The Ratio of Independent Generalized Gamma Random Variables with Applications

Vusi Bilankulu
University of Pretoria, South Africa
vusi.bilankulu@up.ac.za

Andriëtte Bekker
University of Pretoria, South Africa
andriette.bekker@up.ac.za

Filipe Marques
Centro de Matemática e Aplicações (CMA), FCT, UNL and
Departamento de Matemática, FCT, UNL
fjm@fct.unl.pt

April 23, 2018

Abstract

This paper originates from a statistic defined as the ratio of independent generalized gamma random variables and shows that it can be represented as the product of independent generalized gamma random variables with some re-parametrization. By decomposing the characteristic function of the negative natural logarithm of the statistic and by using the distribution of the difference of two independent generalized integer gamma random variables as a basis, accurate and computationally appealing near-exact distributions are derived for the statistic. In the process, a new parameter is introduced in the near-exact distributions which allows to control the degree of precision of these approximations. Furthermore, the performance of the near-exact distributions is assessed using a measure of proximity between cumulative distribution functions and also by comparison with the exact and empirical distributions. We illustrate the use of the proposed approximations on the distribution of the ratio of generalized variances in a multivariate multiple regression setting and with an example of application related with SISO Networks.

1 Introduction

The generalized gamma distribution was introduced by Stacy [1]. It is a generalization of well-known distributions such as gamma, chi-squared, exponential, Rayleigh, Weibull and Nakagami-m. Due to its adaptability, either in this generalized form or one of its special cases, the generalized gamma distribution has received much interest and wide applications in areas such as hydrological processes [2], wireless communication [3], reliability analysis, economics and life testing. In [4], the author unknowingly reintroduced the generalized gamma distribution as a general fading distribution, the so-called $\alpha - \mu$ distribution.

In many of the applications of the generalized gamma distribution, the product or ratio of independent generalized gamma random variables appears naturally. For instance, [2] considered both the product and ratio of independent generalized gamma random variables to model the magnitude of a drought and relative duration of a drought events respectively. In multi-hop wireless relaying systems, the end-to-end signal-to-noise ratio (SNR) and the rate offset can be modeled as a function of the product of independent generalized gamma, Rayleigh or Nakagami-m random variables (see [3]). Signal-to-interference ratio (SIR) can be modelled as the ratio of either independent generalized gamma, independent Rayleigh or independent Nakagami-m random variables (see [5]). In [6] the authors modeled the intensity of the Poisson process in the Poisson-gamma hierarchical generalized linear model as product of independent gamma random variables. The product and ratio of independent generalized gamma distributed random variables also appear fundamental in statistical theory, see [7] for further details.

The exact distribution and approximate distribution of the product of independent generalized gamma distributed random variables has been studied by a number of authors. In [8] the exact distribution was derived in terms of Fox's H-function which are still not computable even with today's powerful computer softwares. By considering special cases of generalized gamma distributed random variables, [9], [10] and [11] expressed the exact distribution in terms of the Meijer's G-function which is be computable using most computer software. However, even with today's powerful computers, the computation of Meijer's G still takes considerably long computational time (see Section 4 for an example of this feature). In [12], an approximation to the product of independent Rayleigh distributed random variables is proposed. Authors in [13] and [14] derived the exact distribution in terms of infinite series from which an approximate distribution can be obtained. However, to get the required accuracy a large number of terms need to be evaluated. In [6] the exact distributions, when all power parameters are of the same sign, is approximated. Another approximate distribution was introduced in [3], however, this approximate distribution was derived under restrictive conditions i.e. when the shape and the power parameters are fixed.

Few authors have attempted to study the distribution of the ratio of independent generalized gamma random variables. Many of those have limited their study to only two or three random variables with some restrictions of parameters (see [1], [2], [15] and [16]). Other authors (e.g. [10]) have derived the distribution in terms of Meijer's$-G$ function. As noted above, this approach has a computational time disadvantage. In [8], it is recognized that the generalized gamma distribution is a special case of the H-function distribution. Thus by studying the ratio of H-function distributed random variables, [8] implicitly studied the exact distribution of the ratio of generalized gamma distributed random variables. However, the results are in terms of the Fox's H-functions and therefore had no practical value.

In this article, it is shown that the ratio of independent generalized gamma random variables can also be represented as the product of independent generalized gamma random variables some with negative parameters and others with positive parameters. By decomposing the characteristic function of the negative natural logarithm of the product of these random variables, a different representation of the exact distribution is obtained. From this new presentation, two near-exact distributions are derived. In the process of decomposing the characteristic function, a new parameter which control the degree of precision for near-exact distribution is obtained.

The rest of the paper is organized as follows; in Section 2, it is shown that the ratio of independent generalized gamma distributed random variables may be represented as the product of independent generalized gamma random variables with some re-parametrization. The statistic of interest, Y, is then defined. Starting from the characteristic function of $Z = -\ln Y$, another representation of the exact distribution of Z is obtained and used to derive novel approximations i.e. near-exact distributions for Z and hence for Y. In Section 3, the performance of near-exact distributions is studied and contrasted against the performance of an empirical distribution. In Section 4, two examples of application are of proposed near-exact distributions are provide in the multivariate multiple linear regression setting and in a problem related with SISO Networks. Computational modules are presented in 5. The conclusion is included in Section 6.

2 The exact distribution and near-exact distributions

2.1 The exact distribution

Let X be a random variable with probability density function (pdf) given by

$$f_X(x; r, \lambda, \delta) = |\delta| \frac{\lambda^{\delta r} x^{\delta r - 1}}{\Gamma(r)} \exp\left(-(\lambda x)^{\delta}\right) \tag{1}$$

where $x \geq 0$, $r > 0$, $\lambda > 0$ and $\delta \neq 0$. Then X is said to follow a generalized gamma distribution and is denoted by $X \sim \mathcal{G}\Gamma(r, \lambda, \delta)$. The s^{th} moment of X is given by

$$E[X^s] = \frac{\Gamma\left(r + \frac{s}{\delta}\right)}{\Gamma(r)} \lambda^s. \tag{2}$$

where $E[\cdot]$ denotes mathematical expectation.

Remark 1 *If* $V = \frac{1}{X}$ *then* $V \sim \mathcal{G}\Gamma(r, \lambda^{-1}, -\delta)$.

Let $X_{1i} \sim \mathcal{G}\Gamma(r_{1i}, \lambda_{1i}, \delta_{1i})$ for $i = 1, 2, \ldots, n_1$ and $\delta_{1i} > 0$. Furthermore, let $X_{2k} \sim \mathcal{G}\Gamma(r_{2k}, \lambda_{2k}, \delta_{2k})$ for $k = 1, 2, \ldots, n_2$ and $\delta_{2k} > 0$. Define the ratio of independent generalized gamma random variables, for $n_1, n_2 > 0$, as

$$G = \left(\prod_{i=1}^{n_1} X_{1i}\right)\left(\prod_{k=1}^{n_2} X_{2k}\right)^{-1}. \tag{3}$$

Following Remark 1, (3) can be conveniently rewritten as

$$G = \prod_{i=1}^{n_1} X_{1i} \prod_{k=1}^{n_2} V_{2k}, \tag{4}$$

where $V_{2k} \sim \mathcal{G}\Gamma\left(r_{2k}, \lambda_{2k}^{-1}, -\delta_{2k}\right)$. Therefore, G can be represented as a product of independent generalized gamma random variables some with negative power parameters and others with positive power parameters (see (4)). The statistic of interest generalizing G, is now defined. For $X_i \sim \mathcal{G}\Gamma\left(r_i, \lambda_i, \delta_i\right)$, $i = 1, 2, \ldots, n$, such that $\delta_i < 0$ and $\delta_k > 0$ for some $i, k \in \{1, 2, 3, \ldots, n\}$, we define

$$Y = \prod_{i=1}^{n} X_i. \tag{5}$$

By considering special cases, [10] expressed the exact density probability function of Y in terms of the Meijer's$-G$ function. Evidence of the computational disadvantage of this approach will be provided in Section 4 through an example in multivariate linear regression models. In its most general form, the exact distribution of Y can be derived in terms of the Fox's H-function (see [10]). Although mathematically elegant, Fox's H-function are not numerically tractable even with the modern software packages. The exact distribution is therefore not practically appealing. To circumvent this difficulty a novel approximation for the distribution of Y is derived. We focus on the transformed variable $Z = -\ln Y$ because its characteristic function is known and since this way we transform the initial problem into a problem of sums or differences of independent random variables which can be addressed using characteristic functions. Following (2), the characteristic function of Z can be written as

$$\phi_Z(t) = E[\exp(jtZ)] = \prod_{i=1}^{n} \frac{\Gamma\left(r_i - \dfrac{jt}{\delta_i}\right)}{\Gamma\left(r_i\right)} \lambda_i^{jt}, \tag{6}$$

where $j = \sqrt{-1}$. Further manipulation of $\phi_Z(t)$ in its current form seems an arduous task. To this end, we decompose $\phi_Z(t)$ as shown in the following theorem.

Theorem 2.1 *The characteristic function of Z can be written as*

$$\phi_Z(t) = \phi_{Z_1}(t)\phi_{Z_2}(t) \tag{7}$$

where

$$\phi_{Z_1}(t) = \prod_{i=1}^{n} \frac{\Gamma\left(r_i + \gamma - \dfrac{jt}{\delta_i}\right)}{\Gamma\left(r_i + \gamma\right)}, \tag{8}$$

and

$$\phi_{Z_2}(t) = \prod_{i=1}^{n} \prod_{k=0}^{\gamma-1} \frac{\delta_i\left(r_i + k\right)}{\left(\delta_i\left(r_i + k\right) - jt\right)} \lambda_i^{jt}, \tag{9}$$

with $\gamma \in \mathbb{N}$.

Proof:
The following algebraic manipulation of the characteristic function of Z given in (6), completes the proof of Theorem 2.1

$$
\begin{aligned}
\phi_Z(t) &= \prod_{i=1}^{n} \frac{\Gamma\left(r_i - \dfrac{jt}{\delta_i}\right)}{\Gamma\left(r_i\right)} \lambda_i^{jt} \\
&= \prod_{i=1}^{n} \left\{ \frac{\Gamma\left(r_i + \gamma - \dfrac{jt}{\delta_i}\right)}{\Gamma\left(r_i + \gamma\right)} \right\} \frac{\Gamma\left(r_i + \gamma\right)}{\Gamma\left(r_i + \gamma - \dfrac{jt}{\delta_i}\right)} \times \frac{\Gamma\left(r_i - \dfrac{jt}{\delta_i}\right)}{\Gamma\left(r_i\right)} \lambda_i^{jt} \\
&= \left\{ \prod_{j=1}^{n} \frac{\Gamma\left(r_i + \gamma - \dfrac{jt}{\delta_i}\right)}{\Gamma\left(r_i + \gamma\right)} \right\} \times \left\{ \prod_{i=1}^{n} \prod_{k=0}^{\gamma-1} \frac{\left(r_i + k\right)}{\left(r_i + k - \dfrac{jt}{\delta_i}\right)} \lambda_i^{it} \right\}.
\end{aligned}
$$

The decomposition of $\phi_Z(t)$ above introduces a new parameter γ which will be used to control the degree of accuracy of the near-exact distributions. In Subsection 2.2, the effect of γ will be evaluated. In expression (7), $\phi_{Z_1}(t)$ and $\phi_{Z_2}(t)$ are the characteristic functions of Z_1 and Z_2 respectively. Then Z may be represented as

$$Z = Z_1 + Z_2 \tag{10}$$

where Z_1 and Z_2 are independent random variables. We conveniently deduce that Z_1 is a sum of n independent log-gamma random variables with parameters $r_i + \gamma$ and 1, multiplied by $1/\delta_i$ respectively. However, further decomposition of (9) is necessary to identify the distribution of Z_2.

Before proceeding, let us define some notations. Let $\beta_i^+ = \delta_i\left(r_i + k\right)$ if $\delta_i > 0$, m_i^+ a multiplicity of β_i^+ and ℓ^+ be the number of distinct values of β_i^+. Analogously, $\beta_i^- = \delta_i\left(r_i + k\right)$ if $\delta_i < 0$, m_i^- a multiplicity of β_i^- and ℓ^- be the number of distinct values of β_i^-. Let $\varphi = \sum_{i=1}^{n} \log\left(\lambda_i\right)$.

Using the above notations, (9) can be further decomposed as

$$
\begin{aligned}
\phi_{Z_2}(t) = &\left\{ \prod_{i=1}^{\ell^+} (\beta_i^+)^{m_i^+} \left(\beta_i^+ - jt\right)^{-m_i^+} \right\} \exp\left(jt\varphi\right) \\
&\times \left\{ \prod_{i=1}^{\ell^-} \left(\beta_i^-\right)^{m_i^-} \left(\beta_i^- - jt\right)^{-m_i^-} \right\}.
\end{aligned} \tag{11}
$$

By letting $\beta_i^* = -\beta_i^-$ (so that $\beta_i^* > 0$) and $\tau = -t$, (11) can be written as

$$\begin{aligned}\phi_{Z_2}(t) &= \left\{ \prod_{i=1}^{\ell^+} (\beta_i^+)^{m_i^+} \left(\beta_i^+ - jt\right)^{-m_i^+} \right\} \exp\left(jt\varphi\right) \\ &\quad \times \left\{ \prod_{i=1}^{\ell^-} (\beta_i^*)^{m_i^-} \left(\beta_i^* - j\tau\right)^{-m_i^-} \right\} \\ &= \phi_{Z_{21}}(t)\phi_{Z_{22}}(\tau) \exp\left(jt\varphi\right), \end{aligned} \tag{12}$$

where

$$\phi_{Z_{21}}(t) = \left\{ \prod_{i=1}^{\ell^+} (\beta_i^+)^{m_i^+} \left(\beta_i^+ - jt\right)^{-m_i^+} \right\}, \tag{13}$$

and

$$\phi_{Z_{22}}(\tau) = \left\{ \prod_{i=1}^{\ell^-} (\beta_i^*)^{m_i^-} \left(\beta_i^* - j\tau\right)^{-m_i^-} \right\}. \tag{14}$$

Note that (13) is the characteristic function of Z_{21} which is a sum of ℓ^+ independent Erlang distributed random variables, where m_i^+ and β_i^+ are shape and rate parameters, respectively. Similarly (from (14)) Z_{22} is a sum of ℓ^- independent Erlang distributed random variables where m_i^- and β_i^* are shape and rate parameters respectively. Both m_i^+ and m_i^- are integer valued parameters, therefore we conclude that

$$Z_{21} \sim \mathcal{GIG}\left(\ell^+, \underline{m}^+, \underline{\beta}^+\right),$$

and

$$Z_{22} \sim \mathcal{GIG}\left(\ell^-, \underline{m}^-, \underline{\beta}^-\right),$$

where $\mathcal{GIG}$ denote the generalized integer gamma distribution [17] and

$$\begin{aligned} \underline{m}^+ &= \left(m_1^+, m_2^+, \ldots, m_{\ell^+}^+\right)', \\ \underline{\beta}^+ &= \left(\beta_1^+, \beta_2^+, \ldots, \beta_{\ell^+}^+\right)', \\ \underline{m}^- &= \left(m_1^-, m_2^-, \ldots, m_{\ell^-}^-\right)', \text{ and} \\ \underline{\beta}^* &= \left(\beta_1^*, \beta_2^*, \ldots, \beta_{\ell^-}^*\right)'. \end{aligned} \tag{15}$$

Since $\tau = -t$, (12) is a characteristic function of a random variable Z_2 such that $Z_2 = Z_{21} - Z_{22}$ with a location parameter φ, and with Z_{21} and Z_{22} independent. Therefore

$$Z_2 \sim \mathcal{SDGIG}\left(\underline{m^+}, \underline{m^-}, \underline{\beta^+}, \underline{\beta^*}, \ell^+, \ell^-, \varphi\right).$$

where $\mathcal{SDGIG}$ denote the shifted difference of generalized integer gamma distribution [18].

Having armed with the above results, now we are in a position to obtain approximations of the distribution of Z (and in turn, approximations of the distribution of Y).

2.2 Near-exact distributions for Z

It has already been shown in (10) that Z can be decomposed into a sum of two independent random variables. To obtain near-exact distributions of Z, only one part of the decomposition will be replaced by its approximation while the other one is left unchanged. In this respect, the exact distribution of Z_2 is known. However, the exact distribution of Z_1 is unknown. Therefore, Z_1 will be replaced by its approximation.

2.2.1 First near-exact distribution

To derived the first near-exact distribution, Z_1 is replaced by its expected value i.e., $E[Z_1]$. As a result, the first near-exact approximate of Z can be obtained as

$$Z^a = E[Z_1] + Z_2,$$

with

$$E[Z_1] = \frac{1}{j} \frac{\partial \phi_{Z_1}(t)}{\partial t}\bigg|_{t=0}.$$

where $\phi_{Z_1}(t)$ is given by (8). Therefore, we conclude that

$$Z^a \sim \mathcal{SDGIG}\left(\underline{m}^+, \underline{m}^-, \underline{\beta}^+, \underline{\beta}^*, \ell^+, \ell^-, \varphi + E[Z_1]\right),$$

Now the first near-exact distribution of Y can be obtained by a trivial transformation and thus omitted.

2.2.2 Second near-exact distribution

Instead of approximating Z_1 by its mean value, here it is approximated by a random variable. A single log-gamma random variable can be represented by a sum of infinite independent shifted exponential random variables (see [18]). Therefore, Z_1 which is a sum of independent log-gamma random variables can represented by an infinite sum of independent shifted Erlang distributed random variables. Erlang distributed random variables are special cases of the gamma distributed random variables. We propose as an approximation for the distribution of Z_1 a single shifted gamma random variable, say W, independent of Z_2 and such that the following system of equations is satisfied

$$\frac{\partial^i \phi_{Z_1}(t)}{\partial t^i}\bigg|_{t=0} = \frac{\partial^i \phi_W(t)}{\partial t^i}\bigg|_{t=0} \qquad i = 1, 2, 3, \tag{16}$$

where ϕ_{Z_1} is given by (8) and the characteristic function of an approximate random variable W of Z_1 is denoted by $\phi_W(t)$ with its form given by

$$\phi_W(t) = \left(1 - \frac{jt}{\psi}\right)^{-\rho} e^{jt\theta}. \tag{17}$$

The values of ψ, ρ and θ are obtained by numerically solving (16).

The solution of the system of equations in (16) may give a positive or a negative value for ψ, therefore the second near-exact approximation for the distribution of Z is, given by

$$Z^b = Z_2 + \text{sign}(\psi) W,$$

where

$$\text{sign}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0. \end{cases}$$

Clearly, Z^b is either a sum or difference of a shifted gamma random variable and an independent $\mathcal{SDGIG}$ distributed random variable. The expressions for the distribution Z^b are available in Appendix 1 of [18]. Now the second near-exact distribution of Y can be obtained by a trivial transformation and thus it is omitted.

3 Computational studies

In this section, computations are performed on various distributions and results are contrasted against each other.

Table 1 summarizes cases which will be used for computations in this paper. In each case, at least one power parameter will be negative in order to ensure that statistic Y (in (5)) is studied.

Table 1: Sets of parameters of independent generalized gamma distribution

Case	r	λ	δ
I	$\left\{\frac{1}{3}, \frac{22}{7}\right\}$	$\left\{\frac{1}{2}, \frac{1}{4}\right\}$	$\{4, -2\}$
II	$\{2, 3, 5\}$	$\{3, 2, 10\}$	$\left\{\frac{1}{2}, 2, -\frac{1}{4}\right\}$
III	$\left\{2, 3, 5, \frac{1}{2}\right\}$	$\left\{3, 2, 10, \frac{2}{7}\right\}$	$\left\{\frac{1}{2}, 2, -\frac{1}{4}, -\frac{1}{3}\right\}$

The performance of near-exact distributions will be analyzed using the proximity measure, Δ defined as

$$\Delta = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \frac{\phi_Z(t) - \phi^*(t)}{t} \right| dt, \tag{18}$$

where $\phi_Z(t)$ and $\phi^*(t)$ are respectively the exact and the approximate characteristic functions of the random variable Z. See [19] for more details on the proximity measure, Δ. Since

$$\sup_{z \in \mathbb{R}} |F_Z(z) - F^*(z)| \leq \Delta,$$

then Δ provides an upper bound on the proximity between $F_Z(z)$ and $F^*(z)$. $F_Z(z)$ and $F^*(z)$ are the exact and approximate cummulative distribution function of the random variable Z respectively.

Tables 2 and 3 provide proximity measures for first and second near-exact distributions respectively. It is evident, from these tables and from low values of Δ, that both near-exact distributions provide good approximations for the exact distribution of Z. Furthermore, the accuracy of the approximation increases with an increase on the value of the precision parameter and hence this parameter can be use to control the degree of accuracy. Between the first and second near-exact distribution, the latter provides better accuracy for a given precision parameter value.

Next, the performance of near-exact, empirical and exact distributions will be contrasted against each other using cdf plots. In Figure 1, we have, for Case I, a normal and a magnified representation of the empirical cdf, obtained using a simulated random sample of size $n = 10^6$, of the cdf of the near-exact distribution obtained setting $\gamma = 20$ and the exact cdf derived using the inversion formulas in [19]. From the normal representation in Figure 1 it is

Table 2: Proximity measures for the first near-exact distribution

Precision parameter γ	Case		
	I	II	III
4	1.6×10^{-2}	5.2×10^{-2}	1.7×10^{-2}
5	1.3×10^{-2}	4.4×10^{-2}	1.4×10^{-2}
10	6.9×10^{-3}	2.5×10^{-2}	7.8×10^{-3}
15	4.8×10^{-3}	1.7×10^{-2}	5.4×10^{-3}
20	3.6×10^{-3}	1.3×10^{-2}	4.1×10^{-3}
50	1.5×10^{-3}	5.5×10^{-3}	1.7×10^{-3}
100	7.5×10^{-4}	2.8×10^{-3}	8.7×10^{-4}
200	3.8×10^{-4}	1.4×10^{-3}	4.4×10^{-4}

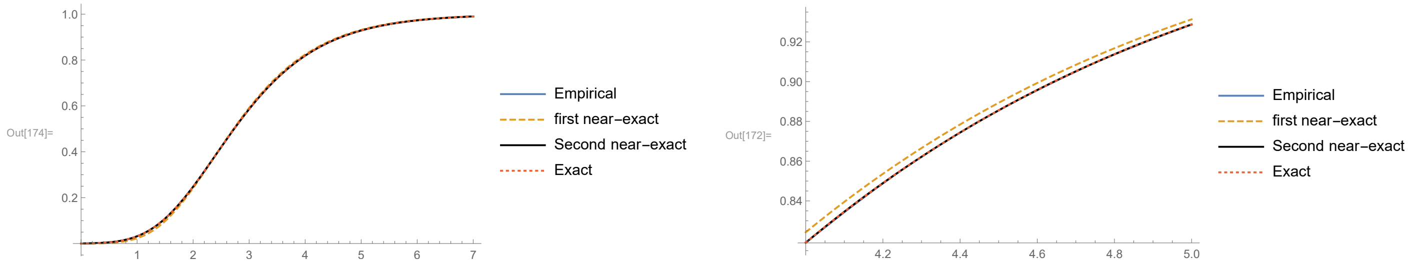
Table 3: Proximity measures for the first near-exact distribution

Precision parameter γ	Case		
	I	II	III
4	9.4×10^{-5}	3.4×10^{-4}	5.3×10^{-5}
5	5.6×10^{-5}	2.3×10^{-4}	3.1×10^{-5}
10	1.0×10^{-5}	5.8×10^{-5}	5.4×10^{-6}
15	3.4×10^{-6}	2.3×10^{-5}	1.9×10^{-6}
20	1.6×10^{-6}	1.1×10^{-5}	8.9×10^{-7}
50	1.2×10^{-7}	9.4×10^{-7}	7.1×10^{-8}
100	1.5×10^{-8}	1.3×10^{-7}	9.6×10^{-9}
200	2.0×10^{-9}	1.7×10^{-8}	1.4×10^{-9}

difficult to distinguish the cdfs however, the magnified version, over a subset of the domain, provides insights into the difference between the approximate and the exact distributions. The second near-exact distribution lies almost exactly over the exact distribution in Figure 1, this illustrates its high degree of accuracy. We may draw similar conclusions for Cases II and III from Figures 2 and 3.

Lastly, near-exact distributions are evaluated in terms of their computational time requirements relative to each other. By their very nature, exact and empirical distributions are time consuming and will not be feasible for regular use. Tables 4, 5 and 6 show, respectively for Cases I, II and III, the computer run-time for the first and second near-exact distributions, for various precision parameter value. For both near-exact distributions, computer run-time increase with an increase in precision parameter. The second near-exact distribution, the more accurate, takes relatively longer time. It can be concluded that with the near-exact distributions, the accuracy is at a cost of computer run-time.

(i) (ii)



PLEASE, define just one legend for both figure AND REMOVE 'Out[174]=' and 'Out[172]=' from figures

Figure 1: Case I - (i) normal representation and (ii) magnified version, over a subset of the domain, of the first near-exact, second near-exact, empirical and exact cdf plots

(i) (ii)

Insert figure NORMAL Insert figure MAGNIFIED version

Figure 2: Case II - (i) normal representation and (ii) magnified version, over a subset of the domain, of the first near-exact, second near-exact, empirical and exact cdf plots

4 Applications

4.1 The distribution of the ratio of generalized variances

The generalized variance was first introduced in [20]. It is a very important concept in multivariate analysis since a number of statistics are in terms of the generalized variance e.g. Wilks lambda. It is a univariate measure of spread for multivariate data. In this section, a ratio of two independent generalized variance is studied. This particular statistic was first studied in its introductory paper [20]. It can be viewed as being analogous with the $F-$ test statistic in univariate analysis of variance ([21], p.292). From this point of view, the statistic should receive wide application in multivariate analysis of variance (MANOVA). However, the Wilks' lambda (denoted Λ) has instead been applied widely to "Its mathematical tractability, and further it is related to the likelihood ratio criterion" ([21], p.292). Below we show that the ratio of generalized variances is distributed as a ratio of independent χ^2 distributed random variables. Therefore, near-exact distribution of the ratio of independent generalized gamma random variables can be useful in modeling the ratio of independent generalized variances and therefore circumventing the issue of mathematical tractability. Whereas Wilks' lambda bears some resemblance with Pillai's trace which is also used in multivariate analysis of variance (MANOVA), the ratio of generalized gamma bears a similar resembles with Hotelling−Lawley trace also used in MANOVA. This provides more rationale to its candidacy as another test statistic in MANOVA. Suppose that A and

(i) Insert figure NORMAL   (ii) Insert figure MAGNIFIED version

Figure 3: Case III - (i) normal representation and (ii) magnified version, over a subset of the domain, of the first near-exact, second near-exact, empirical and exact cdf plots

Table 4: Case I: Run-time for near-exact distributions

Distribution type	Precision parameter			
	4	10	15	20
First near-exact	0.016	0.062	0.109	0.172
Second near-exact	0.671	1.547	2.390	3.266

B are two independent $p \times p$ random matrices such that $A \sim W(n_A, \Sigma)$ and $B \sim W(n_B, \Sigma)$. It follows that $\frac{|A|}{|\Sigma|} \cong \prod_{j=1}^{p} Y_{n_A-j+1}$ where $Y_{n_A-j+1} \sim \chi^2(n_A - j + 1)$ (see [22], p.100 theorem 3.2.15). Similarly, $\frac{|B|}{|\Sigma|} \cong \prod_{j=1}^{p} Y_{n_B-j+1}$ where $Y_{n_B-j+1} \sim \chi^2(n_B - j + 1)$.

Let

$$W = \frac{|A|}{|B|}. \tag{19}$$

Another representation of (19) is

$$W = \frac{|A|/|\Sigma|}{|B|/|\Sigma|} \cong \frac{\prod_{j=1}^{p} Y_{n_A-j+1}}{\prod_{j=1}^{p} Y_{n_B-j+1}}.$$

Therefore (4.1) is distributed as a ratio of χ^2 distributed random variables i.e. a special case of (5). By application of results provided in [10], the pdf of W (see (4.1)) terms of the Meijer's$-G$ function is

$$f_W(w) = \prod_{j=1}^{p} \frac{1}{\Gamma(\frac{n_A-j+1}{2})\Gamma(\frac{n_B-j+1}{2})} G^{n_B\ n_A}_{n_A\ n_B}[w|^{-\frac{n_B}{2},-\frac{n_B-1}{2},\ldots,-\frac{n_B-p+1}{2}}_{\frac{n_A}{2}-1,\frac{n_A-1}{2}-1,\ldots,\frac{n_B-p+1}{2}-1}] \tag{20}$$

By first noting that W in (19) can be represented as a product of independent beta random variables, [23] arrived the same pdf of W as (20).

Below, an application of the test statistic W in multivariate multiple regression is suggested.

Consider the well-known a regression model

$$\underset{}{\boldsymbol{Y}} = \underset{n\times(r+1)}{\boldsymbol{Z}} \underset{(r+1)\times m}{\boldsymbol{\beta}} + \underset{n\times m}{\boldsymbol{\epsilon}} \tag{21}$$

Such that

Table 5: Case II: Run-time for near-exact distributions

Distribution type	Precision parameter			
	4	10	15	20
First near-exact	INSERT VALUES	?	?	?
Second near-exact	INSERT VALUES	?	?	?

Table 6: Case III: Run-time for near-exact distributions

Distribution type	Precision parameter			
	4	10	15	20
First near-exact	INSERT VALUES	?	?	?
Second near-exact	INSERT VALUES	?	?	?

$$E(\boldsymbol{Y}) = \boldsymbol{Z}\boldsymbol{\beta} \tag{22}$$

Let $\boldsymbol{Z} = [\underset{n\times(q+1)}{\boldsymbol{Z_1}} \mid \underset{n\times(r-q)}{\boldsymbol{Z_2}}]$ and $\boldsymbol{\beta} = [\underset{(q+1)\times m}{\boldsymbol{\beta_1}} \mid \underset{(r-q)\times m}{\boldsymbol{\beta_2}}]$ is an $(r+1)\times m$ matrix of unknown parameters. Our interest is to test the hypothesis

$$H_0 : \boldsymbol{\beta}_{(2)} = \mathbf{0} \quad \text{against} \quad H_1 : \boldsymbol{\beta}_{(2)} \neq \mathbf{0}\,. \tag{23}$$

Let

$$\begin{aligned} \hat{\boldsymbol{\Sigma}} &= n^{-1}(\boldsymbol{Y} - \boldsymbol{z}\hat{\boldsymbol{\beta}})'(\boldsymbol{Y} - \boldsymbol{z}\hat{\boldsymbol{\beta}}) \\ \hat{\boldsymbol{\Sigma}}_{\mathbf{1}} &= n^{-1}(\boldsymbol{Y} - \boldsymbol{z}\hat{\boldsymbol{\beta}}_{(1)})'(\boldsymbol{Y} - \boldsymbol{z}\hat{\boldsymbol{\beta}}_{(1)}) \end{aligned} \tag{24}$$

such that $\boldsymbol{E} = n\hat{\boldsymbol{\Sigma}} \sim W(n-r-1, \boldsymbol{\Sigma})$ independently of $\boldsymbol{H} = n(\hat{\boldsymbol{\Sigma}}_{\mathbf{1}} - \hat{\boldsymbol{\Sigma}}) \sim W(r-q, \boldsymbol{\Sigma})$ with n the number of observations, r the number of regressors in a full model and $q+1$ the number regressors in a model under the null hypothesis in (23). $\boldsymbol{Z}$ is assumed to be a full rank matrix. However, if it is not full rank but has rank r_1+1 then r is replaced above by r_1 and $q+1$ is replaced by rank of $\boldsymbol{Z}_1$. See ([24], page 396) for details.

In ([24], p.397, Example 7.9) Wilks lambda was used to test H_0. The distribution of Wilks lambda was approximated using the so called χ^2- approximation *i.e.* $-2\ln\Lambda \dot{\sim} \chi^2$ with $m(q+1)$ degrees of freedom where "$\dot{\sim}$" denotes "approximately distributed as". The disadvantage of this approximation is that it performs very poorly when the sample size is small. We propose the following test statistics for H_0

$$W^* = \frac{|H|}{|E|}. \tag{25}$$

Note that W^* is an application of the test statistic W (see (19)) in multivariate multiple regression analysis. Therefore, near-exact distributions proposed in Subsection 2.2 can be use to approximate the distribution of W^*.

In the example below, H_0 will be tested at 5% level with near-exact distributions used to approximate associated $p-value$. The efficiency of near-exact distributions in the calculation of $p - values$ will be contrasted against $p - values$ calculation using 20.

Example 4.1 *In [24], p.397, Example 7.9, the author considers 18 male and female customers rated the service in three locations of a larger restaurant chain. The ratings were converted into an index. The following are the sample estimated of the study. A statistical test for the significance of interaction was conducted with the following sample statistics*

$$\boldsymbol{E} = \begin{bmatrix} 2977.39 & 1021.72 \\ 1021.72 & 2050.95 \end{bmatrix}$$

and

$$\boldsymbol{H} = \begin{bmatrix} 441.76 & 246.16 \\ 246.16 & 366.12 \end{bmatrix}$$

Therefore $W^* = 0.0205273$. Table 7 presents the $p-$values from χ^2 approximation (critical value is 3.42 with 8 degrees of freedom [24]), near-exact distribution (precision parameter is) and exact distribution. Despite the differences between the values in Table 7 , all the $p-$values point to the non rejection of the null hypothesis.

Table 7: $p-$values provided by the different approximations and the exact p-values obtained using the inversion formulas in [19]

Method	$p-$value
χ^2 approximation	0.51
First near-exact	0.22697
Second near-exact	0.23357
Exact	0.23357

4.2 Application to SISO Networks

In [25] the authors consider a 2-dimensional strip-shaped network with randomly distributed finite number of nodes. They assume that number of nodes in given area is described by stationary Poisson point process. It is considered that the message propagates in a single-input single-output (SISO) fashion over multiple hops. This process is illustrated in Figures 1 and 2 of [25]. As mentioned by the authors, the received power (considering only two channel distortions; fading and path loss) is given by

$$P_r = \frac{P_t X}{d^\alpha} \tag{26}$$

where

- P_t is the transmit power;

- X characterizes the phenomenon of fading, and is assumed to have an exponential distribution, with density

$$f_X(x) = \mu \exp\{-\mu x\}$$

which may be obtained from expression (1) by making $\delta = 1$, $\lambda = \mu$, $r = 1$, thus, as already referred, it is a particular case of the generalized gamma distribution;

- and finally $d^\alpha = Y$, where d represents the Euclidean distance between the transmitter and the receiver and α is the path loss exponent is, modelled by a generalized gamma distribution. Expression (6) in [25] may be obtain from expression (1) in this paper by considering $r = k$, $\lambda = 1/\theta$ and $\delta = \beta$.

Thus, the received power in (26) is the ratio o two random variables, one with an exponential distribution-which is a particular case of the generalized gamma distribution-and the other with a generalized gamma distribution. In [25] the authors give an exact representation for the cdf of X/Y in terms of the Fox H-function which is not easy to use in practical terms. Therefore, we propose the use of the first or second near-exact distribution developed in this paper to approximate the distribution of X/Y. Clearly, our results may also be used to approximate the distribution of products of ratios, that is to approximate $\prod_{i=1}^{p} X_i/Y_i$.

To illustrate the use of our approximations in this application we present Figure 4 with the cdfs of X/Y for the cases $k = 2, \theta = 15^\beta$, $\mu = 1$ and $\beta = 0.8, 1, 2$. These are the same cases addressed in Figure 3 of [25].

INSERT FIGURE PLEASE

Figure 4: cdf of X/Y for $k = 2, \theta = 15^\beta$, $\mu = 1$ and $\beta = 0.8, 1, 2$

In Table 8 it is possible to observe the computing times for the cdfs of both near-exact distributions, computed at the same specific value (empirical 0.05 quantile evaluated from a simulated sample of size 10^6) for the cases considered in Figure 4.

Table 8: Computing times for the first and second near-exact cdfs - for $k = 2, \theta = 15^\beta$ and $\mu = 1$

	Computing time in seconds		
	$\beta = 0.8$	$\beta = 1$	$\beta = 2$
First near-exact	INSERT VALUES		
Second near-exact	INSERT VALUES		

5 Computational modules

Still being formatted for a more user-friendly layout. Will be added later

6 Conclusion

Two near-exact distributions were developed to approximate the distribution of the ratio of independent generalized gamma random variables. Both distributions can be implemented computationally and therefore used in practice. The second near-exact distribution is more accurate than the first one, however requires a longer computer run-time. The precision of both near-exact distributions may be adjusted using the parameter γ. Large values of γ provide high precision but with natural costs in computing time. The second near-exact distribution should be used in problems where a high level of precision is required and the first one for simple and practical problems where speed of execution may be important. In the examples provided it is clear the important role that the ratio of generalized gamma variables may have in practical problems and applications.

7 Acknowledgements

This research was partially funded by National Research Fund (Vulnerable discipline: Academic statistics and Re: CPRR160403161466 grant No. 105840) and STATOMET and by the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through the project UID/MAT/00297/2013 (Centro de Matemática e Aplicações).

References

[1] E.W. Stacy. A generalization of the gamma distribution. *The Annals of Mathematical Statistics*, 33(3):1187–1192, 1962.

[2] M. M. Ali, J. Woo, and S. Nadarajah. Generalized gamma variables with drought application. *Journal of the Korean Statistical Society*, 37:37–45, 2008.

[3] Y. Chen, G.K. Karagiannidis, H. Lu, and N. Cao. Novel approximations to the statistics of products of independent random variables and their applications in wireless communications. *IEEE Trans. Veh. Technol.*, 61(2):443–454, February 2012.

[4] M. D. Yacoub. The α-μ distribution: A physical fading model for the Stacy distribution. *IEEE Transactions on Vehicular Technology*, 56(1):27–34, 2007.

[5] E. Mekic, N. Sekulovic, M. Bandjur, M. Stefanovic, and P. Spalevic. The distribution of ratio of random variable and product of two random variables and its application in performance analysis of multi-hop relaying communications over fading channels. *Przeglad Elektrotechniczny*, 88(7a):133–137, 2012.

[6] F.J. Marques and F. Loingeville. Improved near-exact distributions for the product of independent Generalized Gamma random variables. *Computational Statistics and Data Analysis*, 102:55–66, 2016.

[7] C. A. Coelho and B. C. Arnold. On the exact and near-exact distribution of the product of generalized Gamma random variables and the generalized variance. *Communications in Statistics - Theory and Methods*, 43(10-12):2007–2033, 2014.

[8] M. D. Springer. *The algebra of random variables*. Springer, 1979.

[9] J. Salo, H.M. El-Sallabi, and P. Vainikainen. The distribution of the product of independent Rayleigh random variables. *IEEE Trans. Antennas Propag.*, 54(2):639–643, 2006.

[10] A. M. Mathai. Products and ratios of generalized gamma variates. *Scandinavian Actuarial Journal*, 2:193–198, 1972.

[11] M. D. Springer and W. E. Thompson. The distribution of products of Beta, Gamma and Gaussian random variables. *SIAM Journal on Applied Mathematics*, 18(4):721–737, June 1970.

[12] H. Lu, Y. Chen, and N Cao. Accurate approximation to the pdf of the product of independent Rayleigh random variables. *IEEE Antennas and Wireless Propagation Letters*, 10:1019–1022, 2011.

[13] C.A. Coelho and J.T. Mexia. On the distribution of the product and ratio of independent Generalized Gamma-Ratio random variables. *Sankhya: The Indian Journal of Statistics*, 69(2):221–255, 2007.

[14] S. Ahmed, L.-L. Yang, and L. Hanzo. Probability distributions of products of Rayleigh and Nakagami-m variables using Mellin transform. In *ICC*, pages 1–5. IEEE, 2011.

[15] H.J. Malik. Exact distribution of the quotient of independent generalized gamma variables. *Canad. Math. Bull.*, 10:463–465, 1967.

[16] M. Ahsen and S. A. Hassan. On the ratio of Exponential and Generalized Gamma random variables with applications to ad hoc SISO networks. In *2016 IEEE 84th Vehicular Technology Conference (VTC-Fall)*, pages 1–5, Sept 2016.

[17] C. A. Coelho. The generalized integer Gamma distribution a basis for distributions in multivariate statistics. *Journal of Multivariate Analysis*, 64(MV971710):86–102, 1998.

[18] F.J. Marques, C. A. Coelho, and M. de Carvalho. On the distribution of linear combinations of independent Gumbel random variables. *Statistics and Computing*, 25:683–701, 2014.

[19] J. Gil-Pelaez. Note on the inversion theorem. *Biometrika*, 38:481–482, 1951.

[20] S. S Wilks. Certain generalizations in the analysis of variance. *Biometrika*, 24(3):471–494, 1932.

[21] A. M. Kshirsagar. *Multivariate Analysis*. Marcel Dekker, New York, 1972.

[22] R. J Muirhead. *Aspects of Multivariate Statistical Theory*. 1982.

[23] T. Pham-Gia. Exact distribution of the generalized wilks's statistic and application. *Journal of Multivariate Analysis*, (99):1698–1716, 2008.

[24] R. A. Johnson and D.W. Wichern. *Applied Multivariate Statistical Analysis*. Pearson, 6th edition, 2007.

[25] M. Ahsen and S. A. Hassan. On the ratio of exponential and generalized gamma random variables with applications to ad hoc siso networks. pages 1–5, Sept 2016.