textit{and } k \neq p-1-2\alpha_1 \\ r^*_k + a^* & \text{for } k = p-1-2\alpha_1 \end{cases} \tag{3.3}$$

with

$$a^* = (p \perp\!\!\!\perp 2)(\alpha_2 - \alpha_1)\left(q - \frac{p-1}{2} + q \left\lfloor \frac{p}{2q} \right\rfloor \right)$$

and

$$r_k^* = \begin{cases} c_k, & k \in \{1,\dots,\alpha+1\} \\ q\left(\lfloor \frac{p}{2} \rfloor - \lfloor \frac{k}{2} \rfloor\right), & k \in \{\alpha+2,\dots,\min(p-2\alpha_1,p-1)\}, \\ & k \in \{2+p-2\alpha_1,\dots,2\lfloor p/2 \rfloor -1;\ \text{step } 2\} \\ q\left(\lfloor \frac{p+1}{2} \rfloor - \lfloor \frac{k}{2} \rfloor\right), & k \in \{1+p-2\alpha_1,\dots,p-1;\ \text{step } 2\}\ , \end{cases} \tag{3.4}$$

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

$$c_k = \left\lfloor \frac{q}{2} \right\rfloor \left((k-1)q - 2\left((q+1) \perp\!\!\!\perp 2\right)\left\lfloor \frac{k}{2} \right\rfloor\right) + \left\lfloor \frac{q}{2} \right\rfloor \left\lfloor \frac{q+k \perp\!\!\!\perp 2}{2} \right\rfloor$$

and

$$\begin{aligned} c_{\alpha+1} = -\left(\left\lfloor \frac{p}{2} \right\rfloor - \alpha \left\lfloor \frac{q}{2} \right\rfloor\right)^2 + q\left(\left\lfloor \frac{p}{2} \right\rfloor - \left\lfloor \frac{\alpha+1}{2} \right\rfloor\right) \\ +(q \perp\!\!\!\perp 2)\left(\alpha \left\lfloor \frac{p}{2} \right\rfloor + \frac{\alpha \perp\!\!\!\perp 2}{4} + \frac{\alpha^2}{4} - \alpha^2 \frac{q}{2}\right). \end{aligned} \tag{3.5}$$

Proof. See Appendix A. □

From the results in Theorems 3.1 and 3.2 we may write

$$\begin{aligned} \Phi_W(t) \;=\; & \underbrace{\prod_{k=1}^{p-1} \left(\frac{(n-k)/n}{(n-k)/n - \mathrm{i}t}\right)^{r_k}}_{\Phi_1(t)} \\ & \underbrace{\times \prod_{j=1}^{\lfloor p/2 \rfloor} \prod_{k=1}^{q} \frac{\Gamma(a_j + b_{jk})}{\Gamma(a_j + b_{jk}^*)} \frac{\Gamma(a_j + b_{jk}^* - n\mathrm{i}t)}{\Gamma(a_j + b_{jk} - n\mathrm{i}t)} \times \left(\prod_{k=1}^{q} \frac{\Gamma(a_p + b_{pk})}{\Gamma(a_p + b_{pk}^*)} \frac{\Gamma(a_p + b_{pk}^* - \frac{n}{2}\mathrm{i}t)}{\Gamma\left(a_p + b_{pk} - \frac{n}{2}\mathrm{i}t\right)}\right)^{p \perp 2}}_{\Phi_2(t)}, \end{aligned} \tag{3.6}$$

for a_j, b_{jk}, a_p, b_{pk}, b_{jk}^* and b_{pk}^* defined in (3.1)-(3.2) and where $\Phi_1(t)$ is the c.f. of a GIG a distribution of depth $p-1$ and $\Phi_2(t)$ is the c.f. of the sum of independent Logbeta r.v.'s (some of them multiplied by n and other multiplied by $n/2$).

It will be based on this expression of the c.f. of W that we will develop the near-exact distributions in the next section.

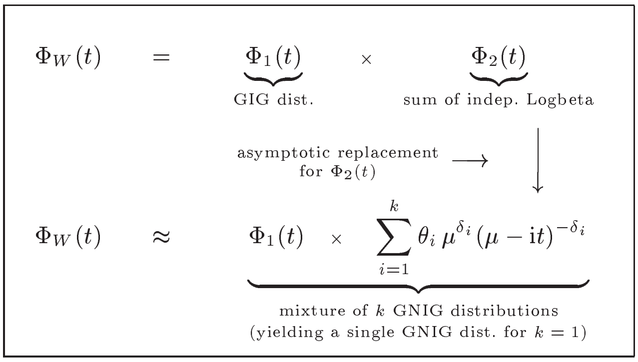

FIG 1. *How the near-exact distributions are built.*

4. Near-exact distributions for W and λ^*. In all cases where we are able to factorize a c.f. into two factors, one of which corresponds to a manageable well known distribution and the other to a distribution that although giving us some problems in terms of being convoluted with the first factor, may however be adequately asymptotically replaced by another c.f. in such a way that the overall c.f. obtained by leaving the first factor unchanged and adequately replacing the second factor may then correspond to a known manageable distribution. This way we will be able to obtain what we call a near-exact c.f. for the r.v. under study.

This is exactly what happens with the c.f. of W. In this section we will show how by keeping $\Phi_1(t)$ in the c.f. of W in (3.6) unchanged and replacing $\Phi_2(t)$ by the c.f. of a Gamma distribution or the mixture of two or three Gamma distributions, matching the first two, four or six derivatives of $\Phi_2(t)$ in order to t at $t = 0$, we will be able to obtain high quality near-exact distributions for W under the form of a GNIG (Generalized Near-Integer Gamma) distribution or mixtures of GNIG distributions (see Figure 1). From these distributions we may then easily obtain near-exact distributions for $\lambda^* = e^{-W}$.

The GNIG distribution of depth $g + 1$ (Coelho, 2004) is the distribution of the r.v.

$$Z = Y + \sum_{i=1}^{g} X_i$$

where the $g + 1$ r.v.'s Y and X_i $(i = 1, \ldots, g)$ are all independent with Gamma distributions, Y with shape parameter r, a positive non-integer, and rate parameter λ and each X_i $(i = 1, \ldots, g)$ with an integer shape parameter r_i and rate parameter λ_i, being all the $g + 1$ rate parameters different. The p.d.f. of Z is given by

$$
\begin{aligned}
&f_Z(z|r_1,\dots,r_g,r;\,\lambda_1,\dots,\lambda_g,\lambda) = \\
&\qquad K\lambda^r \sum_{j=1}^{g} \mathrm{e}^{-\lambda_j z} \sum_{k=1}^{r_j} \Bigg\{ c_{j,k} \frac{\Gamma(k)}{\Gamma(k+r)} z^{k+r-1} \\
&\qquad\qquad {}_1F_1(r,k+r,-(\lambda-\lambda_j)z) \Bigg\},
\end{aligned} \tag{4.1}
$$

for $z > 0$, and the c.d.f. given by

$$
\begin{aligned}
F_Z(z|r_1,\dots,r_g,r;\,\lambda_1,\dots,\lambda_g,\lambda) &= \lambda^r \frac{z^r}{\Gamma(r+1)} {}_1F_1(r,r+1,-\lambda z) \\
&-K\lambda^r \sum_{j=1}^{g} \mathrm{e}^{-\lambda_j z} \sum_{k=1}^{r_j} c^*_{j,k} \sum_{i=0}^{k-1} \Bigg\{ \frac{z^{r+i}\lambda_j^i}{\Gamma(r+1+i)} \\
&\qquad {}_1F_1(r,r+1+i,-(\lambda-\lambda_j)z) \Bigg\}
\end{aligned} \tag{4.2}
$$

also for $z > 0$, where

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

with $c_{j,k}$ given by (11) through (13) in Coelho (1998). 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.

In the next Theorem we develop near-exact distributions for W under the form of a single $GNIG$ distribution or a mixture of two or three $GNIG$ distributions, using the procedure described in Figure 1.

THEOREM 4.1. *Using for $\Phi_2(t)$ in the characteristic function of $W = -\log\lambda$ in (3.6), the approximations:*

i) $\lambda^s(\lambda-\mathrm{i}t)^{-s}$ *with $s,\lambda>0$, such that, for $h=1,2$,*

$$
\left.\frac{\partial^h}{\partial t^h}\,\lambda^s(\lambda-it)^{-s}\right|_{t=0} = \left.\frac{\partial^h}{\partial t^h}\,\Phi_2(t)\right|_{t=0}; \tag{4.3}
$$

ii) $\sum_{k=1}^{2}\theta_k\,\mu^{s_k}(\mu-\mathrm{i}t)^{-s_k}$, *where $\theta_2 = 1-\theta_1$ with θ_k, s_k, $\mu>0$, such that, for $h=1,\dots,4$,*

$$
\left.\frac{\partial^h}{\partial t^h}\sum_{k=1}^{2}\theta_k\left(\frac{\mu}{\mu-\mathrm{i}t}\right)^{s_k}\right|_{t=0} = \left.\frac{\partial^h}{\partial t^h}\,\Phi_2(t)\right|_{t=0}; \tag{4.4}
$$

iii) $\sum_{k=1}^{3} \theta_k^* \, \nu^{s_k^*} (\nu - \mathrm{i}t)^{-s_k^*}$, *where* $\theta_3^* = 1 - \theta_1^* - \theta_2^*$ *with* θ_k^*, $s_k^*, \nu > 0$, *such that, for* $h = 1, \ldots, 6$,

$$\frac{\partial^h}{\partial t^h} \sum_{k=1}^{3} \theta_k^* \, \nu^{s_k^*} (\nu - \mathrm{i}t)^{-s_k^*} \bigg|_{t=0} = \frac{\partial^h}{\partial t^h} \, \Phi_2(t) \bigg|_{t=0} ; \tag{4.5}$$

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

i) *a GNIG distribution of depth* p *with c.d.f.*

$$F(w|r_1, \ldots, r_{p-1}, s; \lambda_1, \ldots, \lambda_{p-1}, \lambda) \, , \tag{4.6}$$

where the r_j *(*$j = 1, \ldots, p-1$*) are given in* (3.3) *and*

$$\lambda_j = \frac{n-j}{n} \, , \quad (j = 1, \ldots, p-1), \tag{4.7}$$

and

$$\lambda = \frac{m_1}{m_2 - m_1^2} \qquad \text{and} \qquad s = \frac{m_1^2}{m_2 - m_1^2} \tag{4.8}$$

with

$$m_h = i^{-h} \, \frac{\partial^h}{\partial t^h} \, \Phi_2(t) \bigg|_{t=0} \, , \qquad h = 1, 2 \, ;$$

ii) *a mixture of two GNIG distributions of depth* p*, with c.d.f.*

$$\sum_{k=1}^{2} \theta_k \, F(w|r_1, \ldots, r_{p-1}, s_k; \lambda_1, \ldots, \lambda_{p-1}, \mu) \, , \tag{4.9}$$

where r_j *and* λ_j *(*$j = 1, \ldots, p-1$*) are given in* (3.3) *and* (4.7) *and* θ_1, μ, s_1 *and* s_2 *are obtained from the numerical solution of the system of four equations*

$$\sum_{k=1}^{2} \theta_k \, \frac{\Gamma(s_k + h)}{\Gamma(s_k)} \, \mu^{-h} = i^{-h} \, \frac{\partial^h}{\partial t^h} \, \Phi_2(t) \bigg|_{t=0} \tag{4.10}$$

$$(h = 1, \ldots, 4)$$

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

iii) or a mixture of three GNIG distributions of depth $p-1$, with c.d.f.

$$\sum_{k=1}^{3} \theta_k^* \, F(w|r_1,\ldots,r_{p-2},s_k^*;\lambda_1,\ldots,\lambda_{p-2},\nu)\,, \tag{4.11}$$

with r_j and λ_j ($j=1,\ldots,p-2$) given by (3.3) and (4.7) and θ_1^*, θ_2^*, ν, s_1^*, s_2^* and s_3^* obtained from the numerical solution of the system of six equations

$$\sum_{k=1}^{3} \theta_j^* \frac{\Gamma(s_k^*+h)}{\Gamma(s_k^*)}\,\nu^{-h} = i^{-h}\frac{\partial^h}{\partial t^h}\,\Phi_2(t)\bigg|_{t=0} \qquad (h=1,\ldots,6) \tag{4.12}$$

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

Proof. If in the c.f. of W we replace $\Phi_2(t)$ by $\lambda^s(\lambda-\mathrm{i}t)^{-s}$ we obtain

$$\Phi_W(t) \;\approx\; \lambda^s(\lambda-\mathrm{i}t)^{-s}\underbrace{\prod_{k=1}^{p-1}\left(\frac{(n-k)/n}{(n-k)/n-\mathrm{i}t}\right)^{r_k}}_{\Phi_1(t)}\,,$$

that is the c.f. of the sum of $p-1$ independent Gamma r.v.'s, $p-2$ of which with integer shape parameters r_j and rate parameters λ_j given by (3.3) and (4.7), and a further Gamma r.v. with rate parameter $s>0$ and shape parameter λ. This c.f. is thus the c.f. of the GNIG distribution of depth p with distribution function given in (4.6). The parameters s and λ are determined in such a way that (4.3) holds. This compels s and λ to be given by (4.8) and makes the two first moments of this near-exact distribution for W to be the same as the two first exact moments of W.

If in the c.f. of W we replace $\Phi_2(t)$ by $\sum_{k=1}^{2}\theta_k\,\mu^{s_k}(\mu-\mathrm{i}t)^{-s_k}$ we obtain

$$\Phi_W(t) \approx \sum_{k=1}^{2}\theta_k\,\mu^{s_k}(\mu-\mathrm{i}t)^{-s_k}\times\underbrace{\prod_{k=1}^{p-1}\left(\frac{(n-k)/n}{(n-k)/n-\mathrm{i}t}\right)^{r_k}}_{\Phi_1(t)}\,,$$

that is the c.f. of the mixture of two GNIG distributions of depth p with distribution function given in (4.9). The parameters θ_1, μ, s_1 and s_2 are defined in such a way that (4.4) holds, giving rise to the evaluation of these parameters as the numerical solution of the system of equations in (4.4) and to a near-exact distribution that matches the first four exact moments of W.

If in the c.f. of W we replace $\Phi_2(t)$ by $\sum_{k=1}^{3} \theta_k^* \nu^{s_k^*}(\nu-\mathrm{i}t)^{-s_k^*}$ we obtain

$$\Phi_W(t) \approx \sum_{k=1}^{3} \theta_k^* \, \nu^{s_k^*}(\nu-\mathrm{i}t)^{-s_k^*} \times \underbrace{\prod_{k=1}^{p-1} \left(\frac{(n-k)/n}{(n-k)/n-\mathrm{i}t} \right)^{r_k}}_{\Phi_1(t)} ,$$

that is the c.f. of the mixture of three GNIG distributions of depth p with distribution function given in (4.11). The parameters θ_1^*, θ_2^*, ν, s_1^*, s_2^* and s_3^* are defined in such a way that (4.5) holds, what gives rise to the evaluation of these parameters as the numerical solution of the system of equations in (4.5), giving rise to a near-exact distribution that matches the first six exact moments of W. □

We should note here that the replacement of the c.f. of a sum of Logbeta r.v.'s by the c.f. of a single Gamma r.v. or the c.f. of a mixture of two or three of such r.v.'s has already been well justified at the end of Section 2.

Corollary. *Distributions with c.d.f.'s given by*

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

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

or

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

where the parameters are the same as in Theorem 3.1, and $0<z<1$ *represents the running value of the statistic* $\lambda^*=e^{-W}$*, may be used as near-exact distributions for this statistic.*

Proof. Since the near-exact distributions developed in Theorem 4.1 were for the r.v. $W=-\log(\lambda^*)$ we only need to mind the relation

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

where $F_{\lambda^*}(\cdot)$ is the c.d.f. of λ^* and $F_W(\cdot)$ is the c.d.f. of W, in order to obtain the corresponding near-exact distributions for λ^*. □

We should also stress that although we advocate the numerical solution of systems of equations (2.4), (4.10) and (4.12), the remarks at the end of Section 3 in Marques and Coelho (2008), also apply here.

5. Numerical studies. In order to evaluate the quality of the approximations developed we use two measures of proximity between c.f.'s which are also measures of proximity between distribution functions or densities.

Let Y be a continuous r.v. defined on S with distribution function $F_Y(y)$, density function $f_Y(y)$ and c.f. $\phi_Y(t)$, and let $\phi_n(t)$, $F_n(y)$ and $f_n(y)$ be respectively the characteristic, distribution and density function of a r.v. X_n. The two measures are

$$\Delta_1 = \frac{1}{2\pi}\int_{-\infty}^{\infty} |\phi_Y(t) - \phi_n(t)|\,\mathrm{d}t$$

and,

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

with

$$\max_{y\in S}|f_Y(y) - f_n(y)| \leq \Delta_1 \quad \text{and} \quad \max_{y\in S}|F_Y(y) - F_n(y)| \leq \Delta_2\,. \tag{5.1}$$

We should note that for continuous r.v.'s,

$$\lim_{n\to\infty} \Delta_1 = 0 \iff \lim_{n\to\infty} \Delta_2 = 0 \tag{5.2}$$

and either one of the limits above imply that

$$X_n \xrightarrow{d} Y\,. \tag{5.3}$$

Indeed both measures and both relations in (5.1) may be derived directly from inversion formulas, and Δ_2 may be seen as based on the Berry-Esseen upper bound on $|F_Y(y) - F_n(y)|$ (Berry, 1941; Esseen, 1945; Loève, 1977, Chap. VI, sec. 21; Hwang, 1998) which may, for any $b > 1/(2\pi)$ and any $T > 0$, be written as

$$\max_{y\in S}|F_Y(y) - F_n(y)| \leq b\int_{-T}^{T}\left|\frac{\phi_Y(t) - \phi_n(t)}{t}\right|\,dt + \frac{C(b)M}{T} \tag{5.4}$$

where $M = max_{y\in S} f_n(y)$ and $C(b)$ is a positive constant that only depends of b. If in (5.4) above we take $T \to \infty$ then we will have Δ_2, since then we may take $b = 1/(2\pi)$.

These measures were used by Grilo and Coelho (2007) to study near-exact approximations to the distribution of the product of independent Beta r.v.'s and by Marques and Coelho (2008) to the study of near-exact distributions for the sphericity l.r.t. statistic.

In this section we will denote Box's asymptotic distribution by 'Box', by 'G', 'M2G' and 'M3G' respectively the asymptotic Gamma distribution developed by Jensen (1991, 1995) and the mixture of two and three Gamma distributions proposed in Section 2 and by 'GNIG', 'M2GNIG' and 'M3GNIG' the near-exact single GNIG distribution and the mixtures of two and three GNIG distributions.

We show in this section the Tables for values of Δ_2 while in Appendix C are the corresponding Tables for the values of Δ_1.

In Table 1, we may see for increasing p (number of variables), with the sample size remaining close to p ($n-p=2$), the continuous degradation (increase) of the values of the measure Δ_2 for Box's asymptotic distribution, even with a value which does not make much sense for $p=50$ (since Δ_2 is an upper-bound on the absolute value of the difference between the approximate and the exact c.d.f.). Also the three asymptotic distributions G, M2G and M3G show slightly increasing values for Δ_2, although remaining within much low values. For the distribution M3G it was not possible to obtain, for $p=20$ and $p=50$, stable solutions for the system of equations (2.4) for $k=3$. This distribution, for values of $p \leq 7$ even shows lower values of Δ_2 than the GNIG near-exact distribution.

Opposite to this behavior, all three near-exact distributions show a sharp improvement (decrease) in their values for Δ_2 for values of p large enough, showing the asymptotic character of these distributions for increasing values of p.

Table 7 in Appendix B, has the corresponding values for Δ_1 and would lead us to draw similar conclusions.

In Tables 2 and 8 we may see the clear asymptotic character of the near-exact distributions for increasing values of q (the number of matrices being tested), with a similar but less marked behavior of the G, M2G and M3G asymptotic distributions, opposite to what happens with Box's asymptotic distribution.

In Tables 3 and 9 we may see how, for increasing sample sizes, the asymptotic character is stronger for the near-exact distributions than for the asymptotic distributions. This asymptotic character being more marked for the near-exact distributions based on mixtures. However, the M3G asymptotic distribution for $p=7$, $q=2$ and $n=50$ even beats the M2GNIG near-exact distribution and the M2G asymptotic distribution performs bet-

TABLE 1
Values of the measure Δ_2 for increasing values of p, with small sample sizes

p	q	n	Box	G	M2G	M3G	GNIG	M2GNIG	M3GNIG
3	2	5	1.364×10^{-1}	1.737×10^{-3}	2.601×10^{-4}	2.826×10^{-5}	3.461×10^{-4}	4.503×10^{-6}	4.503×10^{-6}
4	2	6	2.058×10^{-1}	2.547×10^{-3}	1.417×10^{-4}	9.342×10^{-6}	9.733×10^{-5}	1.623×10^{-7}	9.621×10^{-10}
5	2	7	2.690×10^{-1}	3.036×10^{-3}	2.092×10^{-4}	1.940×10^{-5}	1.952×10^{-5}	2.261×10^{-7}	9.605×10^{-9}
6	2	8	3.256×10^{-1}	3.377×10^{-3}	2.601×10^{-4}	2.826×10^{-5}	6.091×10^{-5}	4.167×10^{-7}	3.404×10^{-9}
7	2	9	3.762×10^{-1}	3.632×10^{-3}	3.000×10^{-4}	3.579×10^{-5}	1.022×10^{-4}	1.037×10^{-6}	1.537×10^{-8}
8	2	10	4.214×10^{-1}	3.831×10^{-3}	3.319×10^{-4}	4.216×10^{-5}	3.493×10^{-5}	1.906×10^{-7}	1.492×10^{-9}
9	2	11	4.621×10^{-1}	3.988×10^{-3}	3.579×10^{-4}	4.755×10^{-5}	2.244×10^{-5}	6.825×10^{-8}	1.940×10^{-10}
10	2	12	4.986×10^{-1}	4.116×10^{-3}	3.794×10^{-4}	5.213×10^{-5}	2.141×10^{-5}	8.573×10^{-8}	5.282×10^{-10}
20	2	22	7.099×10^{-1}	4.658×10^{-3}	4.749×10^{-4}	——×—	4.070×10^{-6}	4.854×10^{-9}	9.293×10^{-12}
35	2	37	9.253×10^{-1}	4.758×10^{-3}	4.950×10^{-4}	——×—	1.208×10^{-6}	5.869×10^{-10}	4.510×10^{-13}
50	2	52	1.286×10^{0}	4.701×10^{-3}	4.864×10^{-4}	——×—	4.184×10^{-7}	9.503×10^{-11}	3.325×10^{-14}

TABLE 2
Values of the measure Δ_2 for increasing values of q, with small sample sizes

p	q	n	Box	G	M2G	M3G	GNIG	M2GNIG	M3GNIG
3	2	5	1.364×10^{-1}	1.737×10^{-3}	2.601×10^{-4}	2.826×10^{-5}	3.461×10^{-4}	4.503×10^{-6}	4.503×10^{-6}
3	5	5	1.669×10^{-1}	1.157×10^{-3}	3.947×10^{-5}	1.697×10^{-6}	4.234×10^{-4}	8.892×10^{-6}	3.004×10^{-7}
3	7	5	1.869×10^{-1}	1.003×10^{-3}	3.283×10^{-5}	1.423×10^{-6}	2.870×10^{-4}	4.725×10^{-6}	1.320×10^{-7}
3	10	5	2.143×10^{-1}	8.566×10^{-4}	2.633×10^{-5}	1.111×10^{-6}	1.796×10^{-4}	2.168×10^{-6}	4.610×10^{-8}
10	2	12	4.986×10^{-1}	4.116×10^{-3}	3.794×10^{-4}	5.213×10^{-5}	2.141×10^{-5}	8.573×10^{-8}	9.293×10^{-10}
10	5	12	5.820×10^{-1}	2.600×10^{-3}	1.796×10^{-4}	——×—	8.878×10^{-7}	2.092×10^{-10}	3.988×10^{-14}
10	7	12	6.376×10^{-1}	2.212×10^{-3}	1.389×10^{-4}	——×—	4.090×10^{-7}	6.378×10^{-11}	1.125×10^{-14}
10	10	12	7.066×10^{-1}	1.861×10^{-3}	1.056×10^{-4}	——×—	1.657×10^{-7}	6.354×10^{-12}	9.907×10^{-16}

TABLE 3
Values of the measure Δ_2 for increasing sample sizes

p	q	n	Box	G	M2G	M3G	GNIG	M2GNIG	M3GNIG
3	2	5	1.364×10^{-1}	1.737×10^{-3}	2.601×10^{-4}	2.826×10^{-5}	3.461×10^{-4}	4.503×10^{-6}	4.503×10^{-6}
3	2	20	4.629×10^{-3}	7.343×10^{-5}	2.186×10^{-8}	——×—	2.010×10^{-5}	1.250×10^{-7}	1.040×10^{-9}
3	2	50	6.684×10^{-4}	1.083×10^{-5}	4.014×10^{-9}	——×—	3.033×10^{-6}	6.261×10^{-9}	5.697×10^{-12}
7	2	9	3.762×10^{-1}	3.632×10^{-3}	3.000×10^{-4}	3.579×10^{-5}	1.022×10^{-4}	1.037×10^{-6}	1.537×10^{-8}
7	2	20	3.897×10^{-2}	3.021×10^{-4}	3.912×10^{-6}	5.604×10^{-8}	2.059×10^{-5}	4.180×10^{-8}	7.620×10^{-11}
7	2	50	4.829×10^{-3}	3.694×10^{-5}	7.172×10^{-8}	8.263×10^{-11}	3.021×10^{-6}	9.071×10^{-10}	4.811×10^{-14}
10	2	12	4.986×10^{-1}	4.116×10^{-3}	3.794×10^{-4}	5.213×10^{-5}	2.141×10^{-5}	8.573×10^{-8}	5.282×10^{-10}
10	2	50	1.090×10^{-2}	5.712×10^{-5}	2.093×10^{-7}	——×—	1.844×10^{-6}	5.352×10^{-10}	1.601×10^{-13}
10	2	100	2.433×10^{-3}	1.269×10^{-5}	1.159×10^{-8}	——×—	4.595×10^{-7}	3.272×10^{-11}	2.338×10^{-15}

Table 4
Values of the measure Δ_2 for large q and large sample sizes

p	q	n	Box	G	M2G	M3G	GNIG	M2GNIG	M3GNIG
3	10	20	8.659×10^{-3}	3.260×10^{-5}	7.104×10^{-8}	1.313×10^{-9}	5.882×10^{-6}	9.890×10^{-9}	3.224×10^{-11}
3	10	50	1.283×10^{-3}	4.763×10^{-6}	1.852×10^{-9}	1.330×10^{-11}	8.294×10^{-7}	4.831×10^{-10}	6.378×10^{-13}
7	10	20	6.879×10^{-2}	1.268×10^{-4}	1.099×10^{-6}	—×—	1.226×10^{-6}	3.699×10^{-10}	1.942×10^{-13}
7	10	50	8.980×10^{-3}	1.505×10^{-5}	2.574×10^{-8}	—×—	2.172×10^{-7}	1.753×10^{-11}	5.605×10^{-15}
10	10	50	2.007×10^{-2}	2.350×10^{-5}	6.342×10^{-8}	—×—	3.343×10^{-8}	1.458×10^{-13}	—×—
10	10	100	4.565×10^{-3}	5.158×10^{-6}	4.263×10^{-9}	—×—	9.210×10^{-9}	4.031×10^{-14}	—×—

ter than the GNIG near-exact distribution for most of the cases, namely those with larger sample sizes.

In Tables 4 and 10, if we compare the values of Δ_2 for different values of p and the same sample size, we may see how, opposite to the asymptotic distributions, the near-exact distributions show a clear asymptotic character for increasing values of p.

In order to be able to compare the asymptotic saddle-point approximation from Jensen (1991, 1995) with the asymptotic and near-exact distributions proposed in this paper we have computed the tail probabilities for these approximations as well as for the Box asymptotic approximation for the exact 0.05 and 0.01 quantiles of λ^* for given values of p, q and n. This alternative approach to the use of the measures Δ_1 and Δ_2 is due to the fact that saddle-point approximations are not able to yield a c.f., indeed they do not even yield neither a p.d.f. nor a c.d.f. but only an approximation for the cumulative probability at a given point.

Since, mainly in order to be able to adequately evaluate the performance of the near-exact distributions we need to have the exact quantiles computed with enough precision (at least 12 decimal places), we have used the Gil-Pelaez (1951) inversion formulas, together with a bisection or secant method, to compute the exact quantiles. However, this computation procedure has its own limitations for values of p larger than 7.

From the tail probabilities in tables 5 and 6 we may see how the saddle-point approximation from Jensen (1991, 1995) although clearly beating his own Gamma asymptotic distribution and even displaying a better behavior than the one reported by the author in Jensen (1991, 1995), it has a worse behavior than the asymptotic distributions M2G and M3G proposed in this paper and it is no match for the near-exact distributions proposed in this paper.

TABLE 5
Absolute value of the difference between the exact tail probability (0.05) *and the tail probabilities for the approximating distributions*

p	q	n	Saddle	G	M2G	M3G	GNIG	M2GNIG	M3GNIG
4	3	6	2.53×10^{-4}	3.07×10^{-4}	1.22×10^{-5}	3.07×10^{-7}	1.08×10^{-6}	1.14×10^{-8}	4.86×10^{-11}
4	10	6	1.67×10^{-4}	3.47×10^{-4}	1.85×10^{-5}	—×—	2.90×10^{-6}	4.20×10^{-9}	7.08×10^{-12}
4	3	30	4.91×10^{-6}	6.07×10^{-6}	1.12×10^{-8}	7.65×10^{-12}	7.84×10^{-8}	1.36×10^{-10}	1.08×10^{-13}
4	10	30	3.07×10^{-6}	6.61×10^{-6}	1.63×10^{-8}	—×—	1.66×10^{-7}	3.75×10^{-11}	2.62×10^{-14}
5	3	7	3.30×10^{-4}	4.65×10^{-4}	2.84×10^{-5}	7.22×10^{-7}	1.04×10^{-5}	3.38×10^{-8}	5.82×10^{-11}
5	10	7	1.98×10^{-4}	4.42×10^{-4}	2.74×10^{-5}	—×—	5.93×10^{-6}	1.64×10^{-8}	5.05×10^{-11}
5	3	30	7.37×10^{-6}	1.10×10^{-5}	3.86×10^{-8}	8.24×10^{-11}	6.67×10^{-7}	2.07×10^{-11}	2.80×10^{-12}
5	10	30	4.15×10^{-6}	9.80×10^{-6}	3.48×10^{-8}	—×—	3.86×10^{-7}	1.78×10^{-10}	1.47×10^{-13}
6	3	8	3.86×10^{-4}	6.01×10^{-4}	4.56×10^{-5}	2.57×10^{-6}	3.47×10^{-6}	6.87×10^{-9}	1.11×10^{-11}
6	10	8	2.20×10^{-4}	5.16×10^{-4}	3.49×10^{-5}	—×—	6.24×10^{-7}	2.25×10^{-10}	5.45×10^{-14}
6	3	30	9.75×10^{-6}	1.65×10^{-5}	8.58×10^{-8}	3.15×10^{-10}	3.61×10^{-7}	8.62×10^{-11}	1.50×10^{-14}
6	10	30	5.25×10^{-6}	1.32×10^{-5}	6.09×10^{-8}	—×—	9.60×10^{-8}	6.01×10^{-12}	1.65×10^{-15}
7	3	9	4.28×10^{-4}	7.16×10^{-4}	6.21×10^{-5}	4.81×10^{-6}	2.13×10^{-6}	1.31×10^{-9}	3.71×10^{-12}
7	10	9	2.35×10^{-4}	5.73×10^{-4}	4.12×10^{-5}	—×—	1.26×10^{-6}	1.02×10^{-9}	7.26×10^{-13}
7	3	30	1.22×10^{-5}	2.26×10^{-5}	1.56×10^{-7}	—×—	3.79×10^{-7}	6.21×10^{-11}	7.18×10^{-13}
7	10	30	6.40×10^{-6}	1.70×10^{-5}	9.54×10^{-8}	—×—	2.10×10^{-7}	4.56×10^{-11}	1.62×10^{-14}

TABLE 6
Absolute value of the difference between the exact tail probability (0.01) *and the tail probabilities for the approximating distributions*

p	q	n	Saddle	G	M2G	M3G	GNIG	M2GNIG	M3GNIG
4	3	6	4.19×10^{-5}	3.83×10^{-4}	1.75×10^{-5}	1.12×10^{-6}	7.20×10^{-7}	2.19×10^{-9}	2.17×10^{-12}
4	10	6	3.15×10^{-5}	2.94×10^{-4}	7.70×10^{-6}	—×—	2.00×10^{-6}	6.13×10^{-10}	1.44×10^{-12}
4	3	30	7.64×10^{-7}	7.64×10^{-6}	1.14×10^{-8}	2.20×10^{-11}	8.92×10^{-8}	9.11×10^{-11}	7.56×10^{-14}
4	10	30	5.81×10^{-7}	5.55×10^{-6}	5.42×10^{-9}	—×—	1.36×10^{-7}	1.05×10^{-11}	1.02×10^{-15}
5	3	7	5.72×10^{-5}	5.07×10^{-4}	2.73×10^{-5}	2.24×10^{-6}	7.32×10^{-6}	8.85×10^{-9}	3.93×10^{-12}
5	10	7	3.75×10^{-5}	3.59×10^{-4}	1.00×10^{-5}	—×—	4.03×10^{-6}	2.09×10^{-9}	1.27×10^{-11}
5	3	30	1.30×10^{-6}	1.17×10^{-5}	2.59×10^{-8}	6.47×10^{-11}	6.69×10^{-7}	2.33×10^{-11}	2.22×10^{-12}
5	10	30	7.84×10^{-7}	7.86×10^{-6}	9.66×10^{-9}	—×—	3.02×10^{-7}	4.17×10^{-11}	1.48×10^{-14}
6	3	8	6.89×10^{-5}	6.04×10^{-4}	3.51×10^{-5}	3.20×10^{-6}	2.24×10^{-6}	9.05×10^{-10}	1.61×10^{-12}
6	10	8	4.17×10^{-5}	4.08×10^{-4}	1.16×10^{-5}	—×—	4.04×10^{-7}	1.50×10^{-11}	2.21×10^{-14}
6	3	30	1.79×10^{-6}	1.60×10^{-5}	4.53×10^{-8}	1.36×10^{-10}	3.27×10^{-7}	3.70×10^{-11}	3.62×10^{-15}
6	10	30	9.88×10^{-7}	1.03×10^{-5}	1.47×10^{-8}	—×—	7.23×10^{-8}	1.16×10^{-12}	1.91×10^{-16}
7	3	9	7.78×10^{-5}	6.81×10^{-4}	4.10×10^{-5}	3.89×10^{-6}	1.32×10^{-6}	1.30×10^{-10}	1.05×10^{-12}
7	10	9	4.48×10^{-5}	4.45×10^{-4}	1.27×10^{-5}	—×—	8.12×10^{-7}	6.03×10^{-11}	3.20×10^{-13}
7	3	30	2.27×10^{-6}	2.06×10^{-5}	6.96×10^{-8}	—×—	3.21×10^{-7}	1.71×10^{-11}	4.69×10^{-15}
7	10	30	1.20×10^{-6}	1.29×10^{-5}	2.08×10^{-8}	—×—	1.55×10^{-7}	7.65×10^{-12}	3.61×10^{-15}

6. Computational implementation of the near-exact distributions developed. The computational implementation of the near-exact distributions developed in this paper is made quite simple if we use an adequate high-level language software like Mathematica®.

```
NEdistW[p_,q_,n_,t_,opts___]:=Module[{d,pr1,pr2,pr3,adn,func,vp,rj,a,xr},
 Options[NEdistW]={dist->1,prcm->300,prcpar->200,prcf->16,ad->0,function->1};
 d = dist/.{opts}/.Options[NEdistW];
 pr1 = prcm/.{opts}/.Options[NEdistW];
 pr2 = prcf/.{opts}/.Options[NEdistW];
 pr3 = prcpar/.{opts}/.Options[NEdistW];
 adn = ad/.{opts}/.Options[NEdistW];
 func = function/.{opts}/.Options[NEdistW];
 vp=If[d==1,Module1[p,q,n,pr1],If[d==2,Module2[p,q,n,pr1],Module3[p,q,n,pr1,
                                                              pr3,adn]]];
 rj = Rj[p,q,n];
 a = Table[(n-k)/n,{k,1,p-1}];
 xr = Rationalize[t,0];
 If[func==3,prd=Product[((n-k)/n)^rj[[k]]*((n-k)/n-I*xr)^(-rj[[k]]),{k,1,p-1}];
  If[d==1,vp[[2]]^vp[[1]]*(vp[[2]]-I*xr)^(-vp[[1]])*prd,
    If[d==2,(vp[[1]]*vp[[4]]^vp[[2]]*(vp[[4]]-I*xr)^(-vp[[2]])+(1-vp[[1]])*
                            vp[[4]]^vp[[3]]*(vp[[4]]-I*xr)^(-vp[[3]]))*prd,
    (vp[[1]]*vp[[6]]^vp[[3]]*(vp[[6]]-I*xr)^(-vp[[3]])+vp[[2]]*vp[[6]]^vp[[4]]*
    (vp[[6]]-I*xr)^(-vp[[4]])+(1-vp[[1]]-vp[[2]])*vp[[6]]^vp[[5]]*
    (vp[[6]]-I*xr)^(-vp[[5]]))*prd]],
  If[func==2,GNIG[r_,b_,l_,a_,w_]:=GNIGpdf[r,b,l,a,w],GNIG[r_,b_,l_,a_,w_]:=
                                                         GNIGcdf[r,b,l,a,w]];
  If[d==1,SetPrecision[GNIG[rj,vp[[1]],a,vp[[2]],xr],pr2],
     If[d==2,SetPrecision[vp[[1]]*GNIG[rj,vp[[2]],a,vp[[4]],xr]+(1-vp[[1]])*
                                       GNIG[rj,vp[[3]],a,vp[[4]],xr],pr2],
    SetPrecision[vp[[1]]*GNIG[rj,vp[[3]],a,vp[[6]],xr]+vp[[2]]*GNIG[rj,vp[[4]],
        a,vp[[6]],xr]+(1-vp[[1]]-vp[[2]])*GNIG[rj,vp[[5]],a,vp[[6]],xr],pr2]]
      ]]]
```

FIG 2. MATHEMATICA *module to implement and compute the c.d.f., p.d.f. or c.f. of the near-exact distributions developed in the paper*

In Figures 2 through 4 we have a set of Mathematica modules that may be used to compute either the c.f., the c.d.f. or the p.d.f. of $W = -\log(\lambda^*)$. The main module is module `NEdistW` in Figure 2, which acts as the interface with the user, who will be able to choose which of the characteristic, cumulative distribution or probability density functions he wants to compute, through the use of an optional argument.

This module has four mandatory arguments which are:

i) p – the number of variables involved, that is, the dimension of the

```
Module1[p_,q_,n_,prcm_:300]:=Module[{mm},
  mm=Table[SetPrecision[MomPhi2[p,q,n,i],prcm],{i,1,2}];
   {mm[[1]]^2/(mm[[2]]-mm[[1]]^2),mm[[1]]/(mm[[2]]-mm[[1]]^2)}
 ]

Module2[p_,q_,n_,prcm_:300]:=Module[{mm},
  mm=Table[SetPrecision[MomPhi2[p,q,n,i],prcm],{i,1,4}];
  Sort[Cases[{p1,r1,r2,m}/.NSolve[
       {mm[[1]]==MomMixGam[{p1},{r1,r2},{m,m},1],
        mm[[2]]==MomMixGam[{p1},{r1,r2},{m,m},2],
        mm[[3]]==MomMixGam[{p1},{r1,r2},{m,m},3],
        mm[[4]]==MomMixGam[{p1},{r1,r2},{m,m},4]},
        {p1,r1,r2,m}],{_Real,_Real,_Real,_Real}]][[2]]
    ]

Module3[p_,q_,n_,prcm_:300,prcpar_:200,ad_:0]:=Module[{mm,vpn},
 mm = Table[SetPrecision[MomPhi2[p,q,n,i],prcm],{i,1,6}];
 vpn = Module2[p,q,n,prcm];
 {p1,p2,r1,r2,r3,m}/.
   FindRoot[{mm[[1]]==MomMixGam[{p1,p2},{r1,r2,r3},{m,m,m},1],
             mm[[2]]==MomMixGam[{p1,p2},{r1,r2,r3}, {m,m,m},2],
             mm[[3]]==MomMixGam[{p1,p2},{r1,r2,r3},{m,m,m},3],
             mm[[4]]==MomMixGam[{p1,p2},{r1,r2,r3},{m,m,m},4],
             mm[[5]]==MomMixGam[{p1,p2},{r1,r2,r3},{m,m,m},5],
             mm[[6]]==MomMixGam[{p1,p2},{r1,r2,r3},{m,m,m},6]},{p1,vpn[[1]]},
             {p2,.99*(1-vpn[[1]])},{r1,vpn[[2]]},{r2,vpn[[3]]},{r3,ad+
      2*vpn[[3]]-vpn[[2]]},{m,vpn[[4]]},
     WorkingPrecision -> prcpar]
  ]
```

Fig 3. Mathematica *modules to compute the parameters for the non-integer Gammas in the near-exact distributions*

matrices being tested,

ii) q – the number of matrices being tested,

iii) n – the sample size minus one,

iv) the point-value at which we want to evaluate the function being computed (c.f., c.d.f. or p.d.f.).

These four mandatory arguments have to be given in the order they are listed above.

This module also has six optional arguments. These are:

i) `function` – the argument that defines which function we want to compute (1 – for the c.d.f., 2 – for the p.d.f.; 3 – for the c.f.); this argument has the default value of 1 and if given a value outside of the above mentioned values the module computes the c.d.f.;

ii) `dist` – the argument that defines which near-exact distribution is to be used (1 – for the single GNIG; 2 – for the M2GNIG; 3 – for the M3GNIG); this argument has the default value of 1 and if given a value out of the above mentioned range will make the module to compute the M3GNIG distribution;

iii) `prcm, prcf` and `prepar` – the arguments with respective default values of 300, 16 and 200 which define the number of precision digits for the computation of respectively: (a) – the exact moments of W to be matched by the near-exact distributions, (b) – the near-exact functions (that is either the c.d.f., the p.d.f. or the c.f., according to the value given for the argument `function`), and (c) – the parameters for the M3GNIG near-exact distribution;

iv) `ad` – the argument with a default value of zero, which may be given any real positive or negative (usually small) value to try to stabilize the convergence of the solution of the system of equations for the M3GNIG distribution (in `Module3`) for cases where this convergence is not attained through the use of the automatically generated initial values.

The modules `Module1`, `Module2` and `Module3` in Figure 3 are called by the module `NEdistW` to compute the parameters respectively for the GNIG, the M2GNIG and the M3GNIG near-exact distributions. While modules `Module1` and `Module2` have as mandatory arguments the values for p, q and n, in this order, and as optional argument, with a default value of 300, the parameter `prcm` described above, `Module3` has as further optional arguments the parameters `prcpar` and `ad`, also described above, given in this order and with default values 200 and zero, respectively.

In Figure 4 we have modules `Phi2`, `MomPhi2` and `MomMixGam`. The former of these modules is used to compute the part of the c.f. of W denoted by $\Phi_2(t)$ in expression (3.6). This module has four mandatory arguments which are the values for p, q, n and the value for the running variable t, which may be not numerically specified. This module is called by module `MomPhi2`, which computes the derivatives of $Phi_2(t)$ at $t = 0$, or rather, the moments, corresponding to the c.f. $\Phi_2(t)$. This module also has four mandatory arguments, p, q, n and a fourth argument which has to bear a positive integer value, specifying the order of the derivative of $\Phi_2(t)$ to be computed.

The module `MomMixGam` is used to compute (usually only symbolically) the moments of a finite mixture of Gamma distributions. It has four mandatory arguments which, for a mixture of ν Gamma distributions are:

```
Phi2[p_,q_,n_,t_]:= Module[{aj,bjk,bjks},
 aj = Table[n+1-2*j,{j,1,Floor[p/2]}];
 bjk = Table[Table[2*j-1+(k-2*j)/q,{k,1,q}],{j,1,Floor[p/2]}];
 bjks = Floor[bjk];
 ap = (n + 1 - p)/2;
 bpk = Table[(p-1)/2-(p+1)/(2*q)+k/q,{k,1,q}];
 bpks = Floor[bpk];
 Product[Product[Gamma[aj[[j]]+bjk[[j,k]]]*Gamma[aj[[j]]+bjks[[j,k]]-n*I*t]/
  (Gamma[aj[[j]]+bjks[[j,k]]]*Gamma[aj[[j]]+bjk[[j,k]]-n*I*t]),{k,1,q}],
   {j,1,Floor[p/2]}]*(Product[Gamma[ap+bpk[[k]]]*Gamma[ap+bpks[[k]]-n/2*I*t]/
          (Gamma[ap+bpks[[k]]]*Gamma[ap+bpk[[k]]-n/2*I*t]),{k,1,q}])^Mod[p,2]
    ]

MomPhi2[p_,q_,n_,h_]:=1/I^h*D[Phi2[p,q,n,t],{t,h}]/.t->0

MomMixGam[p_,r_,m_,h_]:=Module[{nt,ptot},
  nt = Length[r];
  ptot = Apply[Plus,p];
  (1-ptot)*Product[r[[nt]]+i,{i,0,h-1}]*m[[nt]]^(-h)+Sum[p[[j]]*
                          Product[r[[j]]+i,{i,0,h-1}]*m[[j]]^(-h),{j,1,nt-1}]
    ]
```

FIG 4. MATHEMATICA *modules to compute $\Phi_2(t)$, its moments and the moments of Gamma mixtures*

- p – a vector of weights of length $\nu-1$, specifying the first $\nu-1$ weights for the distributions in the mixture, being the ν-th weight computed as $1-\sum_{k=1}^{\nu-1} p_k$;
- r – a vector of length ν with the shape parameters of the Gamma distributions in the mixture;
- m – a vector of length ν with the rate parameters of the Gamma distributions in the mixture;
- h – a positive integer specifying the order of the moment to be computed.

Module `MomPhi2` is called by modules `Module1`, `Module2` and `Module3`, while module `MomMixGam` is called only by modules `Module2` and `Module3` in order to compute the values for the parameters for the single Gamma distribution or for the mixture of two or three Gamma distributions used to build the near-exact distributions with the first 2, 4 or 6 moments matching the first 2, 4 or 6 exact moments of W.

One of the most useful applications of the modules above is to compute near-exact p-values. Let us suppose that we had for $p=5$, $q=4$ and $n=15$ a computed value of 37.2026 for W. An example of a call to module `NEdistW` to compute the near-exact p-value for the M3GNIG distribution, would be

```
1-NEdistW[5,4,15,37.2026,dist->3,function->1]
```

or, equivalently,

```
1-NEdistW[5,4,15,37.2026,dist->3] .
```

The result obtained would be 0.0405706732106333, in about 1.30 seconds (average time for 5 executions in a double core processor of 1.66GHz). If we would like to obtain the near-exact p-value for the simple GNIG near-exact distribution, then the command

```
1-NEdistW[5,4,15,37.2026]
```

would be good enough, given the default values assumed for the optional arguments `function` and `dist`. The result obtained in this case would be, 0.0405678479033504, in about 0.43 seconds (average time for 5 executions with the same processor as above).

7. Conclusions. The results presented in this paper, together with the ones already published on the Wilks Λ statistic (Coelho, 2004; Alberto and Coelho, 2007; Grilo and Coelho, 2007) and on the sphericity l.r.t. statistic (Marques and Coelho, 2008), show that the distribution of the negative logarithm of the three main l.r.t. statistics used in Multivariate Analysis may be written as the sum of a GIG distribution with an independent sum of independent Logbeta r.v.'s. These results are intended to be used as the basis for two future works: one on a common approach for the more common l.r.t. statistics used in Multivariate Analysis which will recall the common traits of these statistics both in terms of their exact and near-exact distributions, and the other on a general approach for two families of generalized sphericity tests, which we may call as multi-sample block-scalar and multi-sample block-matrix sphericity tests, their common links and particular cases.

All the near-exact distributions developed in this paper show a very good performance, with the ones based on mixtures showing an outstanding behaviour. For the approximate distributions developed in this paper, the near-exact distributions clearly outperform their asymptotic counterparts, for a given number of exact moments matched.

Moreover, opposite to the usual asymptotic distributions, the near-exact distributions developed show a marked asymptotic behavior not only for increasing sample sizes but also for increasing values of p (the number of variables) and for increasing values of q (the number of matrices being tested). Yet, all the near-exact distributions proposed may be easily used to compute near-exact quantiles.

Indeed, two of the great features of the near-exact distributions are the facts that, opposite to the usual asymptotic approximations and mainly to the single chi-square approximations which show quite bad fits for small samples (see for example, Lo 2008; Zhang and Boos 1992), the near-exact distributions show a very good fit even for small samples, and their closeness to the exact distribution even increases when the dimension increases.

We should stress here that also the two new asymptotic distributions proposed show an asymptotic behavior for increasing values of q (the number of matrices being tested).

Thus, as a final comment, and given the values of the measures Δ_1 and Δ_2 obtained for the distributions, we would say that we may use the asymptotic distributions proposed in this paper in practical applications that may need a not so high degree of precision, although higher than the one that the usual asymptotic distributions deliver. For applications that may need a high degree of precision in the computation of quantiles we may then use the more elaborate near-exact distributions, mainly those based on mixtures of GNIG distributions, which anyway allow for an easy computation of quantiles. The distribution M3GNIG, given its excellent performance and manageability, may even be used as a replacement of the exact distribution.

As a conclusion we may say that given its quite good precision, actually higher than the asymptotic and saddle-point approximation in Jensen (1991, 1995), as well as the easiness in computing its parameters, the M2G asymptotic distribution seems to be the most adequate for cases where a moderate precision is needed, while the near-exact distributions are appropriate for situations where further precision is needed. The implementation of these distributions is rendered rather (almost, nearly, somewhat) easy when using one of the extended precision and symbolic computation softwares nowadays commonly available, like Mathematica®.

APPENDIX A: PROOFS OF THEOREM 3.1 AND THEOREM 3.2 IN SECTION 3

A.1. Proving Theorem 3.1. In order to prove Theorem 3.1 in section 3 we will first show how $\Phi_W(t)$ may be factorized in two factors, one of them being $\Phi_2(t)$ in Theorem 3.1 and a second factor which will be later on identified with the c.f. of a sum of Exponential distributions.

Lemma 1. *The characteristic function of* $W = -\log \lambda^*$ *may be written as*

$$\Phi_W(t) = \underbrace{\prod_{j=1}^{\lfloor p/2\rfloor}\prod_{k=1}^{q}\frac{\Gamma(a_j+b_{jk}^*)}{\Gamma(a_j)}\frac{\Gamma(a_j-nit)}{\Gamma(a_j+b_{jk}^*-nit)} \times\left\{\prod_{k=1}^{q}\frac{\Gamma(a_p+b_{pk}^*)}{\Gamma(a_p)}\frac{\Gamma\left(a_p-\frac{n}{2}\mathrm{i}t\right)}{\Gamma\left(a_p+b_{pk}^*-\frac{n}{2}\mathrm{i}t\right)}\right\}^{p\perp 2}}_{\Phi_1(t)}$$
$$\times\underbrace{\prod_{j=1}^{\lfloor p/2\rfloor}\prod_{k=1}^{q}\frac{\Gamma(a_j+b_{jk})}{\Gamma(a_j+b_{jk}^*)}\frac{\Gamma(a_j+b_{jk}^*-nit)}{\Gamma(a_j+b_{jk}-nit)} \times\left\{\prod_{k=1}^{q}\frac{\Gamma(a_p+b_{pk})}{\Gamma(a_p+b_{pk}^*)}\frac{\Gamma\left(a_p+b_{pk}^*-\frac{n}{2}\mathrm{i}t\right)}{\Gamma\left(a_p+b_{pk}-\frac{n}{2}\mathrm{i}t\right)}\right\}^{p\perp 2}}_{\Phi_2(t)}, \tag{A.1}$$

where $p \perp\!\!\!\perp 2 = p-2\left\lfloor\frac{p}{2}\right\rfloor = Mod(p,2)$ *represents the remainder of the integer division of* p *by 2,*

$$a_j = n+1-2j\,, \qquad b_{jk} = 2j-1+\frac{k-2j}{q}\,, \qquad b_{jk}^* = \lfloor b_{jk}\rfloor\,, \tag{A.2}$$

$$a_p = \frac{n+1-p}{2}\,, \qquad b_{pk} = \frac{pq-q-p+2k-1}{2q}\,, \qquad b_{pk}^* = \lfloor b_{pk}\rfloor\,, \tag{A.3}$$

Proof. Using the fact that

$$\Gamma(2z) = \pi^{-1/2}\,2^{2z-1}\,\Gamma(z)\,\Gamma(z+1/2)$$

we may write, from (1.7), the c.f. of $W = -\log \lambda^*$ as

$$\Phi_W(t) = q^{-npqit/2} \prod_{j=1}^{\lfloor p/2 \rfloor} \left\{ \frac{\Gamma(nq+1-2j)}{\Gamma(nq+1-2j-nqit)} \prod_{k=1}^{q} \frac{\Gamma(n+1-2j-nit)}{\Gamma(n+1-2j)} \right\}$$

$$\times \left\{ \frac{\Gamma\left(\frac{nq}{2}+\frac{1}{2}-\frac{p}{2}\right)}{\Gamma\left(\frac{nq}{2}+\frac{1}{2}-\frac{p}{2}-\frac{nq}{2}\mathrm{i}t\right)} \prod_{k=1}^{q} \frac{\Gamma\left(\frac{n}{2}+\frac{1}{2}-\frac{p}{2}-\frac{n}{2}\mathrm{i}t\right)}{\Gamma\left(\frac{n}{2}+\frac{1}{2}-\frac{p}{2}\right)} \right\}^{p \perp 2} .$$

Then, using

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

and taking into account that since for any $p \in I\!N$, $\lfloor p/2 \rfloor + \lfloor (p+1)/2 \rfloor = p$,

$$q^{-npqit/2} \left(q^{\frac{nq}{2}\mathrm{i}t}\right)^{\lfloor \frac{p+1}{2} \rfloor - \lfloor \frac{p}{2} \rfloor} \prod_{j=1}^{\lfloor p/2 \rfloor} q^{nqit}$$

$$= q^{-npqit/2+nq\lfloor p/2 \rfloor \mathrm{i}t/2+nq\lfloor \frac{p+1}{2} \rfloor \mathrm{i}t/2} = 1 ,$$

we may write

$$\Phi_W(t) = \prod_{j=1}^{\lfloor p/2 \rfloor} \prod_{k=1}^{q} \left\{ \frac{\Gamma\left(n+\frac{1}{q}-\frac{2j}{q}+\frac{k-1}{q}\right)}{\Gamma\left(n+\frac{1}{q}-\frac{2j}{q}+\frac{k-1}{q}-nit\right)} \frac{\Gamma(n+1-2j-nit)}{\Gamma(n+1-2j)} \right\}$$

$$\times \left\{ \prod_{k=1}^{q} \frac{\Gamma\left(\frac{n}{2}+\frac{1}{2q}-\frac{p}{2q}+\frac{k-1}{q}\right)}{\Gamma\left(\frac{n}{2}+\frac{1}{2q}-\frac{p}{2q}+\frac{k-1}{q}-\frac{n}{2}\mathrm{i}t\right)} \frac{\Gamma\left(\frac{n}{2}+\frac{1}{2}-\frac{p}{2}-\frac{n}{2}\mathrm{i}t\right)}{\Gamma\left(\frac{n}{2}+\frac{1}{2}-\frac{p}{2}\right)} \right\}^{p \perp 2}$$

Taking then a_j, b_{jk}, a_p and b_{pk} defined as in (A.2) and (A.3), we may write

$$\Phi_W(t) = \prod_{j=1}^{\lfloor p/2 \rfloor} \prod_{k=1}^{q} \frac{\Gamma(a_j+b_{jk})}{\Gamma(a_j+b_{jk}-nit)} \frac{\Gamma(a_j-nit)}{\Gamma(a_j)}$$

$$\times \left\{ \prod_{k=1}^{q} \frac{\Gamma(a_p+b_{pk})}{\Gamma\left(a_p+b_{pk}-\frac{n}{2}\mathrm{i}t\right)} \frac{\Gamma\left(a_p-\frac{n}{2}\mathrm{i}t\right)}{\Gamma(a_p)} \right\}^{p \perp 2}$$

that taking b^*_{jk} and b^*_{pk} given by (A.2) and (A.3) may be written as

$$\Phi_W(t) = \prod_{j=1}^{\lfloor p/2\rfloor}\prod_{k=1}^{q}\left\{\frac{\Gamma(a_j+b^*_{jk})}{\Gamma(a_j)}\,\frac{\Gamma(a_j-nit)}{\Gamma(a_j+b^*_{jk}-nit)}\right.$$
$$\left.\times\frac{\Gamma(a_j+b_{jk})}{\Gamma(a_j+b^*_{jk})}\,\frac{\Gamma(a_j+b^*_{jk}-nit)}{\Gamma(a_j+b_{jk}-nit)}\right\}$$
$$\times\left(\prod_{k=1}^{q}\frac{\Gamma(a_p+b_{pk})}{\Gamma(a_p+b^*_{pk})}\,\frac{\Gamma(a_p+b^*_{pk}-\frac{n}{2}\mathrm{i}t)}{\Gamma\left(a_p+b_{pk}-\frac{n}{2}\mathrm{i}t\right)}\right)^{p\perp 2}$$
$$\times\left(\prod_{k=1}^{q}\frac{\Gamma(a_p+b^*_{pk})}{\Gamma\left(a_p\right)}\,\frac{\Gamma\left(a_p-\frac{n}{2}\mathrm{i}t\right)}{\Gamma(a_p+b^*_{pk}-\frac{n}{2}\mathrm{i}t)}\right)^{p\perp 2},$$

what, after some small rearrangements yields the result in the Lemma. □

Let us now take

$$\text{(A.4)}\qquad \Phi_1(t) = \underbrace{\prod_{j=1}^{\lfloor p/2\rfloor}\prod_{k=1}^{q}\frac{\Gamma(a_j+b^*_{jk})}{\Gamma(a_j)}\,\frac{\Gamma(a_j-nit)}{\Gamma(a_j+b^*_{jk}-nit)}}_{\Phi_{1,1}(t)}\times\underbrace{\left\{\prod_{k=1}^{q}\frac{\Gamma(a_p+b^*_{pk})}{\Gamma\left(a_p\right)}\,\frac{\Gamma\left(a_p-\frac{n}{2}\mathrm{i}t\right)}{\Gamma(a_p+b^*_{pk}-\frac{n}{2}\mathrm{i}t)}\right\}^{p\perp 2}}_{\Phi_{1,2}(t)}.$$

We will now show that $\Phi_1(t)$ is indeed the c.f. of the sum of independent Exponential r.v.'s and we will identify the different Exponential distributions involved, by adequately decomposing first $\Phi_{1,1}(t)$, in Lemma 2, and then $\Phi_{1,2}(t)$, in Lemma 3.

Lemma 2. *We may write*

$$\Phi_{1,1}(t) = \prod_{j=1}^{\lfloor p/2\rfloor}\prod_{k=1}^{q}\prod_{l=1-\lfloor\frac{k-2j}{q}\rfloor}^{2j-1}\frac{n-l}{n-l-nit}.$$

Proof. Applying now

$$\text{(A.5)}\qquad \frac{\Gamma(a+n)}{\Gamma(a)} = \prod_{l=0}^{n-1}(a+l)$$

and noticing that

$$b^*_{jk} = \left\lfloor 2j - 1 + \frac{k-2j}{q} \right\rfloor = 2j - 1 + \left\lfloor \frac{k-2j}{q} \right\rfloor ,$$

we may write

$$\begin{aligned}
\Phi_{1,1}(t) &= \prod_{j=1}^{\lfloor p/2 \rfloor} \prod_{k=1}^{q} \frac{\Gamma(a_j+b^*_{jk})}{\Gamma(a_j)} \frac{\Gamma(a_j-n\mathrm{i}t)}{\Gamma(a_j+b^*_{jk}-n\mathrm{i}t)} \\
&= \prod_{j=1}^{\lfloor p/2 \rfloor} \prod_{k=1}^{q} \prod_{l=0}^{b^*_{kj}-1} \frac{n-2j+1+l}{n-2j+1+l-n\mathrm{i}t} \\
&= \prod_{j=1}^{\lfloor p/2 \rfloor} \prod_{k=1}^{q} \prod_{l=1}^{b^*_{kj}} \frac{n-2j+l}{n-2j+l-n\mathrm{i}t}
\end{aligned}$$

$$\begin{aligned}
&= \prod_{j=1}^{\lfloor p/2 \rfloor} \prod_{k=1}^{q} \prod_{l=1}^{2j-1+\left\lfloor \frac{k-2j}{q} \right\rfloor} \frac{n-2j+l}{n-2j+l-n\mathrm{i}t} \\
&= \prod_{j=1}^{\lfloor p/2 \rfloor} \prod_{k=1}^{q} \prod_{l=1}^{2j-1+\left\lfloor \frac{k-2j}{q} \right\rfloor} \frac{n-l+\left\lfloor \frac{k-2j}{q} \right\rfloor}{n-l+\left\lfloor \frac{k-2j}{q} \right\rfloor-n\mathrm{i}t} ,
\end{aligned}$$

that by a simple change in the limits of the last product yields the desired result. □

LEMMA 3. *For* $\Phi_{1,2}(t) = \left(\Phi^*_{1,2}(t)\right)^{p\perp 2}$, *with* a_p *and* b^*_{pk} *defined in* (A.3), *we may write,*

$$\begin{aligned}
\Phi^*_{1,2}(t) &= \prod_{k=1}^{q} \frac{\Gamma(a_p+b^*_{pk})}{\Gamma(a_p)} \frac{\Gamma\left(a_p-\frac{n}{2}\mathrm{i}t\right)}{\Gamma(a_p+b^*_{pk}-\frac{n}{2}\mathrm{i}t)} \\
&= \prod_{k=1}^{q} \prod_{l=1}^{b^*_{pk}} \frac{\frac{2(a_p+l-1)}{n}}{\frac{2(a_p+l-1)}{n}-\mathrm{i}t} .
\end{aligned} \tag{A.6}$$

PROOF. Using (A.5) we may always write, for $b^*_{pk} = \lfloor b_{pk} \rfloor$,

$$\begin{aligned}
\Phi^*_{1,2}(t) &= \prod_{k=1}^{q} \frac{\Gamma(a_p+b^*_{pk})\,\Gamma\left(a_p-\frac{n}{2}\mathrm{i}t\right)}{\Gamma(a_p)\,\Gamma\left(a_p+b^*_{pk}-\frac{n}{2}\mathrm{i}t\right)} \\
&= \prod_{k=1}^{q} \prod_{l=1}^{b^*_{pk}} \frac{a_p+l-1}{a_p+l-1-\frac{n}{2}\mathrm{i}t} .
\end{aligned}$$

□

The results in Lemmas 1 through 3 prove Theorem 3.1.

A.2. Proving Theorem 3.2. To prove Theorem 3.2 we need the following three Lemmas. In the first two Lemmas we identify the different Exponential distributions involved in $\Phi_{1,1}(t)$ and obtain analytic expressions for their counts. The first Lemma refers to even q and the second to odd q.

LEMMA 4. *For even q we may write*

$$\Phi_{1,1}(t) = \prod_{j=\alpha+2}^{2\lfloor p/2\rfloor-1} \left(\frac{\frac{n-j}{n}}{\frac{n-j}{n}-\mathrm{i}t}\right)^{q(\lfloor p/2\rfloor-\lfloor j/2\rfloor)} \prod_{k=1}^{\alpha+1} \left(\frac{\frac{n-k}{n}}{\frac{n-k}{n}-\mathrm{i}t}\right)^{a_k+\gamma_k} \tag{A.7}$$

where $\alpha = \left\lfloor \frac{p-1}{q} \right\rfloor$,

$$a_k = \begin{cases} q^2/4\,, & k=1,\dots,\alpha \\ \left(q-\left(\lfloor \frac{p}{2}\rfloor - \alpha\lfloor\frac{q}{2}\rfloor\right)\right)\left(\lfloor\frac{p}{2}\rfloor-\alpha\lfloor\frac{q}{2}\rfloor\right), & k=\alpha+1\,, \end{cases} \tag{A.8}$$

and

$$\gamma_k = \lfloor \tfrac{q}{2}\rfloor \left((k-1)q - 2\left\lfloor\tfrac{k}{2}\right\rfloor\right) \qquad (k=1,\dots,\alpha+1)\,. \tag{A.9}$$

PROOF. Since for $k=1,\dots,q$ and $j=1,\dots,\lfloor p/2\rfloor$, $\nu = -\left\lfloor\frac{k-2j}{q}\right\rfloor$ takes the values $0,1,2,\dots,\alpha$, with $\alpha = \left\lfloor\frac{p-1}{q}\right\rfloor$, we may write, from the result in Lemma 2,

$$\begin{aligned}\Phi_{1,1}(t) &= \prod_{j=1}^{\lfloor p/2\rfloor}\prod_{k=1}^{q}\prod_{l=1-\left\lfloor\frac{k-2j}{q}\right\rfloor}^{2j-1} \frac{n-l}{n-l-nit}\\ &= \prod_{j=1}^{\lfloor p/2\rfloor}\prod_{k=1}^{q}\left(\prod_{l=1-\left\lfloor\frac{k-2j}{q}\right\rfloor}^{\min(\alpha+1,2j-1)} \frac{n-l}{n-l-nit}\prod_{l=\alpha+2}^{2j-1}\frac{n-l}{n-l-nit}\right)\\ &= \prod_{j=1}^{\lfloor p/2\rfloor}\prod_{l=\alpha+2}^{2j-1}\left(\frac{n-l}{n-l-nit}\right)^{q} \prod_{j=1}^{\lfloor p/2\rfloor}\prod_{k=1}^{q}\prod_{l=1-\left\lfloor\frac{k-2j}{q}\right\rfloor}^{\lfloor(\alpha+1,2j-1)} \frac{n-l}{n-l-nit}\end{aligned} \tag{A.10}$$

where

$$\prod_{j=1}^{\lfloor p/2\rfloor}\prod_{l=\alpha+2}^{2j-1}\left(\frac{n-l}{n-l-nit}\right)^{q}=\prod_{j=\alpha+2}^{2\lfloor p/2\rfloor-1}\left(\frac{n-j}{n-j-nit}\right)^{q(\lfloor p/2\rfloor-\lfloor j/2\rfloor)} \tag{A.11}$$

and, for even q (for which although $\lfloor q/2\rfloor = q/2$, we will use the notation $\lfloor q/2\rfloor$ to make it more uniform with the notation for odd q), taking

$$\begin{aligned}\mathsf{Prod}(\nu;\alpha,q,j) = &\prod_{l=1+\nu}^{\lfloor(\alpha+1,2j-1)}\left(\frac{n-l}{n-l-nit}\right)^{q-\left(2\left(j-\nu\left\lfloor\frac{q}{2}\right\rfloor\right)-1\right)}\\ &\times\prod_{l=2+\nu}^{\lfloor(\alpha+1,2j-1)}\left(\frac{n-l}{n-l-nit}\right)^{2\left(j-\nu\left\lfloor\frac{q}{2}\right\rfloor\right)-1},\end{aligned} \tag{A.12}$$

we have, for $\varphi(\nu,j,q) = 2(j+\nu\lfloor q/2\rfloor)-1$,

$$\begin{aligned}&\prod_{j=1}^{\lfloor p/2\rfloor}\prod_{k=1}^{q}\prod_{l=1-\left\lfloor\frac{k-2j}{q}\right\rfloor}^{\lfloor(\alpha+1,2j-1)}\frac{n-l}{n-l-nit}\\ &\quad=\prod_{\nu=0}^{\alpha-1}\prod_{j=1+\nu\lfloor q/2\rfloor}^{(\nu+1)\lfloor q/2\rfloor}\mathsf{Prod}(\nu;\alpha,q,j)\prod_{j=1+\alpha\lfloor q/2\rfloor}^{\lfloor p/2\rfloor}\mathsf{Prod}(\alpha;\alpha,q,j)\\ &\quad=\prod_{\nu=0}^{\alpha-1}\prod_{j=1}^{\lfloor q/2\rfloor}\left\{\prod_{l=1+\nu}^{\lfloor(\alpha+1,\varphi(\nu,j,q))}\left(\frac{n-l}{n-l-nit}\right)^{q-(2j-1)}\right.\\ &\qquad\qquad\left.\times\prod_{l=2+\nu}^{\lfloor(\alpha+1,\varphi(\nu,j,q))}\left(\frac{n-l}{n-l-nit}\right)^{2j-1}\right\}\\ &\qquad\qquad\times\prod_{j=1}^{\lfloor p/2\rfloor-\alpha\lfloor q/2\rfloor}\left(\frac{n-(\alpha+1)}{n-(\alpha+1)-nit}\right)^{q-(2j-1)}\end{aligned}$$

$$= \prod_{j=1}^{\lfloor q/2\rfloor}\prod_{\nu=0}^{\alpha-1}\prod_{l=2+\nu}^{\lfloor(\alpha+1,2j+\nu q-1)}\left(\frac{n-l}{n-l-nit}\right)^{q}$$

$$\times\prod_{j=1}^{\lfloor q/2\rfloor}\prod_{\nu=0}^{\alpha-1}\left(\frac{n-(\nu+1)}{n-(\nu+1)-nit}\right)^{q-(2j-1)}$$

$$\times\prod_{j=1}^{\lfloor p/2\rfloor-\alpha\lfloor q/2\rfloor}\left(\frac{n-(\alpha+1)}{n-(\alpha+1)-nit}\right)^{q-(2j-1)}$$

where, for even q, and for γ_k and α_k given respectively by (A.9) and (A.8),

i) $$\prod_{j=1}^{\lfloor q/2\rfloor}\prod_{\nu=0}^{\alpha-1}\prod_{l=2+\nu}^{\min(\alpha+1,2j+\nu q-1)}\left(\frac{n-l}{n-l-nit}\right)^{q}=\prod_{k=2}^{\alpha+1}\left(\frac{n-k}{n-k-nit}\right)^{\gamma_k},$$

ii) $$\prod_{j=1}^{\lfloor q/2\rfloor}\prod_{\nu=0}^{\alpha-1}\left(\frac{n-(\nu+1)}{n-(\nu+1)-nit}\right)^{q-(2j-1)}=\prod_{k=1}^{\alpha}\left(\frac{n-k}{n-k-nit}\right)^{a_k},$$

iii) $$\prod_{j=1}^{\lfloor p/2\rfloor-\alpha\lfloor q/2\rfloor}\left(\frac{n-(\alpha+1)}{n-(\alpha+1)-nit}\right)^{q-(2j-1)}=\left(\frac{n-(\alpha+1)}{n-(\alpha+1)-nit}\right)^{a_{\alpha+1}},$$

so that we may write $\Phi_{1,1}(t)$ as in (A.7). □

LEMMA 5. *For odd q, and once again for $\alpha=\left\lfloor\frac{p-1}{q}\right\rfloor$, we may write*

(A.13) $$\Phi_{1,1}(t) = \prod_{j=\alpha+2}^{2\lfloor p/2\rfloor-1}\left(\frac{(n-j)/n}{(n-j)/n-\mathrm{it}}\right)^{q(\lfloor p/2\rfloor-\lfloor j/2\rfloor)} \times\prod_{k=1}^{\alpha+1}\left(\frac{(n-k)/n}{(n-k)/n-\mathrm{it}}\right)^{a_k+\gamma_k}$$

where

(A.14) $$\begin{cases} a_k=\lfloor\frac{q}{2}\rfloor\left\lfloor\frac{q+k\perp 2}{2}\right\rfloor, \quad k=1,\dots,\alpha \\ a_{\alpha+1}=\left(q-\left(\lfloor\frac{p}{2}\rfloor-\alpha\lfloor\frac{q}{2}\rfloor-\lfloor\frac{\alpha}{2}\rfloor\right)\right)\left(\lfloor\frac{p}{2}\rfloor-\alpha\lfloor\frac{q}{2}\rfloor-\left\lfloor\frac{\alpha+1}{2}\right\rfloor\right), \end{cases}$$

and

(A.15) $$\gamma_k=\lfloor q/2\rfloor(k-1)q \quad (k=1,\dots,\alpha+1)\,.$$

PROOF. Following the same lines used in the proof of Lemma 3, we obtain once again expressions (A.10) and (A.11). Then, taking $\mathsf{Prod}(var;\alpha,q,j)$ given by (A.12) and taking into account that $\lfloor 1+\frac{\nu}{2}\rfloor - \left\lfloor\frac{\nu+1}{2}\right\rfloor = (\nu+1)\perp\!\!\!\perp 2$ and that for odd q, $2\left\lfloor\frac{q}{2}\right\rfloor = q-1$, we have,

$$
\begin{aligned}
\prod_{j=1}^{\lfloor p/2\rfloor}\prod_{k=1}^{q}\prod_{l=1-\left\lfloor\frac{k-2j}{q}\right\rfloor}^{\lfloor(\alpha+1,2j-1)}\frac{n-l}{n-l-nit}
&= \prod_{\nu=0}^{\alpha-1}\prod_{j=1+\nu\lfloor q/2\rfloor+\lfloor(1+\nu)/2\rfloor}^{(\nu+1)\lfloor q/2\rfloor+\lfloor 1+\nu/2\rfloor}\mathsf{Prod}\left(\nu;\alpha,q+1,j+\tfrac{1}{2}\right)\\
&\qquad\times\prod_{j=1+\alpha\lfloor q/2\rfloor+\lfloor(1+\alpha)/2\rfloor}^{\lfloor p/2\rfloor}\mathsf{Prod}\left(\nu;\alpha,q+1,j+\tfrac{1}{2}\right)\\
&= \left[\prod_{\nu=0}^{\alpha-1}\prod_{j=1+\lfloor(1+\nu)/2\rfloor}^{\lfloor q/2\rfloor+\lfloor 1+\nu/2\rfloor}\left\{\prod_{l=1+\nu}^{\lfloor(\alpha+1,\varphi(\nu,j,q))}\left(\frac{n-l}{n-l-nit}\right)^{q-(2j-(\nu+1))}\right.\right.\\
&\qquad\left.\left.\times\prod_{l=2+\nu}^{\lfloor(\alpha+1,\varphi(\nu,j,q))}\left(\frac{n-l}{n-l-nit}\right)^{2j-(\nu+1)}\right\}\right]\\
&\qquad\times\prod_{j=1+\lfloor(1+\alpha)/2\rfloor}^{\lfloor p/2\rfloor-\alpha\lfloor q/2\rfloor}\left(\frac{n-(\alpha+1)}{n-(\alpha+1)-nit}\right)^{q-(2j-(\alpha+1))}\\
&= \left\{\prod_{\nu=0}^{\alpha-1}\prod_{j=1+\left\lfloor\frac{1+\nu}{2}\right\rfloor}^{\left\lfloor\frac{q}{2}\right\rfloor+\left\lfloor 1+\frac{\nu}{2}\right\rfloor}\prod_{l=2+\nu}^{\lfloor(\alpha+1,2j+\nu q-\nu-1)}\left(\frac{n-l}{n-l-nit}\right)^{q}\right\}\\
&\qquad\times\prod_{\nu=0}^{\alpha-1}\prod_{j=1+\lfloor(1+\nu)/2\rfloor}^{\lfloor q/2\rfloor+\lfloor 1+\nu/2\rfloor}\left(\frac{n-(\nu+1)}{n-(\nu+1)-nit}\right)^{q-2j+\nu+1}\\
&\qquad\times\prod_{j=1+\lfloor(1+\alpha)/2\rfloor}^{\lfloor p/2\rfloor-\alpha\lfloor q/2\rfloor}\left(\frac{n-(\alpha+1)}{n-(\alpha+1)-nit}\right)^{q-2j+\alpha+1}
\end{aligned}
$$

where, for odd q,, and for γ_k and α_k given respectively by (A.15) and (A.14), and yet for $\delta = (\nu+1)\perp\!\!\!\perp 2$,

i) $$\prod_{\nu=0}^{\alpha-1} \prod_{j=1+\lfloor(1+\nu)/2\rfloor}^{\lfloor q/2\rfloor+\lfloor 1+\nu/2\rfloor} \prod_{l=2+\nu}^{\lfloor(\alpha+1,2j+\nu q-\nu-1)} \left(\frac{n-l}{n-l-n\mathrm{i}t}\right)^{q}$$
$$= \prod_{\nu=0}^{\alpha-1} \prod_{j=1}^{\lfloor q/2\rfloor+\delta} \prod_{l=2+\nu}^{\lfloor(\alpha+1,2j+\nu q-\delta} \left(\frac{n-l}{n-l-n\mathrm{i}t}\right)^{q} = \prod_{k=2}^{\alpha+1} \left(\frac{n-k}{n-k-n\mathrm{i}t}\right)^{\gamma_k},$$

ii) $$\prod_{\nu=0}^{\alpha-1} \prod_{j=1+\lfloor(1+\nu)/2\rfloor}^{\lfloor q/2\rfloor+\lfloor 1+\nu/2\rfloor} \left(\frac{n-(\nu+1)}{n-(\nu+1)-n\mathrm{i}t}\right)^{q-2j+\nu+1}$$
$$= \prod_{\nu=0}^{\alpha-1} \prod_{j=1}^{\lfloor q/2\rfloor+\delta} \left(\frac{n-(\nu+1)}{n-(\nu+1)-n\mathrm{i}t}\right)^{q-2j-\delta} = \prod_{k=1}^{\alpha} \left(\frac{n-k}{n-k-n\mathrm{i}t}\right)^{a_k},$$

iii) $$\prod_{j=1+\lfloor\frac{\alpha+1}{2}\rfloor}^{\lfloor p/2\rfloor-\alpha\lfloor q/2\rfloor} \left(\frac{n-(\alpha+1)}{(n-(\alpha+1)-n\mathrm{i}t}\right)^{q-2j+\alpha+1}$$
$$= \prod_{j=1}^{\lfloor\frac{p}{2}\rfloor-\alpha\lfloor\frac{q}{2}\rfloor-\lfloor\frac{\alpha+1}{2}\rfloor} \left(\frac{n-(\alpha+1)}{n-(\alpha+1)-n\mathrm{i}t}\right)^{q-2j-\alpha\perp 2+1} = \left(\frac{n-(\alpha+1)}{n-(\alpha+1)-n\mathrm{i}t}\right)^{a_{\alpha+1}},$$

so that we may write $\Phi_{1,1}(t)$ as in (A.13). □

Equalities i), ii) and iii) in the proofs of Lemmas 4 and 5 are quite straightforward to verify and may be proven by induction but the proofs are not shown since they are a bit long and tedious.

In the next Corollary we show how we can put together in one single expression the results from Lemmas 4 and 5, for both even and odd q.

Corollary. *For both even and odd q we may write*

(A.16)
$$\Phi_{1,1}(t) = \prod_{j=\alpha+2}^{2\lfloor p/2\rfloor-1} \left(\frac{(n-j)/n}{(n-j)/n-\mathrm{i}t}\right)^{q(\lfloor p/2\rfloor-\lfloor j/2\rfloor)}$$
$$\times \prod_{k=1}^{\alpha+1} \left(\frac{(n-k)/n}{(n-k)/n-\mathrm{i}t}\right)^{a_k+\gamma_k}$$

where

$$\text{(A.17)}\qquad \begin{cases} a_k = \lfloor \frac{q}{2} \rfloor \left\lfloor \frac{q+k\perp 2}{2} \right\rfloor , \qquad k=1,\ldots,\alpha \\ a_{\alpha+1} = (q-(\lfloor \frac{p}{2} \rfloor - \alpha \lfloor \frac{q}{2} \rfloor))\, (\lfloor \frac{p}{2} \rfloor - \alpha \lfloor \frac{q}{2} \rfloor) \\ \qquad\qquad + (q \perp\!\!\!\perp 2) \left(\alpha \lfloor \frac{p}{2} \rfloor - \alpha^2 \lfloor \frac{q}{2} \rfloor - \frac{\alpha^2}{4} + \frac{\alpha\perp 2}{4} - q \left\lfloor \frac{\alpha+1}{2} \right\rfloor \right) , \end{cases}$$

and, for $k=1,\ldots,\alpha+1$*,*

$$\text{(A.18)}\qquad \gamma_k = \left\lfloor \frac{q}{2} \right\rfloor \left((k-1)q - 2\,((q+1) \perp\!\!\!\perp 2) \left\lfloor \frac{k}{2} \right\rfloor \right) .$$

PROOF. Since for $k \in I\!N_0$ and even q we have $\left\lfloor \frac{q+k\perp 2}{2} \right\rfloor = \lfloor \frac{q}{2} \rfloor$, and since for even q we also have

$$\frac{q^2}{4} = \left\lfloor \frac{q}{2} \right\rfloor \left\lfloor \frac{q}{2} \right\rfloor ,$$

from (A.8) and (A.14), for any q, even or odd, we may write

$$a_k = \left\lfloor \frac{q}{2} \right\rfloor \left\lfloor \frac{q+k \perp\!\!\!\perp 2}{2} \right\rfloor , \qquad k=1,\ldots,\alpha\,,$$

while from (A.9) and (A.15) we may write (A.18). Finally, since

$$\left\lfloor \frac{\alpha+1}{2} \right\rfloor + \left\lfloor \frac{\alpha}{2} \right\rfloor = \alpha$$

and

$$\left\lfloor \frac{\alpha}{2} \right\rfloor \left\lfloor \frac{\alpha+1}{2} \right\rfloor = \begin{cases} \frac{\alpha^2}{4} & \alpha \text{ even} \\ \frac{\alpha^2-1}{4} & \alpha \text{ odd} \end{cases}$$

we have

$$\begin{aligned} &(q-(\lfloor \tfrac{p}{2} \rfloor - \alpha \lfloor \tfrac{q}{2} \rfloor - \lfloor \tfrac{\alpha}{2} \rfloor)) \left(\lfloor \tfrac{p}{2} \rfloor - \alpha \lfloor \tfrac{q}{2} \rfloor - \left\lfloor \tfrac{\alpha+1}{2} \right\rfloor \right) \\ &\quad = (q-(\lfloor \tfrac{p}{2} \rfloor - \alpha \lfloor \tfrac{q}{2} \rfloor))\, (\lfloor \tfrac{p}{2} \rfloor - \alpha \lfloor \tfrac{q}{2} \rfloor) - q \left\lfloor \tfrac{\alpha+1}{2} \right\rfloor \\ &\qquad + \lfloor \tfrac{p}{2} \rfloor \left\lfloor \tfrac{\alpha+1}{2} \right\rfloor - \alpha \lfloor \tfrac{q}{2} \rfloor \left\lfloor \tfrac{\alpha+1}{2} \right\rfloor - \lfloor \tfrac{\alpha}{2} \rfloor \left\lfloor \tfrac{\alpha+1}{2} \right\rfloor \\ &\qquad\qquad + \lfloor \tfrac{p}{2} \rfloor \lfloor \tfrac{\alpha}{2} \rfloor - \alpha \lfloor \tfrac{q}{2} \rfloor \lfloor \tfrac{\alpha}{2} \rfloor \\ &\quad = (q-(\lfloor \tfrac{p}{2} \rfloor - \alpha \lfloor \tfrac{q}{2} \rfloor))\, (\lfloor \tfrac{p}{2} \rfloor - \alpha \lfloor \tfrac{q}{2} \rfloor) - q \left\lfloor \tfrac{\alpha+1}{2} \right\rfloor \\ &\qquad\qquad + \lfloor \tfrac{p}{2} \rfloor\, \alpha - \lfloor \tfrac{q}{2} \rfloor\, \alpha^2 - \tfrac{\alpha^2}{4} + \tfrac{\alpha\perp 2}{4} . \end{aligned}$$

□

In the next Lemma we show how $\Phi^*_{1,2}(t)$ may also be seen as the c.f. of a GIG distribution by identifying the different Exponential distributions involved and obtaining analytic expressions for their counts.

LEMMA 6. *We may write, for $\Phi^*_{1,2}(t)$ defined as in Lemma 3,*

$$\Phi^*_{1,2}(t) = \left(\frac{(n-k^*)/n}{(n-k^*)/n - \mathrm{i}t}\right)^{\gamma(\alpha_2-\alpha_1)} \prod_{\substack{l=1+p-2\alpha_1 \\ \text{step } 2}}^{p-1} \left(\frac{(n-l)/n}{(n-l)/n-\mathrm{i}t}\right)^q, \tag{A.19}$$

where

$$\gamma = q - \left(\frac{p-1}{2} - \beta\right), \quad \text{with} \quad \beta = \left\lfloor \frac{p}{2q} \right\rfloor q, \tag{A.20}$$

$$\alpha_1 = \left\lfloor \frac{q-1}{q}\,\frac{p-1}{2} \right\rfloor, \quad \alpha_2 = \left\lfloor \frac{q-1}{q}\,\frac{p+1}{2} \right\rfloor, \tag{A.21}$$

and

$$k^* = 1 + p - 2\alpha_2\,. \tag{A.22}$$

PROOF. In Lemma 3 we have $b^*_{pk} = \lfloor b_{pk} \rfloor$, with

$$b_{pk} = \frac{pq-q-p+2k-1}{2q} = \frac{q-1}{q}\,\frac{p-1}{2} + \frac{k-1}{q}$$

so that, given that here p is odd and thus $\frac{p-1}{2}$ is an integer,

$$\begin{aligned} b^*_{pk} &= \left\lfloor \frac{q-1}{q}\,\frac{p-1}{2} + \frac{k-1}{q} \right\rfloor \\ &= \begin{cases} \alpha_1 & \text{for} \quad k = 1, \dots, \frac{p-1}{2} - \beta \\ \alpha_2 & \text{for} \quad k = \frac{p+1}{2} - \beta, \dots, q \end{cases} \end{aligned}$$

with β given by (A.20) and α_1 and α_2 given by (A.21).

We may thus write, taking into account the definitions of a_p in (A.3), γ

in (A.21) and α_1 and α_2 in (A.22),

$$\begin{aligned}
\Phi^*_{1,2}(t) &= \prod_{k=1}^{\frac{p-1}{2}-\beta} \prod_{l=1}^{\alpha_1} \frac{a_p+l-1}{a_p+l-1-\frac{n}{2}\mathrm{i}t} \prod_{k=\frac{p+1}{2}-\beta}^{q} \prod_{l=1}^{\alpha_2} \frac{a_p+l-1}{a_p+l-1-\frac{n}{2}\mathrm{i}t} \\
&= \prod_{l=1}^{\alpha_1} \left(\frac{a_p+l-1}{a_p+l-1-\frac{n}{2}\mathrm{i}t}\right)^{\frac{p-1}{2}-\beta} \prod_{l=1}^{\alpha_2} \left(\frac{a_p+l-1}{a_p+l-1-\frac{n}{2}\mathrm{i}t}\right)^{q-\left(\frac{p+1}{2}-\beta\right)} \\
&= \prod_{l=1}^{\alpha_1} \left(\frac{a_p+l-1}{a_p+l-1-\frac{n}{2}\mathrm{i}t}\right)^{q} \left(\frac{a_p+\alpha_2-1}{a_p+\alpha_2-1-\frac{n}{2}\mathrm{i}t}\right)^{\gamma(\alpha_2-\alpha_1)} \\
&= \prod_{l=1}^{\alpha_1} \left(\frac{(n-p-1+2l)/n}{(n-p-1+2l)/n-\mathrm{i}t}\right)^{q} \left(\frac{(n-p-1+2\alpha_2)/n}{(n-p-1+2\alpha_2)/n-\mathrm{i}t}\right)^{\gamma(\alpha_2-\alpha_1)} \\
&= \prod_{\substack{l^*=p+1-2\alpha_1 \\ \text{step } 2}}^{p-1} \left(\frac{(n-l^*)/n}{(n-l^*)/n-\mathrm{i}t}\right)^{q} \left(\frac{(n-k^*)/n}{(n-k^*)/n-\mathrm{i}t}\right)^{\gamma(\alpha_2-\alpha_1)}
\end{aligned}$$

where the last equality is obtained by taking $l^* = p+1-2l$ and k^* given by (A.22). □

From the above Corollary and Lemma 6, while for $k = 1,\dots,\alpha$, r_k is obtained just by adding γ_k in (A.18) and a_k in the first row of (A.17), for $k = \alpha+1$ we have, from (A.17),

$$\begin{aligned}
a_{\alpha+1} = &\left(q-\left(\left\lfloor\tfrac{p}{2}\right\rfloor-\alpha\left\lfloor\tfrac{q}{2}\right\rfloor\right)\right)\left(\left\lfloor\tfrac{p}{2}\right\rfloor-\alpha\left\lfloor\tfrac{q}{2}\right\rfloor\right) \\
&\qquad +(q \perp\!\!\!\perp 2)\left(\alpha\left\lfloor\tfrac{p}{2}\right\rfloor-\alpha^2\left\lfloor\tfrac{q}{2}\right\rfloor-\tfrac{\alpha^2}{4}+\tfrac{\alpha \perp\!\!\!\perp 2}{4}-q\left\lfloor\tfrac{\alpha+1}{2}\right\rfloor\right) \\
= &\; q\left\lfloor\tfrac{p}{2}\right\rfloor-q\alpha\left\lfloor\tfrac{q}{2}\right\rfloor-\left(\left\lfloor\tfrac{p}{2}\right\rfloor-\alpha\left\lfloor\tfrac{q}{2}\right\rfloor\right)^2-(q \perp\!\!\!\perp 2)q\left\lfloor\tfrac{\alpha+1}{2}\right\rfloor \\
&\qquad +(q \perp\!\!\!\perp 2)\left(\alpha\left\lfloor\tfrac{p}{2}\right\rfloor-\alpha^2\left\lfloor\tfrac{q}{2}\right\rfloor-\tfrac{\alpha^2}{4}+\tfrac{\alpha \perp\!\!\!\perp 2}{4}\right)
\end{aligned}$$

while, from (A.18),

$$\gamma_{\alpha+1} = q\alpha\left\lfloor\frac{q}{2}\right\rfloor - 2\left\lfloor\frac{q}{2}\right\rfloor\left\lfloor\frac{\alpha+1}{2}\right\rfloor\left((q+1) \perp\!\!\!\perp 2\right),$$

where

$$2\left\lfloor\tfrac{q}{2}\right\rfloor\left\lfloor\tfrac{\alpha+1}{2}\right\rfloor\left((q+1) \perp\!\!\!\perp 2\right) = \left((q+1) \perp\!\!\!\perp 2\right) q\left\lfloor\tfrac{\alpha+1}{2}\right\rfloor$$

so that $a_{\alpha+1}+\gamma_{\alpha+1}$ comes given by (3.5), since

$$(q \perp\!\!\!\perp 2)q\left\lfloor\tfrac{\alpha+1}{2}\right\rfloor + \left((q+1) \perp\!\!\!\perp 2\right) q\left\lfloor\tfrac{\alpha+1}{2}\right\rfloor = q\left\lfloor\tfrac{\alpha+1}{2}\right\rfloor .$$

For $k=\alpha+2,\dots,\min(p-2\alpha_1,p-1)$ and $k=2+p-2\alpha_1,\dots,2\left\lfloor\frac{p}{2}\right\rfloor-1$, with step 2, we have to consider the result in the first row in (A.16), while for $k=1+p-2\alpha_1,\dots,p-1$, with step 2, we have to consider this same result together with the result in Lemma 6 and notice that

$$\Phi_{1,2}(t)=\left(\Phi^*_{1,2}(t)\right)^{p\perp 2}$$

and that, as such, the exponent q in (A.19) in Lemma 6 only appears for odd p and that

$$q\left(\left\lfloor\frac{p}{2}\right\rfloor-\left\lfloor\frac{k}{2}\right\rfloor\right)+q(p\perp\!\!\!\perp 2)\;=\;q\left(\left\lfloor\frac{p+1}{2}\right\rfloor-\left\lfloor\frac{k}{2}\right\rfloor\right)\,.$$

Finally, from (A.19) and (A.22) in Lemma 6, we see that for $k=p+1-2\alpha_2$ we have to add, for odd p, $\gamma(\alpha_2-\alpha_1)$, where $\gamma=q-\left(\frac{p-1}{2}-q\left\lfloor\frac{p}{2q}\right\rfloor\right)$, to the value of r_k to obtain r^*_k. It happens that the only possible values for $\alpha_2-\alpha_1$ are either zero or 1, so that it will be only for $\alpha_2-\alpha_1=1$ or $1+p-2\alpha_2=p-1-2\alpha_1$ that we will have to add $\gamma(\alpha_2-\alpha_1)$ to r_k.

This concludes the proof of Theorem 3.2.

APPENDIX B: PLOTS OF P.D.F.'S AND C.D.F.'S OF BOX'S ASYMPTOTIC DISTRIBUTION FOR $W = -\mathrm{LOG}(\lambda^*)$

Figures 5 and 6 in this Appendix show plots of p.d.f.'s and c.d.f.'s of Box's asymptotic distribution corresponding to the c.f. in (2.1) for $W = -\log(\lambda^*)$, for some combinations of p, q and n for which this distribution is not a proper distribution. In all the four cases presented the p.d.f. displays some negative values and consequently for $p = 5, 7$ and 10 the c.d.f. assumes values above 1, while for $p = 50$ it has negative values for smaller values of the argument.

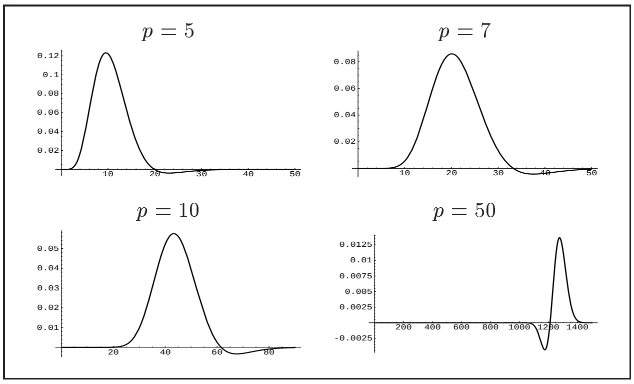

FIG 5. *Plots of p.d.f.'s of Box's asymptotic distribution for* $W = -\log(\lambda^*)$ *for* $q = 2$ *and* $n = p + 2$.

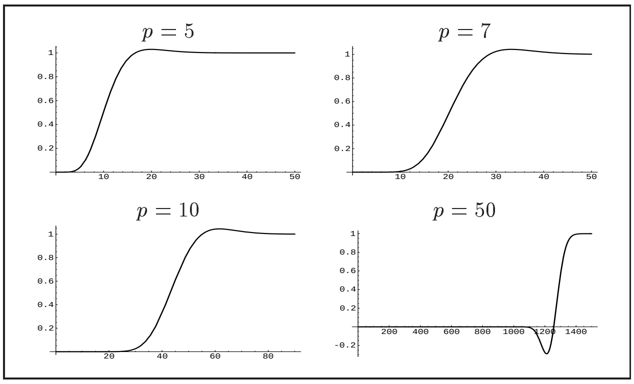

FIG 6. *Plots of c.d.f.'s of Box's asymptotic distribution for* $W = -\log(\lambda^*)$ *for* $q = 2$ *and* $n = p + 2$.

APPENDIX C: TABLES WITH VALUES OF Δ_1

Tables 7 through 10 in this Appendix have the values of the measure Δ_1 for the asymptotic Box, G, M2G and M3G and the near-exact GNIG, M2GNIG and M3GNIG distributions corresponding to the cases treated in Section 5.

TABLE 7
Values of the measure Δ_1 for increasing values of p, with small sample sizes

p	q	n	Box	G	M2G	M3G	GNIG	M2GNIG	M3GNIG
3	2	5	5.100×10^{-2}	1.942×10^{-3}	1.057×10^{-4}	1.378×10^{-5}	3.926×10^{-4}	8.756×10^{-6}	8.756×10^{-6}
4	2	6	5.151×10^{-2}	1.513×10^{-3}	1.134×10^{-4}	9.118×10^{-6}	6.086×10^{-5}	1.413×10^{-7}	1.085×10^{-9}
5	2	7	5.094×10^{-2}	1.229×10^{-3}	1.128×10^{-4}	1.261×10^{-5}	8.648×10^{-6}	1.461×10^{-7}	8.152×10^{-9}
6	2	8	4.955×10^{-2}	1.034×10^{-3}	1.057×10^{-4}	1.378×10^{-5}	1.991×10^{-5}	1.904×10^{-7}	1.932×10^{-9}
7	2	9	4.773×10^{-2}	8.912×10^{-4}	9.760×10^{-5}	1.394×10^{-5}	2.653×10^{-5}	3.726×10^{-7}	6.828×10^{-9}
8	2	10	4.571×10^{-2}	7.816×10^{-4}	8.984×10^{-5}	1.365×10^{-5}	7.580×10^{-6}	5.747×10^{-8}	5.580×10^{-10}
9	2	11	4.362×10^{-2}	6.947×10^{-4}	8.276×10^{-5}	1.314×10^{-5}	4.152×10^{-6}	1.751×10^{-8}	6.154×10^{-11}
10	2	12	4.152×10^{-2}	6.239×10^{-4}	7.642×10^{-5}	1.256×10^{-5}	3.431×10^{-6}	1.898×10^{-8}	1.447×10^{-10}
20	2	22	2.377×10^{-2}	2.943×10^{-4}	4.022×10^{-5}	——×—	2.666×10^{-7}	4.324×10^{-10}	1.012×10^{-12}
35	2	37	1.463×10^{-2}	1.526×10^{-4}	2.142×10^{-5}	——×—	3.974×10^{-8}	2.605×10^{-11}	2.430×10^{-14}
50	2	52	1.523×10^{-2}	9.874×10^{-5}	1.382×10^{-5}	——×—	8.971×10^{-9}	2.741×10^{-12}	1.161×10^{-15}

TABLE 8
Values of the measure Δ_1 for increasing values of q, with small sample sizes

p	q	n	Box	G	M2G	M3G	GNIG	M2GNIG	M3GNIG
3	2	5	5.100×10^{-2}	1.942×10^{-3}	1.057×10^{-4}	1.378×10^{-5}	3.926×10^{-4}	8.756×10^{-6}	8.756×10^{-6}
3	5	5	2.929×10^{-2}	4.239×10^{-4}	1.938×10^{-5}	1.007×10^{-6}	1.566×10^{-4}	4.447×10^{-6}	1.830×10^{-7}
3	7	5	2.687×10^{-2}	2.924×10^{-4}	1.280×10^{-5}	6.685×10^{-7}	8.435×10^{-5}	1.872×10^{-6}	6.352×10^{-8}
3	10	5	2.524×10^{-2}	2.006×10^{-4}	8.247×10^{-6}	4.191×10^{-7}	4.238×10^{-5}	6.879×10^{-7}	1.771×10^{-8}
10	2	12	4.152×10^{-2}	6.239×10^{-4}	7.642×10^{-5}	1.256×10^{-5}	3.431×10^{-6}	1.898×10^{-8}	1.447×10^{-10}
10	5	12	2.563×10^{-2}	2.168×10^{-4}	2.008×10^{-5}	——×—	7.559×10^{-8}	2.407×10^{-11}	5.573×10^{-15}
10	7	12	2.257×10^{-2}	1.532×10^{-4}	1.292×10^{-5}	——×—	2.875×10^{-8}	6.034×10^{-12}	1.288×10^{-15}
10	10	12	1.974×10^{-2}	1.066×10^{-4}	8.127×10^{-6}	——×—	9.588×10^{-9}	4.935×10^{-13}	9.298×10^{-17}

Table 9

Values of the measure Δ_1 for increasing sample sizes

p	q	n	Box	G	M2G	M3G	GNIG	M2GNIG	M3GNIG
3	2	5	5.100×10^{-2}	1.942×10^{-3}	1.057×10^{-4}	1.378×10^{-5}	3.926×10^{-4}	8.756×10^{-6}	8.756×10^{-6}
3	2	20	2.355×10^{-3}	1.111×10^{-4}	5.564×10^{-8}	——×—	3.040×10^{-5}	2.849×10^{-7}	3.113×10^{-9}
3	2	50	3.581×10^{-4}	1.724×10^{-5}	9.763×10^{-9}	——×—	4.826×10^{-6}	1.498×10^{-8}	1.768×10^{-11}
7	2	9	4.773×10^{-2}	8.912×10^{-4}	9.760×10^{-5}	1.394×10^{-5}	2.653×10^{-5}	3.726×10^{-7}	6.828×10^{-9}
7	2	20	6.963×10^{-3}	1.138×10^{-4}	1.980×10^{-6}	3.425×10^{-8}	7.832×10^{-6}	2.157×10^{-8}	4.763×10^{-11}
7	2	50	9.812×10^{-4}	1.596×10^{-5}	4.175×10^{-8}	5.771×10^{-11}	1.309×10^{-6}	5.310×10^{-10}	3.406×10^{-14}
10	2	12	4.152×10^{-2}	6.239×10^{-4}	7.642×10^{-5}	1.256×10^{-5}	3.431×10^{-6}	1.898×10^{-8}	1.447×10^{-10}
10	2	50	1.506×10^{-3}	1.623×10^{-5}	7.965×10^{-8}	——×—	5.258×10^{-7}	2.051×10^{-10}	7.399×10^{-14}
10	2	100	3.559×10^{-4}	3.825×10^{-6}	4.684×10^{-9}	——×—	1.387×10^{-7}	1.326×10^{-11}	1.141×10^{-15}

Table 10

Values of the measure Δ_1 for large q and large sample sizes

p	q	n	Box	G	M2G	M3G	GNIG	M2GNIG	M3GNIG
3	10	20	1.268×10^{-3}	9.834×10^{-6}	2.872×10^{-8}	6.017×10^{-10}	1.776×10^{-6}	4.018×10^{-9}	1.588×10^{-11}
3	10	50	1.951×10^{-4}	1.494×10^{-6}	7.791×10^{-10}	6.342×10^{-12}	2.602×10^{-7}	2.040×10^{-10}	3.269×10^{-13}
7	10	20	4.219×10^{-3}	1.550×10^{-5}	1.794×10^{-7}	——×—	1.503×10^{-7}	6.064×10^{-11}	3.831×10^{-14}
7	10	50	6.052×10^{-4}	2.039×10^{-6}	4.654×10^{-9}	——×—	2.946×10^{-8}	3.176×10^{-12}	1.222×10^{-15}
10	10	50	9.398×10^{-4}	2.203×10^{-6}	7.933×10^{-9}	——×—	3.137×10^{-9}	1.828×10^{-14}	——×—
10	10	100	2.229×10^{-4}	5.052×10^{-7}	5.569×10^{-10}	——×—	9.023×10^{-10}	5.274×10^{-15}	——×—

REFERENCES

ALBERTO, R. P. and COELHO, C. A. (2007). Study of the quality of several asymptotic and near-exact approximations based on moments for the distribution of the Wilks Lambda statistic. *J. Statist. Plann. Inference* **137** 1612–1626.

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

ANDERSON, T. W., FANG, K.-T. and HSU, H. (1986). Maximum-likelihood estimates and likelihood-ratio criteria for multivariate elliptically contoured distributions. *Canad. J. Statist.* **14** 55–59.

ANDERSON, T. W. and FANG, K.-T. (1990). Inference in multivariate elliptically contoured distributions based on maximum likelihood. *Statistical Inference in Elliptically Contoured and Related Distributions*, K.-T. Fang, T. W. Anderson, eds., 201–216.

BARTLETT, M. S. (1937). Properties of sufficiency and statistical tests. *Proc. R. Soc. Lond. Ser. A Math. Phys. Sci.* **160** 268–282.

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

BOX, G. E. P. (1949). A general distribution theory for a class of likelihood criteria. *Biometrika* **36** 317–346.

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. (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., ALBERTO, R. P., GRILO, L. M. (2006). A mixture of Generalized Integer Gamma distributions as the exact distribution of the product of an odd number of independent Beta random variables. *J. Interdiscip. Math.* **9** 229–248.

ESSEEN, C.-G. (1945). Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian Law. *Acta Math.* **77** 1–125.

GIL-PELAEZ, J. (1951). Note on the inversion theorem. *Biometrika* **38** 481–482.

GRILO, L. M. and COELHO, C. A. (2007). Development and study of two near-exact approximations to the distribution of the product of an odd number of independent Beta random variables. *J. Statist. Plann. Inference* **137** 1560–1575.

HWANG, H.-K. (1998). On convergence rates in the central limit theorems for combinatorial structures. *European J. Combin.* **19** 329–343.

JAMSHIDIANA, M. and SCHOTT, J. R. (2007). Testing equality of covariance matrices when data are incomplete. *Comput. Statist. Data Anal.* **51** 4227–4239.

JENSEN, D. R. and GOOD, I. J. (1981). Invariant distributions associated with matrix laws under structural symmetry. *J. Roy. Statist. Soc. Ser. B* **43** 327–332.

JENSEN, J. L. (1991). A large deviation-type approximation for the 'Box class' of likelihood ratio criteria. *J. Amer. Statist. Assoc.* **86** 437–440.

JENSEN, J. L. (1995). Correction to 'A large deviation-type approximation for the 'Box class' of likelihood ratio criteria'. *J. Amer. Statist. Assoc.* **90** 812.

KARIYA, T. (1981). Robustness of multivariate tests. *Ann. Statist.* **9** 1267–1275.

LO, Y. (2008). A likelihood ratio test of a homoscedastic normal mixture against a heteroscedastic normal mixture. *Stat. Comput.* **18** 233–240.

LOÈVE, M. (1977). *Probability Theory*, Vol. I, 4th ed. Springer, New York.

MARQUES, F. J. and COELHO, C. A.. Near-exact distributions for the sphericity likelihood ratio test statistic. *J. Statist. Plann. Inference* **138** 726–741.

MUIRHEAD, R. J. (1982). *Aspects of Multivariate Statistical Theory.* Wiley, New York.

Rigby, R. A. (1997). Bayesian Discrimination Between Two Multivariate Normal Populations With Equal Covariance Matrices. *J. Amer. Statist. Assoc.* **92** 1151–1154.

Srivastava, M. S. (1967). Classification into multivariate normal populations when the population means are linearly restricted. *Ann. Inst. Statist. Math.* **19** 473–478.

Zhang, J. and Boos, D. (1992). Bootstrap critical values for testing homogeneity of covariance matrices. *J. Amer. Statist. Assoc.* **87** 425–429.