Development and comparative study of two near-exact approximations to the distribution of the product of an odd number of independent Beta random variables

Luís M. Grilo [a,1,*] Carlos A. Coelho [b,1]

[a]*Dep. of Mathematics, Polytechnic Institute of Tomar, Portugal*
lgrilo@ipt.pt

[b]*Mathematics Department, Faculty of Sciences and Technology – The New University of Lisbon, Portugal*
cmac@fct.unl.pt

Abstract

Using the concept of near-exact approximation to a distribution we developed two different near-exact approximations to the distribution of the product of an odd number of particular independent Beta random variables. One of them is a particular Generalized Near-Integer Gamma (GNIG) distribution and the other is a mixture of two GNIG distributions. These near-exact distributions are mostly adequate to be used as a basis for approximations of distributions of several statistics used in Multivariate Analysis. By factoring the characteristic function of the logarithm of the product of the Beta random variables, and then replacing a suitably chosen factor of that characteristic function by an adequate asymptotic result it is possible to obtain what we call a near-exact characteristic function, which gives rise to the near-exact approximation to the exact distribution. Depending on the asymptotic result used to replace the chosen parts of the characteristic function, one may obtain different near-exact approximations. Moments from the two near-exact approximations developed are compared with the exact ones. The two approximations are also compared with each other, namely in terms of moments and quantiles.

Key words: Near-exact approximations, Generalized Near-integer Gamma distribution, independent random variables, mixtures, Multivariate Analysis.

1 Introduction

It is our aim to obtain a close approximation to the exact distribution of the product of an odd number of independent r.v.'s (random variables) with Beta distributions with parameters yielding particular relationships, in a concise manageable form, so that at least the computation of quantiles is rendered reasonably easy.

Let

$$Y_j \sim Beta\left(a_j, \frac{b}{2}\right) \qquad j = 1, \ldots, p$$

be independent Beta r.v.'s with parameters a_j and $b/2$, with p and b positive odd integers, where $a_j = c + p/2 - j/2 \quad (j = 1, \ldots, p)$, $c \in I\!\!R^+$ (where $I\!\!R^+$ is the set of positive reals). We are then interested in the distribution of

$$W' = \prod_{j=1}^{p} Y_j \qquad \text{or} \qquad W = -\log W' = -\sum_{j=1}^{p} \log Y_j \,.$$

We may note that a few likelihood ratio test statistics used in Multivariate Analysis, such as the Wilks Λ statistic (Wilks, 1932, 1935) have a distribution of this kind (Anderson, 1984). However, the exact form of the distribution of W and W' is too complicated for practical use, since it usually involves hypergeometric functions of higher order. The exact form of the distribution of W', as well as the distribution of the sum of independent Gamma r.v.'s with different rate parameters, have already been studied in the past by several authors but all results were obtained under the form of series expansions (Kabe, 1962; Tretter and Walster, 1975; Nandi, 1977; Gupta and Richards, 1979, 1983).

When p and b are both even, or only one of them is odd, although some of the series expansions have very good convergence properties, namely those based on Chi-square mixtures (Gupta and Richards, 1983), the exact distribution for W and W' was obtained in a concise manageable form, without any series expansions by Coelho (1998, 1999). In Coelho (1998, 1999) the distribution of W' is presented as a particular case of the GIG (Generalized Integer Gamma) distribution, that is the distribution of the sum of independent Gamma r.v.'s with different rate parameters and integer shape parameters.

When p and b are both odd, Coelho (2004) obtained a near-exact approximation for the distribution of W under the form of a GNIG (Generalized Near-Integer Gamma) distribution, that is the distribution of the sum of a GIG r.v. with an independent Gamma r.v. with non-integer shape parameter. These approximations were built in the following way: after factorizing the c.f. (characteristic function) of W we replace one of the factors, by an

approximation with good convergence properties, in such a way that we obtain a near-exact c.f. that corresponds to a known distribution. In this paper we obtain two other near-exact approximations for the distribution of W and W', both of them even closer to the exact distribution. The factor that is replaced corresponds to the c.f. of a Logbeta r.v. (a r.v. whose exponential has a Beta distribution), being in one case approximated by the c.f. of the sum of two Gamma r.v.'s that matches the first three exact moments and in the other case approximated by the c.f. of the mixture of two Gamma r.v.'s that matches the first four exact moments. By joining this factor, with the remaining unchanged part of the c.f., we get what we call a near-exact c.f. for W. For the first case this c.f. corresponds to a GNIG distribution, while in the second case it corresponds to a mixture of two GNIG distributions.

In the following section we set forth some distributions that will be used in the sections ahead.

2 Some useful distributions

2.1 The GNIG distribution and the mixture of two GNIG distributions

Let Z_1 be a r.v. with a GIG distribution (Coelho, 1998) of depth g, with shape parameters $r_1, \ldots, r_g \in I\!N$ (where $I\!N$ is the set of positive integers) and all different rate parameters $\lambda_1, \ldots, \lambda_g \in I\!R^+$, that is,

$$Z_1 \sim GIG\left(r_1, \ldots, r_g; \lambda_1, \ldots, \lambda_g\right)$$

and let Z_2 be a r.v. with a Gamma distribution with positive non-integer shape parameter r and rate parameter $\lambda \in I\!R^+$, that is,

$$Z_2 \sim G(r, \lambda).$$

Furthermore, let Z_1 and Z_2 be independent and $\lambda \neq \lambda_j$ $(j = 1 \ldots, g)$. Then the distribution of

$$Z = Z_1 + Z_2$$

is a GNIG distribution (Coelho, 2004) of depth $g + 1$. We will denote this by

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

The p.d.f. (probability density function) of Z is given by

$$f_Z(z) = K\lambda^r \sum_{j=1}^{g} e^{-\lambda_j z} \sum_{k=1}^{r_j} \left\{ c_{j,k} \frac{\Gamma(k)}{\Gamma(k+r)} z^{k+r-1} {}_1F_1(r, k+r, -(\lambda-\lambda_j)z) \right\}, \tag{1}$$

$$(z > 0)$$

3

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

$$
\begin{aligned}
F_Z(z) = \lambda^r \frac{z^r}{\Gamma(r+1)} {}_1F_1(r, r+1, -\lambda z) \\
- K\lambda^r \sum_{j=1}^{g} e^{-\lambda_j z} \sum_{k=1}^{r_j} c_{j,k}^* \sum_{i=0}^{k-1} \frac{z^{r+i}\lambda_j^i}{\Gamma(r+1+i)} {}_1F_1(r, r+1+i, -(\lambda - \lambda_j)z) \quad (2)
\end{aligned}
$$

$$(z > 0)$$

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_{jk} given by (11) through (13) in Coelho (1998). In the above expressions $_1F_1(a, b; z)$ is the Kummer confluent hypergeometric function (Abramowitz and Stegun, 1974). Such functions have usually very good convergence properties and are nowadays handled by a number of software packages.

We may note that if $r \in I\!N$, the GNIG distribution turns itself into a GIG distribution, so that we may look at the GNIG distribution as a generalization of the GIG distribution.

Let us suppose now that the r.v. W has a distribution that is a mixture of two GNIG distributions of depth g, the first one with shape parameters $r_1^*, \ldots, r_g^*$ and rate parameters $\lambda_1^*, \ldots, \lambda_g^*$, with weight θ $(0 < \theta < 1)$ and the second one with shape parameters $r_1, \ldots, r_g$ and rate parameters $\lambda_1, \ldots, \lambda_g$, with weight $1 - \theta$. We will denote this fact by

$$
W \sim MGNIG\left(\theta; r_1^*, \ldots, r_g^*, \lambda_1^*, \ldots, \lambda_g^*; r_1, \ldots, r_g, \lambda_1, \ldots, \lambda_g\right).
$$

2.2   The Logbeta distribution

Let now X be a r.v. with a Beta distribution with parameters $\alpha > 0$ and $\beta > 0$. We will denote this by

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

The h-th moment of X is given by

$$
E(X^h) = \frac{B(\alpha + h, \beta)}{B(\alpha, \beta)} = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)} \frac{\Gamma(\alpha + h)}{\Gamma(\alpha + \beta + h)} \quad (\alpha + h > 0). \qquad (3)
$$

If we take $Y = -\log X$ then we will say that Y has a Logbeta distribution with parameters α and β (Johnson, Kotz and Balakrishnan, 1995), what is denoted by

$$Y \sim Logbeta(\alpha, \beta).$$

The p.d.f. of Y is

$$f_Y(y) = \frac{1}{B(\alpha, \beta)} \, e^{-\alpha y}(1 - e^{-y})^{\beta - 1}, \qquad y > 0. \tag{4}$$

Since the Gamma functions in (3) are still defined for any real or complex $h \in V_\epsilon(0)$, the c.f. of Y is given by

$$\phi_Y(t) = E(e^{itY}) = E(e^{-it \log X}) = E(X^{-it}) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)} \frac{\Gamma(\alpha - it)}{\Gamma(\alpha + \beta - it)}, \tag{5}$$

where $i = (-1)^{1/2}$ and $t \in I\!\!R$ (where $I\!\!R$ is the real set). Based on (5), we may obtain the first four moments of Y as

$$\mu_1' = E(Y) = \psi(\alpha + \beta) - \psi(\alpha)$$

$$\mu_2' = E(Y^2) = \psi'(\alpha) - \psi'(\alpha + \beta) + [\psi(\alpha + \beta) - \psi(\alpha)]^2$$

$$\mu_3' = E(Y^3) = \psi''(\alpha + \beta) - \psi''(\alpha) + [\psi(\alpha + \beta) - \psi(\alpha)]^3$$

$$+ 3 [\psi(\alpha + \beta) - \psi(\alpha)] [\psi'(\alpha) - \psi'(\alpha + \beta)] \tag{6}$$

$$\mu_4' = E(Y^4) = \psi'''(\alpha) - \psi'''(\alpha + \beta) + [\psi(\alpha) - \psi(\alpha + \beta)]^4$$

$$+ 6 [\psi(\alpha) - \psi(\alpha + \beta)]^2 [\psi'(\alpha) - \psi'(\alpha + \beta)]$$

$$+ 4 [\psi(\alpha) - \psi(\alpha + \beta)] [\psi''(\alpha) - \psi''(\alpha + \beta)]$$

$$+ 3 [\psi'(\alpha) - \psi'(\alpha + \beta)]^2,$$

where $\psi(x) = \frac{d}{dx} \log \Gamma(x)$ is the digamma function, $\psi'(x) = \frac{d^2}{dx^2} \log \Gamma(x) = \frac{d}{dx} \psi(x)$ is the trigamma function and $\psi''(x) = \frac{d}{dx} \psi'(x)$ is the quadrigamma function, and so on.

3 Two near-exact approximations to the distribution of the product of particular independent Beta random variables

In this section we will obtain two near-exact approximations for the distribution of the product of an odd number of independent Beta r.v.'s with same second parameter and whose first parameter evolves by 1/2.

Theorem 1 *Let*

$$Y_j \sim Beta\left(a_j, \frac{b}{2}\right), \qquad j = 1, ..., p \tag{7}$$

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

$$W_1' = \prod_{j=1}^{p} Y_j \qquad and \qquad W_1 = -\log W_1' = -\sum_{j=1}^{p} \log Y_j. \tag{8}$$

Let us further consider

$$Y_j^* \sim Beta\left(a_j^*, \frac{p}{2}\right), \qquad j = 1, ..., b$$

b independent r.v.'s where $a_j^* = c + \frac{b}{2} - \frac{j}{2}$ $(j = 1, ..., b)$ and let

$$W_2' = \prod_{j=1}^{b} Y_j^* \qquad and \qquad W_2 = -\log W_2' = -\sum_{j=1}^{b} \log Y_j^*.$$

Then a near-exact approximation to the distribution of W_1 and W_2 is a GNIG distribution of depth $p + b - 1$, symbolically,

$$W_k \stackrel{ne}{\approx} GNIG\left(r_1^*, ..., r_{p+b-3}^*, r_{p+b-2}^* = 1, r_{p+b-1}^*; \lambda_1, ..., \lambda_{p+b-3}, \lambda_{p+b-2}, \lambda_{p+b-1}\right),$$
$$(k = 1, 2) \tag{9}$$

with rate parameters

$$\lambda_j = c + \frac{j}{2} - \frac{1}{2}, \qquad (j = 1, ..., p + b - 3) \tag{10}$$

and shape parameters

$$r_j^* = \begin{cases} r_j & j = 1, ..., p - 1 \\ & j = p + 1, ..., p + b - 4 \quad step\ 2 \\ & j = p + b - 3 \\ r_j + 1 & j = p, ..., p + b - 5 \quad step\ 2 \end{cases} \tag{11}$$

(with sequences indexed by j being used only if the upper bound is larger than the lower bound), where

$$r_j = \begin{cases} h_j & j = 1, 2 \\ r_{j-2} + h_j & j = 3, ..., p + b - 3 \end{cases} \qquad (12)$$

with

$$h_j = \begin{cases} 1 & j = 1, ..., \min(p - 1, b) \\ 0 & j = 1 + \min(p - 1, b), ..., \max(p - 1, b) \\ -1 & j = 1 + \max(p - 1, b), ..., p + b - 3 \end{cases} \qquad (13)$$

and yet r^*_{p+b-1}, λ_{p+b-2} and λ_{p+b-1} obtained by numerical solution of the system of equations

$$\begin{cases} \mu'_1 = \dfrac{1}{\lambda_{p+b-2}} + \dfrac{r^*_{p+b-1}}{\lambda_{p+b-1}} \\[4mm] \mu'_2 = \dfrac{2\lambda^2_{p+b-1} + 2\lambda_{p+b-2}\lambda_{p+b-1}r^*_{p+b-1} + \lambda^2_{p+b-2}r^*_{p+b-1}(1+r^*_{p+b-1})}{\lambda^2_{p+b-2}\lambda^2_{p+b-1}} \\[4mm] \mu'_3 = \dfrac{6\lambda^3_{p+b-1} + 6\lambda_{p+b-2}\lambda^2_{p+b-1}r^*_{p+b-1} + 3\lambda^2_{p+b-2}\lambda_{p+b-1}r^*_{p+b-1}(1+r^*_{p+b-1})}{\lambda^3_{p+b-2}\lambda^3_{p+b-1}} \\[4mm] \qquad + \dfrac{\lambda^3_{p+b-2}r^*_{p+b-1}(2+3r^*_{p+b-1}+r^{*2}_{p+b-1})}{\lambda^3_{p+b-2}\lambda^3_{p+b-1}} \end{cases} \qquad (14)$$

where on the left hand side we have the first three moments of a Logbeta r.v. with parameters $a_1 + \frac{b}{2} - \frac{3}{2}$ and $\frac{3}{2}$, obtained from (6) by replacing α and β by the appropriate values, and on the right hand side the first three moments of the sum of two Gamma r.v.'s, the first one with shape parameter $r^*_{p+b-2} = 1$ and rate parameter λ_{p+b-2} and the second one with shape parameter r^*_{p+b-1} and rate parameter λ_{p+b-1}.

The other near-exact approximation to the distribution of W_1 and W_2 is obtained as a mixture of two GNIG distributions of depth $p+b-2$, symbolically,

$$W_k \overset{ne}{\approx} MGNIG\left(\theta; r^*_1, ..., r^*_{p+b-3}, r_{p+b-2}; \lambda_1, ..., \lambda_{p+b-3}, \lambda_{p+b-2};\right.$$

$$\left. r^*_1, ..., r^*_{p+b-3}, r'_{p+b-2}; \lambda_1, ..., \lambda_{p+b-3}, \lambda'_{p+b-2}\right), \quad (15)$$

$$(k = 1, 2)$$

where the shape parameters $r_1^*, \ldots, r_{p+b-3}^*$ are given by (11) through (13), the rate parameters $\lambda_1, \ldots \lambda_{p+b-3}$ by (10), r'_{p+b-2} is set equal to r_{p+b-2}, and θ, r_{p+b-2}, λ_{p+b-2} and λ'_{p+b-2} are obtained through the numerical solution of the system of equations

$$
\begin{cases}
\mu'_1 = \theta \dfrac{\Gamma\left(r_{p+b-2}+1\right)}{\Gamma\left(r_{p+b-2}\right)} \dfrac{1}{\lambda_{p+b-2}} + (1-\theta) \dfrac{\Gamma\left(r_{p+b-2}+1\right)}{\Gamma\left(r_{p+b-2}\right)} \dfrac{1}{\lambda'_{p+b-2}} \\[2ex]
\mu'_2 = \theta \dfrac{\Gamma\left(r_{p+b-2}+2\right)}{\Gamma\left(r_{p+b-2}\right)} \dfrac{1}{\lambda^2_{p+b-2}} + (1-\theta) \dfrac{\Gamma\left(r_{p+b-2}+2\right)}{\Gamma\left(r_{p+b-2}\right)} \dfrac{1}{\lambda'^2_{p+b-2}} \\[2ex]
\mu'_3 = \theta \dfrac{\Gamma\left(r_{p+b-2}+3\right)}{\Gamma\left(r_{p+b-2}\right)} \dfrac{1}{\lambda^3_{p+b-2}} + (1-\theta) \dfrac{\Gamma\left(r_{p+b-2}+3\right)}{\Gamma\left(r_{p+b-2}\right)} \dfrac{1}{\lambda'^3_{p+b-2}} \\[2ex]
\mu'_4 = \theta \dfrac{\Gamma\left(r_{p+b-2}+4\right)}{\Gamma\left(r_{p+b-2}\right)} \dfrac{1}{\lambda^4_{p+b-2}} + (1-\theta) \dfrac{\Gamma\left(r_{p+b-2}+4\right)}{\Gamma\left(r_{p+b-2}\right)} \dfrac{1}{\lambda'^4_{p+b-2}}
\end{cases}
\tag{16}
$$

where on the left hand side we have now the first four moments of a Logbeta r.v. with parameters $a_1 + \frac{b}{2} - \frac{3}{2}$ and $\frac{3}{2}$, obtained from (6) by replacing α and β by the appropriate values, and on the right hand side the first four moments of a mixture with weights θ $(0 < \theta < 1)$ and $1 - \theta$, of two Gamma r.v.'s, the first one with shape parameter r_{p+b-2} and rate parameter λ_{p+b-2} and the second one with shape parameter r_{p+b-2} and rate parameter λ'_{p+b-2}.

Proof: From (3), the h-th moment of Y_j is given by

$$
E\left(Y_j^h\right) = \frac{\Gamma\left(a_j + \frac{b}{2}\right)}{\Gamma\left(a_j\right)} \frac{\Gamma\left(a_j + h\right)}{\Gamma\left(a_j + \frac{b}{2} + h\right)},
$$

so that, given the independence of the p r.v.'s Y_j, we have

$$
E\left(W_1'^h\right) = E\left(\prod_{j=1}^p Y_j^h\right) = \prod_{j=1}^p E\left(Y_j^h\right) = \prod_{j=1}^p \frac{\Gamma\left(a_j + \frac{b}{2}\right)}{\Gamma\left(a_j\right)} \frac{\Gamma\left(a_j + h\right)}{\Gamma\left(a_j + \frac{b}{2} + h\right)}.
$$

The c.f. of W_1 is thus given by

$$
\begin{aligned}
\phi_{W_1}(t) &= E\left(e^{itW_1}\right) = E\left(e^{-it \log W_1'}\right) \\[2ex]
&= E\left(W_1'^{-it}\right) = \prod_{j=1}^p \frac{\Gamma\left(a_j + \frac{b}{2}\right)}{\Gamma\left(a_j\right)} \frac{\Gamma\left(a_j - it\right)}{\Gamma\left(a_j + \frac{b}{2} - it\right)},
\end{aligned}
\tag{17}
$$

where $\mathrm{i} = (-1)^{1/2}$ and $t \in \mathbb{R}$. But then, splitting the product and taking $a_{j+1} = c + \frac{p}{2} - \frac{j}{2} - \frac{1}{2} = a_j - \frac{1}{2}$ we may write the c.f. of W_1 as

$$
\phi_{W_1}(t) = \frac{\Gamma\left(a_1 + \frac{b}{2}\right)}{\Gamma(a_1)} \frac{\Gamma(a_1 - \mathrm{i}t)}{\Gamma\left(a_1 + \frac{b}{2} - \mathrm{i}t\right)} \prod_{j=2}^{p} \frac{\Gamma\left(a_j + \frac{b}{2}\right)}{\Gamma(a_j)} \frac{\Gamma(a_j - \mathrm{i}t)}{\Gamma\left(a_j + \frac{b}{2} - \mathrm{i}t\right)}
$$

$$
= \frac{\Gamma\left(a_1 + \frac{b}{2}\right)}{\Gamma\left(a_1 + \frac{b}{2} - \frac{3}{2}\right)} \frac{\Gamma\left(a_1 + \frac{b}{2} - \frac{3}{2}\right)}{\Gamma(a_1)} \frac{\Gamma\left(a_1 + \frac{b}{2} - \frac{3}{2} - \mathrm{i}t\right)}{\Gamma\left(a_1 + \frac{b}{2} - \mathrm{i}t\right)} \frac{\Gamma(a_1 - \mathrm{i}t)}{\Gamma\left(a_1 + \frac{b}{2} - \frac{3}{2} - \mathrm{i}t\right)}
$$

$$
\prod_{j=1}^{p-1} \frac{\Gamma\left(a_{j+1} + \frac{b}{2}\right)}{\Gamma(a_{j+1})} \frac{\Gamma(a_{j+1} - \mathrm{i}t)}{\Gamma\left(a_{j+1} + \frac{b}{2} - \mathrm{i}t\right)}
$$

$$
= \frac{\Gamma\left(a_1 + \frac{b}{2}\right)}{\Gamma\left(a_1 + \frac{b-3}{2}\right)} \frac{\Gamma\left(a_1 + \frac{b-3}{2}\right)}{\Gamma(a_1)} \frac{\Gamma\left(a_1 + \frac{b-3}{2} - \mathrm{i}t\right)}{\Gamma\left(a_1 + \frac{b}{2} - \mathrm{i}t\right)} \frac{\Gamma(a_1 - \mathrm{i}t)}{\Gamma\left(a_1 + \frac{b-3}{2} - \mathrm{i}t\right)}
$$

$$
\prod_{j=1}^{p-1} \frac{\Gamma\left(a_j + \frac{b-1}{2}\right)}{\Gamma\left(a_j - \frac{1}{2}\right)} \frac{\Gamma\left(a_j - \frac{1}{2} - \mathrm{i}t\right)}{\Gamma\left(a_j + \frac{b-1}{2} - \mathrm{i}t\right)}.
$$

But then, since b is an odd integer and thus $\frac{b-3}{2}$ is an integer, we may use, for any real or complex a, the identity

$$
\frac{\Gamma(a + \frac{b-3}{2})}{\Gamma(a)} = \prod_{j=0}^{\frac{b-3}{2} - 1} (a + j),
$$

to write the c.f. of W_1 as

$$
\phi_{W_1}(t) = \frac{\Gamma\left(a_1 + \frac{b}{2}\right)}{\Gamma\left(a_1 + \frac{b-3}{2}\right)} \frac{\Gamma\left(a_1 + \frac{b-3}{2} - \mathrm{i}t\right)}{\Gamma\left(a_1 + \frac{b}{2} - \mathrm{i}t\right)} \prod_{j=0}^{\frac{b-3}{2} - 1} (a_1 + j) \prod_{j=0}^{\frac{b-3}{2} - 1} (a_1 - \mathrm{i}t + j)^{-1}
$$

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

$$
= \frac{\Gamma\left(a_1 + \frac{b}{2}\right)}{\Gamma\left(a_1 + \frac{b-3}{2}\right)} \frac{\Gamma\left(a_1 + \frac{b-3}{2} - \mathrm{i}t\right)}{\Gamma\left(a_1 + \frac{b}{2} - \mathrm{i}t\right)} \prod_{j=0}^{\frac{b-5}{2}} (a_1 + j)(a_1 - \mathrm{i}t + j)^{-1}
$$

$$
\prod_{j=1}^{p-1} \frac{\Gamma\left(a_j + \frac{b-1}{2}\right)}{\Gamma\left(a_j - \frac{1}{2}\right)} \frac{\Gamma\left(a_j - \frac{1}{2} - \mathrm{i}t\right)}{\Gamma\left(a_j + \frac{b-1}{2} - \mathrm{i}t\right)}.
$$

9

Finally, using, the identity (Coelho, 2004)

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

valid for odd p, with r_j $(j = 1, \ldots, p + b - 3)$ given by (12) and (13), we have

$$
\begin{aligned}
\phi_{W_1}(t) &= \frac{\Gamma\left(c + \frac{p-1}{2} + \frac{b}{2}\right)}{\Gamma\left(c + \frac{p-1}{2} + \frac{b-3}{2}\right)} \frac{\Gamma\left(c + \frac{p-1}{2} + \frac{b-3}{2} - \mathrm{i}t\right)}{\Gamma\left(c + \frac{p-1}{2} + \frac{b}{2} - \mathrm{i}t\right)} \\
&\quad \prod_{j=0}^{\frac{b-5}{2}} \left(c + \frac{p}{2} - \frac{1}{2} + j\right) \left(c + \frac{p}{2} - \frac{1}{2} + j - \mathrm{i}t\right)^{-1} \\
&\quad \prod_{j=1}^{p+b-3} \left(c + \frac{j}{2} - \frac{1}{2}\right)^{r_j} \left(c + \frac{j}{2} - \frac{1}{2} - \mathrm{i}t\right)^{-r_j} \\
&= \frac{\Gamma\left(c + \frac{p-1}{2} + \frac{b}{2}\right)}{\Gamma\left(c + \frac{p-1}{2} + \frac{b-3}{2}\right)} \frac{\Gamma\left(c + \frac{p-1}{2} + \frac{b-3}{2} - \mathrm{i}t\right)}{\Gamma\left(c + \frac{p-1}{2} + \frac{b}{2} - \mathrm{i}t\right)} \\
&\quad \prod_{\substack{j=p \\ \text{step } 2}}^{p+b-5} \left(c + \frac{j}{2} - \frac{1}{2}\right) \left(c + \frac{j}{2} - \frac{1}{2} - \mathrm{i}t\right)^{-1} \\
&\quad \prod_{j=1}^{p+b-3} \left(c + \frac{j}{2} - \frac{1}{2}\right)^{r_j} \left(c + \frac{j}{2} - \frac{1}{2} - \mathrm{i}t\right)^{-r_j} \\
&= \frac{\Gamma\left(c + \frac{p-1}{2} + \frac{b}{2}\right)}{\Gamma\left(c + \frac{p-1}{2} + \frac{b-3}{2}\right)} \frac{\Gamma\left(c + \frac{p-1}{2} + \frac{b-3}{2} - \mathrm{i}t\right)}{\Gamma\left(c + \frac{p-1}{2} + \frac{b}{2} - \mathrm{i}t\right)} \\
&\quad \prod_{j=1}^{p+b-3} \left(c + \frac{j}{2} - \frac{1}{2}\right)^{r_j^*} \left(c + \frac{j}{2} - \frac{1}{2} - \mathrm{i}t\right)^{-r_j^*}, \qquad (18)
\end{aligned}
$$

where r_j^* $(j = 1, \ldots, p + b - 3)$ is given by (11) through (13).

Now, we keep unchanged the last product in (18) and we replace

$$\frac{\Gamma\left(c + \frac{p-1}{2} + \frac{b}{2}\right)}{\Gamma\left(c + \frac{p-1}{2} + \frac{b-3}{2}\right)} \frac{\Gamma\left(c + \frac{p-1}{2} + \frac{b-3}{2} - \mathrm{i}t\right)}{\Gamma\left(c + \frac{p-1}{2} + \frac{b}{2} - \mathrm{i}t\right)}, \qquad (19)$$

that is the c.f. of a Logbeta r.v. with parameters $c + \frac{p-1}{2} + \frac{b-3}{2}$ and $\frac{3}{2}$, by

$$\lambda_{p+b-2} \left(\lambda_{p+b-2} - \mathrm{i}t\right)^{-1} \lambda_{p+b-1}^{r_{p+b-1}^*} \left(\lambda_{p+b-1} - \mathrm{i}t\right)^{-r_{p+b-1}^*}, \qquad (20)$$

that is the c.f. of the sum of two Gamma r.v.'s, one with shape parameter 1 and rate parameter λ_{p+b-2} and the other with shape parameter r^*_{p+b-1} and rate parameter λ_{p+b-1}. The replacement is made in such a way that the c.f.$^{\rm s}$ in (19) and (20) have the first three derivatives with respect to t at $t=0$ equal. This means that the distributions to which they correspond will have the same first three moments. This leads us to obtain r^*_{p+b-1}, λ_{p+b-1} and λ_{p+b-2} as the solutions of the system of equations (14).

This way a near-exact c.f. of W_1 is obtained under the form

$$\phi_{W_1}(t) \simeq \lambda_{p+b-2} \left(\lambda_{p+b-2} - \mathrm{i}t\right)^{-1} \lambda_{p+b-1}^{r^*_{p+b-1}} \left(\lambda_{p+b-1} - \mathrm{i}t\right)^{-r^*_{p+b-1}}$$

$$\prod_{j=1}^{p+b-3} \left(c + \frac{j}{2} - \frac{1}{2}\right)^{r^*_j} \left(c + \frac{j}{2} - \frac{1}{2} - \mathrm{i}t\right)^{-r^*_j}, \tag{21}$$

that is the c.f. of the GNIG distribution of depth $p+b-1$ in (9), the sum of a GIG distribution of depth $p+b-3$ with a Gamma r.v. with shape parameter equal to 1 and rate parameter λ_{p+b-2} and another Gamma r.v. with shape parameter $r^*_{p+b-1} \in I\!\!R^+ \backslash I\!\!N$ and rate parameter λ_{p+b-1}. This turns out to be the sum of a GIG distribution of depth $p+b-2$ with the latter Gamma distribution. If by chance $r^*_{p+b-1} \in I\!\!N$ then the near-exact approximation to the distribution of W_1 we just obtained turns into a GIG distribution of depth $p+b-1$. In all the above we supposed that λ_{p+b-2} and λ_{p+b-1} were different from each other and different from all the other λ_j ($j = 1, \ldots, p+b-3$). We may note that if by chance it happens that any of λ_{p+b-2} or λ_{p+b-1} is equal to any of the λ_j ($j = 1, \ldots, p+b-3$) or equal to each other this will only reduce the depth of the GNIG distribution by one. In any case, this near-exact approximation will have the same first three moments as the exact distribution.

In fact, depending on the asymptotic result used in (18), we may obtain different near-exact approximations to the distribution of W_1. This way, we decided now to replace the part of the c.f. of W_1 in (19) by

$$\theta \frac{\lambda_{p+b-2}^{r_{p+b-2}}}{\left(\lambda_{p+b-2} - \mathrm{i}t\right)^{r_{p+b-2}}} + (1-\theta) \frac{\lambda_{p+b-2}'^{r'_{p+b-2}}}{\left(\lambda_{p+b-2}' - \mathrm{i}t\right)^{r'_{p+b-2}}}, \tag{22}$$

that is the c.f. of the mixture of two Gamma r.v.'s with shape parameters r_{p+b-2} and r'_{p+b-2} and rate parameters λ_{p+b-2} and λ'_{p+b-2}. The approximation is done in such a way that the c.f.'s in (19) and (22) have the same first four derivatives with respect to t at $t=0$, that is, the corresponding distributions have the same first four moments. We set $r'_{p+b-2} = r_{p+b-2}$ and then obtain the

11

four parameters θ, r_{p+b-2}, λ_{p+b-2} and λ'_{p+b-2} through the numerical solution of the system of equations (16).

In this case we have

$$\phi_{W_1}(t) \simeq \left[\theta \frac{\lambda_{p+b-2}^{r_{p+b-2}}}{(\lambda_{p+b-2} - \mathrm{i}t)^{r_{p+b-2}}} + (1-\theta) \frac{\lambda_{p+b-2}'^{r_{p+b-2}}}{\left(\lambda'_{p+b-2} - \mathrm{i}t\right)^{r_{p+b-2}}} \right]$$
$$\prod_{j=1}^{p+b-3} \left(c + \frac{j}{2} - \frac{1}{2}\right)^{r_j^*} \left(c + \frac{j}{2} - \frac{1}{2} - \mathrm{i}t\right)^{-r_j^*}, \qquad (23)$$

that is the c.f. of the mixture of two GNIG distributions of depth $p+b-2$ in (15), with $r'_{p+b-2} = r_{p+b-2}$. In the above we supposed that λ_{p+b-2} and λ'_{p+b-2} were different from all the other λ_j $(j = 1, \ldots, p+b-3)$. If, by chance, this does not happen, this will only reduce the depth of the corresponding GNIG distribution by one. In any case, this near-exact approximation will have the same first four moments as the exact distribution.

That the c.f. of W_2 and W_1 are the same (that is, that p and b are interchangeable), may be seen through the use of the equality

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

which is valid for any integer p and b (Coelho, 1998, 1999), by writing the c.f. of W_2 as

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

Since $\phi_{W_1}(t) = \phi_{W_2}(t)$, W_1 and W_2 have the same distribution and thus the same near-exact c.f.$^{\mathrm{s}}$ and distributions. The near-exact distributions for W_1' and W_2' may then be obtained through the transformation $W_k' = \mathrm{e}^{-W_k}$ $(k=1,2)$. $\square$

The near-exact p.d.f. and c.d.f. of W_1, for the first approximation in Theorem 1 above, corresponding to the c.f. in (21), are given by (1) and (2) respectively, with $Z = W_1$, $g = p + b - 2$, $r_j = r_j^*$ $(j = 1, \ldots, p + b - 3)$, $r_g = r_{p+b-2}^* = 1$, $r = r_{p+b-1}^*$, $\lambda_j = c - \frac{1}{2} + \frac{j}{2}$ $(j = 1, \ldots, p + b - 3)$, $\lambda_g = \lambda_{p+b-2}$ and $\lambda = \lambda_{p+b-1}$, that is, this near-exact p.d.f. of W_1 is given by

$$f(w_1) = K\lambda_{p+b-1}^{r_{p+b-1}^*} \sum_{j=1}^{p+b-2} e^{-\lambda_j w_1} \sum_{k=1}^{r_j^*} \left\{ c_{j,k} \frac{\Gamma(k)}{\Gamma(k+r_{p+b-1}^*)} w_1^{k+r_{p+b-1}^*-1} \right.$$
$$\left. {}_1F_1(r_{p+b-1}^*, k+r_{p+b-1}^*, -(\lambda_{p+b-1}-\lambda_j)w_1) \right\},$$
$$(w_1 > 0)$$

and the near-exact c.d.f., by

$$F(w_1) = \lambda_{p+b-1}^{r_{p+b-1}^*} \frac{w_1^{r_{p+b-1}^*}}{\Gamma(r_{p+b-1}^* + 1)} {}_1F_1(r_{p+b-1}^*, r_{p+b-1}^* + 1, -\lambda_{p+b-1}w_1)$$
$$- K\lambda_{p+b-1}^{r_{p+b-1}^*} \sum_{j=1}^{p+b-2} e^{-\lambda_j w_1} \sum_{k=1}^{r_j^*} c_{j,k}^* \sum_{i=0}^{k-1} \frac{w_1^{r_{p+b-1}^*+i} \lambda_j^i}{\Gamma(r_{p+b-1}^* + 1 + i)}$$
$$_1F_1(r_{p+b-1}^*, r_{p+b-1}^* + 1 + i, -(\lambda_{p+b-1} - \lambda_j)w_1)$$
$$(w_1 > 0)$$

with

$$K = \prod_{j=1}^{p+b-2} \lambda_j^{r_j}.$$

The near-exact p.d.f. and c.d.f. of W_1, for the second approximation in Theorem 1 above, corresponding to the c.f. in (23), are given by

$$f(w_1) = \theta K\lambda_{p+b-2}^{r_{p+b-2}} \sum_{j=1}^{p+b-3} e^{-\lambda_j w_1} \sum_{k=1}^{r_j^*} \left\{ c_{j,k} \frac{\Gamma(k)}{\Gamma(k+r_{p+b-2})} w_1^{k+r_{p+b-2}-1} \right.$$
$$\left. {}_1F_1(r_{p+b-2}, k+r_{p+b-2}, -(\lambda_{p+b-2}-\lambda_j)w_1) \right\}$$
$$+(1-\theta) K\lambda_{p+b-2}'^{r_{p+b-2}} \sum_{j=1}^{p+b-3} e^{-\lambda_j w_1} \sum_{k=1}^{r_j^*} \left\{ c_{j,k} \frac{\Gamma(k)}{\Gamma(k+r_{p+b-2})} w_1^{k+r_{p+b-2}-1} \right.$$
$$\left. {}_1F_1(r_{p+b-2}, k+r_{p+b-2}, -(\lambda_{p+b-2}'-\lambda_j)w_1) \right\}$$
$$(w_1 > 0)$$

13

and

$$F(w_1) = \theta \left\{ \lambda_{p+b-2}^{r_{p+b-2}} \frac{w_1^{r_{p+b-2}}}{\Gamma(r_{p+b-2}+1)} {}_1F_1(r_{p+b-2}, r_{p+b-2}+1, -\lambda_{p+b-2}w_1) \right.$$

$$- K\lambda_{p+b-2}^{r_{p+b-2}} \sum_{j=1}^{p+b-3} e^{-\lambda_j w_1} \sum_{k=1}^{r_j^*} c_{j,k}^* \sum_{i=0}^{k-1} \frac{w_1^{r_{p+b-2}+i} \lambda_j^i}{\Gamma(r_{p+b-2}+1+i)}$$

$$\left. {}_1F_1(r_{p+b-2}, r_{p+b-2}+1+i, -(\lambda_{p+b-2} - \lambda_j)w_1) \right\}$$

$$+(1-\theta) \left\{ \lambda_{p+b-2}'^{r_{p+b-2}} \frac{w_1^{r_{p+b-2}}}{\Gamma(r_{p+b-2}+1)} {}_1F_1(r_{p+b-2}, r_{p+b-2}+1, -\lambda_{p+b-2}'w_1) \right.$$

$$- K\lambda_{p+b-2}'^{r_{p+b-2}} \sum_{j=1}^{p+b-3} e^{-\lambda_j w_1} \sum_{k=1}^{r_j^*} c_{j,k}^* \sum_{i=0}^{k-1} \frac{w_1^{r_{p+b-2}+i} \lambda_j^i}{\Gamma(r_{p+b-2}+1+i)}$$

$$\left. {}_1F_1(r_{p+b-2}, r_{p+b-2}+1+i, -(\lambda_{p+b-2}' - \lambda_j)w_1) \right\}$$

$$(w_1 > 0)$$

with

$$K = \prod_{j=1}^{p+b-3} \lambda_j^{r_j} \, .$$

The corresponding near-exact p.d.f.'s and c.d.f.'s of W_1' may then be obtained through the transformation $W_1' = e^{-W_1}$.

Although, at first sight, these expressions may look a bit complex, they are actually much manageable. In fact, from these c.d.f.'s near-exact quantiles may be easily computed.

4 Evaluation of the quality and comparison of the two near-exact approximations

It is now important to analyze how good the proposed approximations are. Let us start with the study of the proximity between a $Logbeta\left(c + \frac{p-1}{2} + \frac{b-3}{2}, \frac{3}{2}\right)$ distribution and a $GNIG\left(1, r_{p+b-1}^*; \lambda_{p+b-2}, \lambda_{p+b-1}\right)$ distribution of depth 2, which correspond to the c.f.'s in (19) and (20), respectively. The parameters r_{p+b-1}^*, λ_{p+b-2}, λ_{p+b-1} are obtained as the numeric solution of the system of equations (14) where μ_1', μ_2' and μ_3' are given by (6).

If we take, for example, $\alpha = c + \frac{p-1}{2} + \frac{b-3}{2} = \frac{152}{10}$, where p and b are both

positive odd integers, with $b \geq 3$, we will obtain for r^*_{p+b-1}, λ_{p+b-2}, λ_{p+b-1} the values in Table 1.

Table 1
Values for shape and rate parameters for the GNIG distribution

$r^*_{p+b-2}=1$	$\lambda_{p+b-2}=15.672643219241610$
$r^*_{p+b-1}=0.499989615106888$	$\lambda_{p+b-1}=15.003755582976448$

Then, in order to analyze the closeness between a $Logbeta\left(c + \frac{p-1}{2} + \frac{b-3}{2}, \frac{3}{2}\right)$ distribution and the mixture of two Gamma distributions, $MG\left(\theta; r_{p+b-2}, \lambda_{p+b-2}; r_{p+b-2}, \lambda'_{p+b-1}\right)$, which correspond respectively to the c.f.$^{\mathrm{s}}$ in (19) and (22), we will take θ, r_{p+b-2}, λ_{p+b-2}, λ'_{p+b-2} in (22) as solutions of the system of equations (16) where μ'_1, μ'_2, μ'_3 and μ'_4 are given by (6) with $\alpha = 152/10$ and $\beta = 3/2$. We obtain for θ, r_{p+b-2}, λ_{p+b-2} and λ'_{p+b-2} the values in Table 2.

Table 2
Values for the weight, shape and rate parameters for the MG distribution

$r_{p+b-2}=1.500000114557350$	$\lambda_{p+b-2}=15.248908384617104$
$r'_{p+b-2} = r_{p+b-2}$	$\lambda'_{p+b-2}=15.657422531150101$
$\theta=0.517644326454732$	

We may note that the choice of the value $3/2$ for the second parameter of the above Logbeta distribution, made in a given moment of the proof of Theorem 1, was not random but rather a judicious one. This choice was based on the assumption that this value should be of the form $k/2$ where k is a positive odd integer. Besides, on one hand, for the sum of two Gamma distributions this value has to be greater than one, since the sum of the two shape parameters $r^*_{p+b-2} = 1$ and r^*_{p+b-1} in (20) will always be greater than one, while, on the other hand, the value $3/2$ is, for both approximations, the value of the form $k/2$, that leads to a closer match between the exact and approximate c.f.$^{\mathrm{s}}$.

The p.d.f.'s for the three distributions, that is, the Logbeta, GNIG and MG distributions, are indeed almost superimposed, as we may see by analyzing Tables 3 and 4. In Table 3 we consider the ordinates of the Logbeta p.d.f. in (4), with parameters $\alpha = 152/10$ and $\beta = 3/2$ and the ordinates of the p.d.f. of a GNIG distribution of depth 2 with parameters given in Table 1 and in Table 4 the ordinates of the Logbeta p.d.f. and the ordinates of the p.d.f. of an MG distribution with parameters given in Table 2. We may notice the close match between the three p.d.f.'s, mainly between the p.d.f.'s for the Logbeta and the MG distributions.

Table 3
Ordinates for the Logbeta and GNIG p.d.f.'s and absolute
value of the differences

x-value	Logbeta Ordinates ($\alpha=152/10$, $\beta=3/2$)	GNIG Ordinates	Absolute value of difference
(0.00)	(0.00)	(0.00)	(0.00)
0.05	7.074907762548123	7.074891356005670	0.000016406542453
0.10	4.621799957250436	4.621799638464716	0.000000318785720
0.15	2.615034081669441	2.615038603240202	0.000004521570762
0.20	1.395122967860079	1.395125739783440	0.000002771923362
0.25	0.720739641900814	0.720740250195227	0.000000608294413
0.30	0.364858983899873	0.364858532419835	0.000000451480037
0.35	0.182137548158239	0.182136880665808	0.000000667492431
0.40	0.089999784336385	0.089999264351411	0.000000519984975
0.45	0.044127428272311	0.044127124974587	0.000000303297724
0.50	0.021504187316464	0.021504053584123	0.000000133732342

Table 4
Ordinates for the Logbeta and MG p.d.f.'s and absolute
value of the differences

x-value	Logbeta Ordinates ($\alpha=152/10$, $\beta=3/2$)	MG Ordinates	Absolute value of difference
(0.00)	(0.00)	(0.00)	(0.00)
0.05	7.074907762548123	7.074907891847342	0.000000129299219
0.10	4.621799957250436	4.621799915035653	0.000000042214784
0.15	2.615034081669441	2.615034040661706	0.000000041007735
0.20	1.395122967860079	1.395122962209558	0.000000005650520
0.25	0.720739641900814	0.720739652020356	0.000000010119543
0.30	0.364858983899873	0.364858994042061	0.000000010142219
0.35	0.182137548158239	0.182137553520275	0.000000005362036
0.40	0.089999784336385	0.089999785671716	0.000000001335331
0.45	0.044127428272311	0.044127427554592	0.000000000717719
0.50	0.021504187316464	0.021504186022274	0.000000001294190

In Table 5 we have the moments for these three distributions. As we can
see the first three moments of the GNIG distribution match the first three
moments of the Logbeta distribution and the first four moments of the MG
distribution match the first four moments of the Logbeta, as it was really
meant by construction of the two approximations. The other moments are all
very close to the exact Logbeta ones, once again with the MG distribution
showing an even better approximation.

Table 5
Comparison of moments of the three distributions

Moment order	Logbeta distribution ($\alpha=152/10$, $\beta=3/2$)	GNIG distribution	MG distribution
1	0.097129744129098	0.097129744129098	0.097129744129098
2	0.015726385953047	0.015726385953047	0.015726385953047
3	0.003565408035466	0.003565408035466	0.003565408035466
4	0.001039464045853	0.001039465187900	0.001039464045853
5	0.000370454752724	0.000370456502839	0.000370454755770
6	0.000156058213794	0.000156060157529	0.000156058220471
7	0.000075868390181	0.000075870359673	0.000075868400193
8	0.000041808769278	0.000041810732914	0.000041808782411
9	0.000025754487269	0.000025756477006	0.000025754503685
10	0.000017537922433	0.000017540003334	0.000017537942764

In Table 6 we may analyze the .90, .95 and .99 quantiles for the three distributions. Once again the mixture of two Gamma distributions shows a better approximation, although both distributions yield remarkably good approximations.

Table 6
Comparison of quantiles of the three distributions

Quantile	Logbeta distribution ($\alpha=152/10$, $\beta=3/2$)	GNIG distribution	MG distribution
.90	0.202409649747600	0.202409639248248	0.202409650573903
.95	0.253045234566776	0.253045104347747	0.253045235827326
.99	0.367409541282429	0.367409219232922	0.367409540486442

We should note that both approximations improve for larger values of $\alpha = c + \frac{p-1}{2} + \frac{b-3}{2}$ being the approximation given by the mixture of two Gamma distributions always the best of the two.

Finally, we compare the near-exact moments for W_1 in (8) given by the two near-exact distributions whose c.f.'s are given by (21) and (23) with the exact ones, obtained from the c.f. in (17). In this case the exact distribution of W_1 is the sum of p independent Logbeta r.v.'s.

We considered three possible different situations that yield $\alpha = c + \frac{p-1}{2} + \frac{b-3}{2} = 152/10$. For all three situations we take $c = 2/10$. In Table 7 we have the moments for $p = 33$ and $b = 1$, what would yield $a_j = \frac{2}{10} + \frac{33}{2} - \frac{j}{2}$ ($j = 1, \ldots, 33$) in (7) (or equivalently, $p = 1$ and $b = 33$, what would yield just one Logbeta distribution, with $a_1 = c$), in Table 8, the moments for $p = 31$ and $b = 3$, what would yield $a_j = \frac{2}{10} + \frac{31}{2} - \frac{j}{2}$ ($j = 1, \ldots, 31$) in (7), and in Table 9 for $p = 17$ and $b = 17$, yielding $a_j = \frac{2}{10} + \frac{17}{2} - \frac{j}{2}$ ($j = 1, \ldots, 17$).

For all three cases it is possible to confirm the equality of the first three moments for the overall GNIG near-exact distribution and of the first four for the overall mixture of two GNIG distributions (MGNIG) and the very close proximity of all the other moments, of course with some relatively small degradation as the order of the moments progress.

Table 7

Comparison of moments of order h for the sum of Logbetas (exact) and the two near-exact distributions: the GNIG in (21) and the MGNIG in (23) for $p = 33$ and $b = 1$, or $p = 1$ and $b = 33$, with $c = 2/10$

h	Sum of Logbetas (exact)	GNIG (near-exact)	MGNIG (near-exact)
1	8.074209799622	8.074209799622	8.074209799622
2	91.398532273409	91.398532273409	91.398532273409
3	1412.625283255628	1412.625283255628	1412.625283255628
4	28435.923374222438	28435.923374223580	28435.923374222438
5	711807.170188010914	711807.170188058215	711807.170188010917
6	21359273.623424165535	21359273.623425790148	21359273.623424165687
7	747606062.549602711864	747606062.549661527388	747606062.549602717994
8	29904460682.860704833626	29904460682.863075216977	29904460682.860705087549
9	1345702404593.773404360869	1345702404593.880233787601	1345702404593.773415877302
10	67285134362214.402856671253	67285134362219.745856866369	67285134362214.403433411033

Table 8

Comparison of moments of order h for the sum of Logbetas (exact) and the
two near-exact distributions: the GNIG in (21) and the MGNIG in (23)
for $p = 31$ and $b = 3$, with $c = 2/10$

h	Sum of Logbetas (exact)	GNIG (near-exact)	MGNIG (near-exact)
1	15.058578518011	15.058578518011	15.058578518011
2	256.938443186038	256.938443186038	256.938443186038
3	5037.371813747536	5037.371813747536	5037.371813747536
4	114617.058355589198	114617.058355590340	114617.058355589198
5	3036894.158503457579	3036894.158503544763	3036894.158503457582
6	93345423.514353025083	93345423.514357535609	93345423.514353025363
7	3297990460.599706707560	3297990460.599914613797	3297990460.599706724521
8	132371103233.040425680975	132371103233.049932904859	132371103233.040426576351
9	5963688861840.865009412834	5963688861841.319697681349	5963688861840.865055538297
10	298299167638986.629019092376	298299167639009.954931173233	298299167638986.631472117996

Table 9

Comparison of moments of order h for the sum of Logbetas (exact) and the
two near-exact distributions: the GNIG in (21) and the MGNIG in (23)
for $p = 17$ and $b = 17$, with $c = 2/10$

h	Sum of Logbetas (exact)	GNIG (near-exact)	MGNIG (near-exact)
1	29.820927206388	29.820927206388	29.820927206388
2	922.355817517058	922.355817517058	922.355817517058
3	29738.684568847095	29738.684568847095	29738.684568847095
4	1005464.874469326057	1005464.874469327199	1005464.874469326057
5	35888096.386680082775	35888096.386680254255	35888096.386680082778
6	1362045465.297717201464	1362045465.297733216942	1362045465.297717202014
7	55355953130.110260320272	55355953130.111472384574	55355953130.110260380312
8	2424157852793.904581954260	2424157852793.986978435253	2424157852793.904587155921
9	114907378069157.741159672017	114907378069163.059301551761	114907378069157.741558075982
10	5909217530677925.017650505511	5909217530678262.697578488520	5909217530677925.046258781012

However, even for the higher moments the error percentage remains at extremely low levels, of no more than 8×10^{-14} percent for the 10th moment, when computed taking the exact moments as a basis, as it may be seen in Table 10. Another remarkable feature is the fact that the error percentages, for a given near-exact distribution, remain quite stable for any combination of p and b values.

Table 10

Error percentages for the 10th moment concerning the two
near-exact distributions: the GNIG in (21) and the
MGNIG in (23) taking as a basis the exact moment

	GNIG (near-exact)	MGNIG (near-exact)
$p = 33,\ b = 1$	-7.941×10^{-14}	-8.572×10^{-18}
$p = 31,\ b = 3$	-7.820×10^{-14}	-8.223×10^{-18}
$p = 17,\ b = 17$	-5.714×10^{-14}	-4.841×10^{-18}

Although we have the analytic expressions for the exact and near-exact c.f.'s of the logarithm of product of p independent Beta r.v.'s, since some of the parameters in the near-exact approximations are computed by numerically solving systems of equations, it doesn't render possible the evaluation of analytic Berry-Esseen type of bounds (Berry, 1941; Esseen, 1945; Loève, 1977; Hwang, 1998) for the differences between the exact and the approximate c.d.f.'s. However, for any given case, once the values for all the parameters of the Beta distributions in the product are specified, we may compute a couple of measures based on the difference between the exact c.f. of Y and the c.f. corresponding to the near-exact approximation under study. These measures will help us in evaluating and comparing the performance of the near-exact distributions proposed. One of these measures is even indeed based on the Berry-Esseen bound.

Let Y be a r.v. with support S and let $\Phi_Y(t)$ and $F_Y(y)$ be respectively the exact c.f. and c.d.f. of Y and let $\Phi(t)$ and $F(y)$ represent respectively the c.f. and c.d.f. corresponding to the approximation under study. Further, let $f(y)$ represent the p.d.f. corresponding to $F(y)$. One measure we may think of is

$$\Delta_1 = \int_{-\infty}^{\infty} |\Phi_Y(t) - \Phi(t)|\, dt$$

while the other, related with this one and based directly on the Berry-Esseen upper bound on $|F_Y(y) - F(y)|$ is

$$\Delta_2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \frac{\Phi_Y(t) - \Phi(t)}{t} \right| dt \, .$$

The Berry-Esseen inequality, which is indeed an upper bound on $|F_Y(y) - F(y)|$, may, for any $b > 1/(2\pi)$ and any $T > 0$, be written as

$$\max_{y \in S} |F_Y(y) - F(y)| \leq b \int_{-T}^{T} \left| \frac{\Phi_Y(t) - \Phi(t)}{t} \right| dt + C(b) \frac{M}{T} \tag{24}$$

where $M = max_{y \in S} f(y)$ and $C(b)$ is a positive constant that only depends on b. But if in (24) above we take $T \to \infty$ then we will have Δ_2 since then we may take $b = 1/(2\pi)$.

In Table 11 we may see the computed values of the two measures Δ_1 and Δ_2 for both near-exact distributions proposed. We considered the same three situations as above, that is, the sum of p Logbeta distributions for $\alpha = 152/10$, with $c = 2/10$, and i) $p = 33$, $b = 1$, ii) $p = 31$, $b = 3$ and iii) $p = 17$, $b = 17$. Once again we may see the better performance of the mixture of two GNIG distributions, with lower values for both measures.

Table 11

Values for the two measures Δ_1 and Δ_2 for the two near-exact distributions: the GNIG in (21) and the MGNIG in (23) for $\alpha = 152/10$, with $c = 2/10$, and several values of p and b

	GNIG		MGNIG	
	Δ_1	Δ_2	Δ_1	Δ_2
$p = 33,\ b = 1$	1.302×10^{-10}	7.887×10^{-12}	2.302×10^{-13}	1.078×10^{-14}
$p = 31,\ b = 3$	3.474×10^{-12}	5.524×10^{-13}	2.294×10^{-15}	2.940×10^{-16}
$p = 17,\ b = 17$	8.080×10^{-13}	1.941×10^{-13}	3.377×10^{-16}	6.853×10^{-17}

5   Conclusions and final remarks

The two near-exact approximations developed lay very close to the exact distribution in terms of c.f., moments, c.d.f. and quantiles and are highly adequate to be used in computing near-exact quantiles.

Although some of the expressions of the near-exact distributions, namely for the c.d.f.'s, may still seem a bit complicated they are not only far more manageable than the exact c.d.f. but also perfectly handled by a number of available softwares, rendering the computation of near-exact quantiles not too hard.

The two near-exact distributions proposed, that is, the single GNIG distribution and the mixture of two GNIG distributions, illustrate well the trade-off between manageability and precision. While the single GNIG near-exact distribution is much simpler both analytically and computationally, the mixture of two GNIG distributions yields a much better approximation.

The mixture of two GNIG distributions leads to the best results. This was somehow expected given that by construction this approximation has the same first four moments equal to the exact ones while the single GNIG distribution only has the first three.

The near-exact distributions presented and the methods of derivation used may be readily applied to obtain near-exact distributions of the generalized Wilks Λ statistic in the case where two or more sets of variables have an odd number of variables, as well as to obtain near-exact approximations to the distributions of other related statistics as the one used to test the equality of two multivariate Normal distributions.

6   Acknowledgments

References

Abramowitz, M. and Stegun, I. A., 1974. Handbook of Mathematical Functions, (Eds.), 9th print. Dover, New York.

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

Berry, A., 1941. The Accuracy of the Gaussian Approximation to the Sum of Independent Variates. Transactions of the American Mathematical Society 49, 122-136.

Coelho, C. A., 1998. The Generalized Integer Gamma Distribution – A Basis for Distributions in Multivariate Statistics. J. Multiv. Anal. 64, 86-102.

Coelho, C. A., 1999. Addendum to the paper "The Generalized Integer Gamma Distribution – A Basis for Distributions in Multivariate Statistics". J. Multiv. Anal. 69, 281-285.

Coelho, C. A., 2004. The Generalized Near-Integer Gamma Distribution. A basis for 'near-exact' approximations to the distributions of statistics which are the product of an odd number of independent Beta random variables. Journal of Multivariate Analysis 89, 191-218.

Esseen, C.-G., 1945. Fourier Analysis of Distribution Functions. A Mathematical Study of the Laplace-Gaussian Law. Acta Mathematica 77, 1-125.

Gupta, R. D. and Richards, D., 1979. Exact distributions of Wilks' Λ under null and non-null (linear) hypotheses. Statistica 39, 333-342.

Gupta, R. D. and Richards, D., 1983. Application of results of Kotz, Johnson and Boyd to the null distribution of Wilks' criterion. In: P. K. Sen (Ed.), Contributions to Statistics: Essays in honour of Norman L. Johnson, 205-210.

Hwang, H.-K., 1998. On convergence rates in the central limit theorems for combinatorial structures. European Journal of Combinatorics 19, 329-343.

Johnson, N. L., Kotz, S. and Balakrishnan, N., 1995. Continuous Univariate Distributions, Vol. 2, 2nd ed.. J. Wiley & Sons, New York.

Kabe, D. G., 1962. On the exact distribution of a class of multivariate test criteria. Ann. Math. Statist. 33, 1197-1200.

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

Nandi, S. B., 1977. The exact null distribution of Wilks' criterion. Sankhyã 39, 307-315.

Tretter, M. J. and Walster, G. W., 1975. Central and noncentral distributions of Wilks' statistic in Manova as mixtures of incomplete Beta functions. Ann. Statist. 3, 467-472.

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

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