Stat Papers (2009) 50:323–338
DOI 10.1007/s00362-007-0083-2

REGULAR ARTICLE

The type I distribution of the ratio of independent "Weibullized" generalized beta-prime variables

Andriëtte Bekker · Jacobus Roux · Thu Pham-Gia

Received: 13 March 2007 / Revised: 18 June 2007 / Published online: 4 August 2007

Abstract In this paper we introduce the distribution of $\frac{X_1^c}{X_1^c+X_2^c}$, with $c > 0$, where X_i, $i = 1, 2$, are independent generalized beta-prime-distributed random variables, and establish a closed form expression of its density. This distribution has as its limiting case the generalized beta type I distribution recently introduced by Nadarajah and Kotz (2004). Due to the presence of several parameters the density can take a wide variety of shapes.

Keywords Gauss hypergeometric function · Generalized beta-prime distribution · Incomplete beta function · Meijer's G-function · Predictive distribution · Income distribution

Mathematics Subject Classification (2000) 62E15 · 62H10

1 Introduction

The importance of the density of a ratio of random variables taken to a certain power, $c > 0$, with parameter c as a weighting factor has been identified in the literature (see e.g., Mathai and Moschopoulos 1997; Nadarajah and Kotz 2004). In the present paper,

A. Bekker (✉) · J. Roux
Department of Statistics, Faculty of Natural and Agricultural Sciences,
University of Pretoria, Pretoria 0002, South Africa
e-mail: andriette.bekker@up.ac.za

T. Pham-Gia
Department of Mathematics and Statistics,
Université de Moncton, Moncton, NB E1A 3EO, Canada
e-mail: thu.pham-gia@umoncton.ca

we introduce the distribution of

$$U_c = \frac{X_1^c}{X_1^c + X_2^c},$$

where $X_i, i = 1, 2$, are independent *generalized beta-prime* (*GBP*) distributed variables, denoted by $X_i \sim GBP(\alpha_i, \beta_i, \lambda_i)$, with probability density function (pdf)

$$f_i(x; \alpha_i, \beta_i, \lambda_i) = \lambda_i^{\alpha_i} \frac{x^{\alpha_i - 1}}{B(\alpha_i, \beta_i)(1 + \lambda_i x)^{\alpha_i + \beta_i}}, \qquad x > 0, \alpha_i, \beta_i, \lambda_i > 0 \quad (1)$$

This generalized beta-prime distribution has attracted useful applications in many areas including system availability and measuring information in predictive distributions, see Dyer (1982), Amaral-Turkman and Dunsmore (1985), Sarhan (1995) and Pham-Gia and Turkkan (2002). Both the generalized beta-prime distribution and the gamma distribution are special cases of the parameter-rich generalized hypergeometric distribution defined by Mathai and Saxena (1966) where the gamma distribution is a limiting case of (1) if λ_i is substituted with λ_i/β and $\beta \to \infty$ (Bekker 1990). This additional parameter, λ_i in the generalized beta-prime density (1) results in a more versatile distribution that's more responsive to modeling needs (Pham-Gia and Duong 1989).

For positive random variables (such as gamma and exponential), their powers X^c, with c positive, are often encountered, resulting in the so-called "Weibullized" distributions (see e.g., Gupta and Nadarajah 2004, p. 118; Malik 1967; McDonald and Xu 1995; Bekker et al. 2000). A particularly interesting feature of the variable X^c, with X having the generalized beta-prime distribution, is the flexibility of its hazard function.

For independent $X_1 \sim gamma(\alpha, \varphi_1)$ and $X_2 \sim gamma(\beta, \varphi_2)$, it is well known that the densities of the type I and type II ratios $X_1/(X_1 + X_2)$ and X_1/X_2 are the standard beta and standard beta-prime, respectively, provided $\varphi_1 = \varphi_2 = \varphi$. In the general case, however, the second ratio has the $GBP(\alpha, \beta, \lambda)$ distribution with $\lambda = \varphi_1/\varphi_2$. Therefore, we refer to the distribution of U_c, as defined before, as the type I distribution of independent " Weibullized" generalized beta-prime variables.

Some properties of the distribution of U_c are presented and it also turns out that density of Nadarajah and Kotz (2004) (see Eq. (7) in their paper) is a special limiting case of the density of U_c. This parameter-rich distribution of U_c can take a wide variety of shapes and can be applied in Bayesian statistics and economics.

2 Mathematical preliminaries

Before we prove the main results, we shall state some results which are necessary in proving our results. It is well known in the literature (see e.g., Gradshteyn and Ryzhik 2000) that the Gauss hypergeometric function and Meijer's G-function are defined as follows:

Definition 1 The Gauss hypergeometric function in 3 parameters a, b and c, denoted by ${}_2F_1$, is defined as follows:

$$
{}_2F_1\left(a,b;c;x\right)=\sum_{n=0}^{\infty}\frac{(a,n)\,(b,n)}{(c,n)}\cdot\frac{x^n}{n!},\qquad |x|<1,
$$

with the Pochhammer coefficient
$(a,n)=a\,(a+1)\,(a+2)\cdots(a+n-1)$, for $n\geq 1$ and $(a,0)=1$.

Definition 2 Meijer's G-function $G(x)=G_{p,q}^{m,n}\left[x\left|\begin{matrix}a_1,\ldots,a_p\\ b_1,\ldots,b_q\end{matrix}\right.\right]$ is defined as

$$
G_{p,q}^{m,n}\left[x\left|\begin{matrix}a_1,\ldots,a_p\\ b_1,\ldots,b_q\end{matrix}\right.\right]
=\frac{1}{2\pi i}\int_L\frac{\prod_{j=1}^{m}\Gamma\left(b_j+s\right)\prod_{j=1}^{n}\Gamma\left(1-a_j-s\right)}{\prod_{j=m+1}^{q}\Gamma\left(1-b_j-s\right)\prod_{j=n+1}^{p}\Gamma\left(a_j+s\right)}x^{-s}ds,
$$

where the integral is taken along the complex contour L. Under some fairly general conditions on the integers m, n, p and q, and on the parameters a_i, $i=1,\ldots,p$ and b_j, $j=1,\ldots,q$, the above integral exists. (For a discussion of the G-function, see Mathai 1993).

Now we state the following results that will be needed (see Mathai and Saxena 1973, p. 37), Gradshteyn and Ryzhik 2000, p. 1025), also Erdélyi et al. 1953, p. 87).

Result 1

$$
{}_2F_1\left(a,b;c;1-x\right)=\frac{\Gamma\left(c\right)}{\Gamma\left(a\right)\Gamma\left(b\right)\Gamma\left(c-a\right)\Gamma\left(c-b\right)}G_{2,2}^{2,2}\left[x\left|\begin{matrix}1-a & 1-b\\ 0 & c-a-b\end{matrix}\right.\right],\tag{2}
$$

provided $c-a, c-b\neq 0,-1,-2,\ldots,$ and a more general relation relates ${}_pF_q$ to the G-function $G_{q+1,p}^{p,1}$.

Result 2 The cumulative distribution function (cdf) of X_i, where $X_i\sim GBP(\alpha_i,\beta_i,\lambda_i)$, is $F(t)=\frac{B_{\frac{\lambda_i t}{1+\lambda_i t}}(\alpha_i,\beta_i)}{B(\alpha_i,\beta_i)}$, with the incomplete beta function given by

$$
B_x\left(a,b\right)=\int_0^x y^{a-1}\left(1-y\right)^{b-1}dy=a^{-1}x^a{}_2F_1\left(a,1-b;a+1;x\right).\tag{3}
$$

3 Main results

In this section the type 1 distribution of the ratio of independent " Weibullized" generalized beta-prime variables will be derived and some properties discussed.

Theorem 1 *Let* $X_i \sim GBP(\alpha_i, \beta_i, \lambda_i)$, $i = 1, 2$, *be independent. The variable* $U_c = \frac{X_1^c}{X_1^c + X_2^c}$, *where* $c > 0$ *is a constant, possesses the density*:

$$f(u) = K_1 (1-u)^{(\alpha_2/c)-1} u^{-(1+\alpha_2/c)} \times {}_2F_1\left[\alpha_2 + \beta_2, \alpha_1 + \alpha_2; \alpha_1 + \alpha_2 + \beta_1 + \beta_2; 1 - \frac{\lambda_2}{\lambda_1}\left(\frac{1-u}{u}\right)^{1/c}\right], \quad 0 < u < 1, \tag{4}$$

where $K_1 = \frac{B(\alpha_1+\alpha_2, \beta_1+\beta_2)}{cB(\alpha_1, \beta_1)B(\alpha_2, \beta_2)}\left(\frac{\lambda_2}{\lambda_1}\right)^{\alpha_2}$.

Proof Let $X_i, i = 1, 2$, be independent nonnegative random variables with densities of the form $A_i x_i^{\alpha_i - 1} g_i(x_i)$. The joint density is $A_1 A_2 x_1^{\alpha_1 - 1} x_2^{\alpha_2 - 1} g_1(x_1) g_2(x_2)$.

Changing the variables from (X_1, X_2) to $\left(U_c = \frac{X_1^c}{X_1^c + X_2^c}, Z = X_1^c + X_2^c\right)$, we note the inverse transformation

$$X_1 = (U_c Z)^{\frac{1}{c}}, \qquad X_2 = ((1 - U_c) Z)^{\frac{1}{c}}.$$

The Jacobian of the transformation is

$$c^{-2} u^{\frac{1}{c}-1} (1-u)^{\frac{1}{c}-1} z^{\frac{2}{c}-1}.$$

The joint density is given by

$$\begin{aligned} &A_1 A_2 c^{-2} u^{\frac{1}{c}-1} (1-u)^{\frac{1}{c}-1} z^{\frac{2}{c}-1} (uz)^{(\alpha_1 - 1)/c} ((1-u) z)^{(\alpha_2 - 1)/c} \\ &\quad \times g_1\left((uz)^{\frac{1}{c}}\right) g_2\left(((1-u) z)^{\frac{1}{c}}\right) \\ &= A_1 A_2 c^{-2} u^{\frac{\alpha_1}{c}-1} (1-u)^{\frac{\alpha_2}{c}-1} z^{(\alpha_1+\alpha_2)/c-1} g_1\left((uz)^{\frac{1}{c}}\right) g_2\left(((1-u) z)^{\frac{1}{c}}\right) \end{aligned}$$

and the marginal density of U_c becomes

$$A_1 A_2 c^{-2} u^{\frac{\alpha_1}{c}-1} (1-u)^{\frac{\alpha_2}{c}-1} \int_0^\infty z^{(\alpha_1+\alpha_2)/c-1} g_1\left((uz)^{\frac{1}{c}}\right) g_2\left(((1-u) z)^{\frac{1}{c}}\right) dz. \tag{5}$$

In this case $A_1 = \lambda_1^{\alpha_1}/B(\alpha_1, \beta_1)$, $A_2 = \lambda_2^{\alpha_2}/B(\alpha_2, \beta_2)$, $g_1(x_1) = 1/(1+\lambda_1 x_1)^{\alpha_1+\beta_1}$ and $g_2(x_2) = 1/(1 + \lambda_2 x_2)^{\alpha_2+\beta_2}$.

Let $p_1 = \lambda_1 u^{\frac{1}{c}}$ and $p_2 = \lambda_2 (1-u)^{\frac{1}{c}}$. Then the above integral in (5) becomes

$$\begin{aligned}
&\int_0^\infty z^{(\alpha_1+\alpha_2)/c-1} \left(1 + p_1 z^{\frac{1}{c}}\right)^{-(\alpha_1+\beta_1)} \left(1 + p_2 z^{\frac{1}{c}}\right)^{-(\alpha_2+\beta_2)} dz \\
&= c p_1^{-(\alpha_1+\beta_1)} p_2^{-(\alpha_2+\beta_2)} \int_0^\infty v^{\alpha_1+\alpha_2-1} \left(p_1^{-1} + v\right)^{-(\alpha_1+\beta_1)} \left(p_2^{-1} + v\right)^{-(\alpha_2+\beta_2)} dv \\
&= c p_1^{-(\alpha_1+\alpha_2)} B(\alpha_1+\alpha_2, \beta_1+\beta_2) \\
&\quad \times {}_2F_1\left(\alpha_2+\beta_2, \alpha_1+\alpha_2; \alpha_1+\alpha_2+\beta_1+\beta_2; 1 - \frac{p_2}{p_1}\right) \qquad (6)
\end{aligned}$$

(see Erdélyi et al. 1954b, p. 233). Using (5) and (6), we arrive at (4).

Remark
1. The density of $T = \frac{X_1}{X_2} = [U_c(1-U_c)]^{\frac{1}{c}}$ utilizing (4) becomes $\frac{ct^{c-1}}{(1+t^c)^2} f\left(\frac{t^c}{1+t^c}\right)$. It involves the Gauss hypergeometric function, instead of Appell's function, as derived by Pham-Gia and Turkkan (2002, Eq. (10)).

2. For given independent gamma variables Y_i, $i = 1, \ldots, 4$, and U_c as defined in the above theorem, we have

$$U_c = \frac{X_1^c}{X_1^c + X_2^c} = \frac{\left(\frac{Y_1}{Y_2}\right)^c}{\left(\frac{Y_1}{Y_2}\right)^c + \left(\frac{Y_3}{Y_4}\right)^c} = \frac{(Y_1 Y_4)^c}{(Y_1 Y_4)^c + (Y_2 Y_3)^c},$$

$Z_1 = Y_1 Y_4$ has the density:

$$f(z; \alpha_1, \alpha_4; \gamma_1, \gamma_4) = \frac{\gamma_1 \gamma_4}{\Gamma(\alpha_1)\Gamma(\alpha_4)} G_{0,2}^{2,0}\left(z\gamma_1\gamma_4 \,|\alpha_1 - 1, \alpha_4 - 1\right), z > 0 \qquad (7)$$

The density of $Z_2 = Y_2 Y_3$ possesses a similar expression, and hence, U_c also has the density of the ratio of the c-th powers of products of independent gamma variables (Springer and Thompson 1970). On the other hand, (7) shows that U_c is the ratio of random variables whose densities are defined by $G_{0,2}^{2,0}$ functions.

The next theorem shows that the distribution of $W = X_1^c / \left(X_1^c + X_2^c\right)$, where X_i, $i = 1, 2$, are *gamma* variables with *same scale* parameter

($X_i \sim gamma\,(\alpha_i, \beta)\,, i = 1, 2$), derived by Nadarajah and Kotz (2004) can be obtained as limiting case from (4). However, Eq. (7) in their paper should be:

$$f(w) = \frac{w^{-(1+\alpha_2/c)}\,(1-w)^{\alpha_2/c-1}}{\alpha_2 c B\,(\alpha_1, \alpha_2)} \left\{ \alpha_2\, {}_2F_1\left(\alpha_2, \alpha_1+\alpha_2; \alpha_2+1; -\left(\frac{1-w}{w}\right)^{1/c}\right) \right.$$
$$- \left(\frac{1-w}{w}\right)^{1/c} \frac{\alpha_2\,(\alpha_1+\alpha_2)}{(\alpha_2+1)}$$
$$\left. \times\, {}_2F_1\left(\alpha_2+1, \alpha_1+\alpha_2+1; \alpha_2+2; -\left(\frac{1-w}{w}\right)^{1/c}\right)\right\}$$

Using relation (9.2.12) in Lebedev (1965, p. 243), the above density reduces to

$$f(w) = \frac{1}{cB\,(\alpha_1, \alpha_2)} w^{-\alpha_2/c-1}\,(1-w)^{\alpha_2/c-1} \left(1 + \left(\frac{1}{w} - 1\right)^{1/c}\right)^{-(\alpha_1+\alpha_2)}, \quad 0 < w < 1 \tag{8}$$

Note the absence of parameter β in this equation.

Theorem 2 *Let $X_i \sim GBP\left(\alpha_i, \beta, \frac{\lambda}{\beta}\right)$, $i = 1, 2$, be independent. Then $U_c = \frac{X_1^c}{X_1^c + X_2^c}$ where $c > 0$ is a constant, has as density (8) if $\beta \to \infty$.*

Proof From (4) the density of U_c for this special case is:

$$f(u) = \frac{B\,(\alpha_1+\alpha_2, 2\beta)}{cB\,(\alpha_1, \beta)\,B\,(\alpha_2, \beta)}\,(1-u)^{(\alpha_2/c)-1}\,u^{-(1+\alpha_2/c)}$$
$$\times {}_2F_1\left[\alpha_2+\beta, \alpha_1+\alpha_2; \alpha_1+\alpha_2+2\beta; 1 - \left(\frac{1-u}{u}\right)^{1/c}\right], \quad 0 < u < 1 \tag{9}$$

Note the absence of the parameter λ in (9). For large values of β,

$$\frac{\Gamma\,(2\beta)\,\Gamma\,(\alpha_1+\beta)\,\Gamma\,(\alpha_2+\beta)}{\Gamma\,(\beta)\,\Gamma\,(\beta)\,\Gamma\,(\alpha_1+\alpha_2+2\beta)} = 2^{-(\alpha_1+\alpha_2)} \tag{10}$$

Furthermore,

$$\lim_{\beta\to\infty} {}_2F_1\left[\alpha_2+\beta,\alpha_1+\alpha_2;\alpha_1+\alpha_2+2\beta;1-\left(\frac{1-u}{u}\right)^{1/c}\right]$$
$$=\lim_{\beta\to\infty}\sum_{k=0}^{\infty}\frac{\left(1+\frac{k-1}{\alpha_2+\beta}\right)\cdots\left(1+\frac{k-k}{\alpha_2+\beta}\right)\left(1-\left(\frac{1-u}{u}\right)^{1/c}\right)^k}{\left(1+\frac{k-1}{\alpha_1+\alpha_2+2\beta}\right)\cdots\left(1+\frac{k-k}{\alpha_1+\alpha_2+2\beta}\right)}\frac{(\alpha_1+\alpha_2)_k(\alpha_2+\beta)^k}{k!(\alpha_1+\alpha_2+2\beta)^k}$$
$$=2^{\alpha_1+\alpha_2}\left(1+\left(\frac{1}{u}-1\right)^{\frac{1}{c}}\right)^{-(\alpha_1+\alpha_2)} \tag{11}$$

From (9) and using (10) and (11), the result follows.

Remark From Theorem 2 follows that $T_c=\left(\frac{X_1}{X_2}\right)^c$ has as limiting case, when $\beta\to\infty$, the density:

$$f(t)=\frac{1}{cB(\alpha_1,\alpha_2)}t^{-\left(\frac{\alpha_2}{c}+1\right)}\left(1+t^{-\frac{1}{c}}\right)^{-(\alpha_1+\alpha_2)},\qquad 0<t<\infty.$$

Theorem 3 *The cumulative distribution function of U_c with pdf (4) is*

$$F(u)=1-K_2\left\{\frac{\lambda_2}{\lambda_1}\left(\frac{1-u}{u}\right)^{1/c}\right\}^{\alpha_2}$$
$$\times G_{3,3}^{2,3}\left(\frac{\lambda_2}{\lambda_1}\left(\frac{1-u}{u}\right)^{1/c}\middle|\begin{array}{ccc}1-\alpha_2, & 1-\alpha_2-\beta_2, & 1-\alpha_1-\alpha_2\\ 0, & \beta_1-\alpha_2, & -\alpha_2\end{array}\right), \tag{12}$$

where $K_2^{-1}=\prod_{i=1}^{2}B(\alpha_i,\beta_i)\,\Gamma(\alpha_i+\beta_i)$.

Proof We have:

$$F(u)=P\left(\frac{X_1^c}{X_1^c+X_2^c}\le u\right)$$
$$=P\left(X_2\ge\left(\frac{1-u}{u}\right)^{1/c}X_1\right)$$
$$=1-E_{X_1}\left\{P\left(X_2\le\left(\frac{1-u}{u}\right)^{1/c}x_1\,|X_1=x_1\right)\right\}$$
$$=1-\frac{\lambda_1^{\alpha_1}}{\prod_{i=1}^{2}B(\alpha_i,\beta_i)}\int_0^{\infty}B_{\frac{x_1\lambda_2\left(\frac{1-u}{u}\right)^{1/c}}{\left[1+x_1\lambda_2\left(\frac{1-u}{u}\right)^{1/c}\right]}}(\alpha_2,\beta_2)\,x_1^{\alpha_1-1}(1+\lambda_1x_1)^{-(\alpha_1+\beta_1)}dx_1$$

$$= 1 - \frac{\lambda_1^{\alpha_1}}{\alpha_2 \prod_{i=1}^{2} B(\alpha_i, \beta_i)} \int_0^\infty x_1^{\alpha_1 - 1} (1 + \lambda_1 x_1)^{-(\alpha_1+\beta_1)} \left\{ \frac{x_1 \lambda_2 \left(\frac{1-u}{u}\right)^{1/c}}{1 + x_1 \lambda_2 \left(\frac{1-u}{u}\right)^{1/c}} \right\}^{\alpha_2}$$

$$\times {}_2F_1\left(\alpha_2, 1 - \beta_2; \alpha_2 + 1; \frac{x_1 \lambda_2 \left(\frac{1-u}{u}\right)^{1/c}}{1 + x_1 \lambda_2 \left(\frac{1-u}{u}\right)^{1/c}}\right) dx_1 \tag{13}$$

by using relation (3). Now, using Kummer relation (Erdélyi et al. 1953, Eq. (2), p. 105), Eq. (13) can be written as

$$F(u) = 1 - \frac{1}{\alpha_2 \prod_{i=1}^{2} B(\alpha_i, \beta_i)} \left\{ \lambda_1 \lambda_2 \left(\frac{1-u}{u}\right)^{1/c} \right\}^{\alpha_2} \int_0^\infty x_1^{\alpha_1+\alpha_2-1} (1+\lambda_1 x_1)^{-(\alpha_1+\beta_1)}$$

$$\times {}_2F_1\left(\alpha_2, \alpha_2 + \beta_2; \alpha_2 + 1; -x_1 \lambda_2 \left(\frac{1-u}{u}\right)^{1/c}\right) dx_1 \tag{14}$$

If we now express Gauss hypergeometric function in terms of Meijer's G-function, as given by (2), and use the expression of the Stieltjes transform of a G-function (Mathai and Saxena 1973, p. 86), (14) becomes expression (12).

Alternatively the cdf of U_c can be obtained in terms of incomplete beta functions as illustrated in the corollary below.

Corollary 1

$$F(u) = K_1 \int_0^u (1-t)^{\frac{\alpha_2}{c} - 1} t^{-\left(1 + \frac{\alpha_2}{c}\right)}$$

$$\times {}_2F_1\left(\alpha_2 + \beta_2, \alpha_1 + \alpha_2; \alpha_1 + \alpha_2 + \beta_1 + \beta_2; 1 - \frac{\lambda_2}{\lambda_1}\left(\frac{1-t}{t}\right)^{1/c}\right) dt$$

$$= K_1 \sum_{k=0}^{\infty} \frac{(\alpha_2 + \beta_2)_k (\alpha_1 + \alpha_2)_k}{(\alpha_1 + \alpha_2 + \beta_1 + \beta_2)_k \, k!} \int_0^u (1-t)^{\frac{\alpha_2}{c} - 1} t^{-\left(1+\frac{\alpha_2}{c}\right)} \left(1 - \frac{\lambda_2}{\lambda_1}\left(\frac{1-t}{t}\right)^{1/c}\right)^k dt$$

where K_1 is defined as before.

Let $z = \frac{\lambda_2}{\lambda_1}\left(\frac{1-t}{t}\right)^{1/c}$, then

$$F(u) = 1 - K_1 \sum_{k=0}^{\infty} c \frac{(\alpha_2 + \beta_2)_k (\alpha_1 + \alpha_2)_k}{(\alpha_1 + \alpha_2 + \beta_1 + \beta_2)_k \, k!} \left(\frac{\lambda_2}{\lambda_1}\right)^{\alpha_2} \int_0^{\frac{\lambda_2}{\lambda_1}\left(\frac{1-u}{u}\right)^{1/c}} z^{\alpha_2} (1-z)^k \, dz$$

$$= 1 - \frac{B(\alpha_1 + \alpha_2, \beta_1 + \beta_2)}{B(\alpha_1, \beta_1) B(\alpha_2, \beta_2)} \sum_{k=0}^{\infty} \frac{(\alpha_2 + \beta_2)_k (\alpha_1 + \alpha_2)_k}{(\alpha_1 + \alpha_2 + \beta_1 + \beta_2)_k \, k!} \frac{B_{\frac{\lambda_2}{\lambda_1}\left(\frac{1-u}{u}\right)^{1/c}}(\alpha_2, k+1)}{B(\alpha_2, k+1)} \tag{15}$$

Expressing the cdf in terms of incomplete beta functions, (15) might appear to be less cumbersome than Eq. (12).

In the next theorem an expression for the hazard function of U_c, $H(u) = \frac{f(u)}{1-F(u)}$, is given. Its inverse $[H(u)]^{-1}$ is the Mills’ ratio, used in economics, to measure the ratio of the tail area of the distribution to its bounding ordinate $f(u)$.

Theorem 4 *Let U_c has pdf (4) then the hazard function is given by*

$$H(u) = \alpha_2 \left(cu(1-u)\right)^{-1} \times \left\{ 1 + K_3 \frac{G^{2,3}_{3,3}\left(\frac{\lambda_2}{\lambda_1}\left(\frac{1-u}{u}\right)^{1/c} \middle| \begin{matrix} 1-\alpha_2, & 1-\alpha_2-\beta_2, & 1-\alpha_1-\alpha_2 \\ 1, & \beta_1-\alpha_2, & -\alpha_2 \end{matrix} \right)}{{}_2F_1\left[\alpha_2+\beta_2, \alpha_1+\alpha_2; \alpha_1+\alpha_2+\beta_1+\beta_2; 1-\frac{\lambda_2}{\lambda_1}\left(\frac{1-u}{u}\right)^{1/c}\right]} \right\}^{-1} \tag{16}$$

where $K_3^{-1} = B(\alpha_1+\alpha_2, \beta_1+\beta_2) \prod_{i=1}^{2} \Gamma(\alpha_i+\beta_i)$.

Proof Using (4) and (12), expression (16) follows immediately. The hazard function in this case can also be expressed in terms of incomplete beta functions, by utilizing (4) and (15).

The moments of U_c can be expressed as values of G-functions, as shown below.

Theorem 5 *Let U_c has pdf (4) then*

$$E(U_c^h) = \frac{K_2}{\Gamma(h)} \left(\frac{\lambda_2}{\lambda_1}\right)^{\alpha_2} (2\pi)^{2(1-c)} c^{\gamma-2} \times G^{2c+1,2c+1}_{2c+1,2c+1}\left(\left(\frac{\lambda_2}{\lambda_1}\right)^c \middle| \begin{matrix} \Delta(c, 1-\alpha_2-\beta_2), & \Delta(c, 1-\alpha_1-\alpha_2), & 1-\alpha_2/c \\ \Delta(c, 0), & \Delta(c, \beta_1-\alpha_2), & h-\alpha_2/c \end{matrix} \right)$$

where $\Delta(n, \zeta)$ represent the set of parameters $\left(\frac{\zeta}{n}, \frac{\zeta+1}{n}, \ldots, \frac{\zeta+n-1}{n}\right)$ For $c = 1$, we have:

$$E(U^h) = \frac{1}{\left(\prod_{i=1}^{2} \Gamma(\alpha_i)\Gamma(\beta_i)\right)\Gamma(h)} \times \left(\frac{\lambda_2}{\lambda_1}\right)^{\alpha_2} G^{3,3}_{3,3}\left(\frac{\lambda_2}{\lambda_1} \middle| \begin{matrix} 1-\alpha_2-\beta_2, & 1-\alpha_1-\alpha_2, & 1-\alpha_2 \\ 0, & \beta_1-\alpha_2, & h-\alpha_2 \end{matrix} \right)$$

Proof From (4) we can write the h-th moment about the origin, $h \geq 1$, as

$$E(U_c^h) = K_1 \int_0^1 u^{h-\alpha_2/c-1} (1-u)^{\alpha_2/c-1} \times {}_2F_1\left[\alpha_2+\beta_2, \alpha_1+\alpha_2; \alpha_1+\alpha_2+\beta_1+\beta_2; 1-\frac{\lambda_2}{\lambda_1}\left(\frac{1-u}{u}\right)^{1/c}\right] du \tag{17}$$

with K_1 as defined as before. Setting $v = \frac{(1-u)}{u}$, (17) can be written as

$$E(U_c^h) = K_1 \int_0^\infty v^{\alpha_2/c-1} (1+v)^{-h} \times {}_2F_1 \left[\alpha_2 + \beta_2, \alpha_1 + \alpha_2; \alpha_1 + \alpha_2 + \beta_1 + \beta_2; 1 - \frac{\lambda_2}{\lambda_1} v^{1/c} \right] dv$$

Using relation (2), and the relation $(1+v)^{-h} = \frac{1}{\Gamma(h)} G_{1,1}^{1,1}\left(v \middle| \begin{matrix} 1-h \\ 0 \end{matrix} \right)$, the above equation becomes:

$$E(U_c^h) = \frac{K_2}{c\Gamma(h)} \left(\frac{\lambda_2}{\lambda_1} \right)^{\alpha_2} \int_0^\infty v^{\alpha_2/c-1} G_{1,1}^{1,1}\left(v \middle| \begin{matrix} 1-h \\ 0 \end{matrix} \right) \times G_{2,2}^{2,2}\left(\frac{\lambda_2}{\lambda_1} v^{1/c} \middle| \begin{matrix} 1-\alpha_2-\beta_2, & 1-\alpha_1-\alpha_2 \\ 0, & \beta_1-\alpha_2 \end{matrix} \right) dv$$

where K_2 is defined as before.

Using the expression of the Mellin transform of the product of two G-functions (see Mathai and Saxena 1973 p. 80), when c is an integer, the result follows. For $c = 1$, the equation reduces easily to a simpler expression.

Alternatively $E(U_c^h)$ can be expressed as follows:

Corollary 2 *From* (17)*, let* $z = \frac{\lambda_2}{\lambda_1} \left(\frac{1-u}{u} \right)^{\frac{1}{c}}$*, then*

$$\begin{aligned} E(U_c^h) &= K_1 \left(\frac{\lambda_1}{\lambda_2} \right)^{\alpha_2} \sum_{k=0}^\infty c \frac{(\alpha_2+\beta_2)_k (\alpha_1+\alpha_2)_k}{(\alpha_1+\alpha_2+\beta_1+\beta_2)_k\, k!} \int_0^\infty z^{\alpha_2-1} \left(1 + \left(\frac{\lambda_1}{\lambda_2} z \right)^c \right)^{-h} (1-z)^k \, dz \\ &= K_1 \left(\frac{\lambda_1}{\lambda_2} \right)^{\alpha_2} \sum_{k=0}^\infty cB(k+1, \alpha_2)\, {}_2F_1 \left(h, \alpha_2; \alpha_2+k+1; -\left(\frac{\lambda_1}{\lambda_2} \right)^c \right) \\ &\quad \times \frac{(\alpha_2+\beta_2)_k (\alpha_1+\alpha_2)_k}{(\alpha_1+\alpha_2+\beta_1+\beta_2)_k\, k!} \end{aligned}$$

where K_1 *is defined as before* (*see* Erdélyi et al. 1954a, p. 337).

4 Discussion

Thanks to the presence of several parameters, density (4) can take a wide variety of shapes: U-shapes, unimodal shapes, concave curves, as can be seen from Figs. 1, 2, 3, 4, 5, which illustrates some of the shapes of density (4) for selected values of $(\alpha_i, \beta_i, \lambda_i)$, $i = 1, 2$, and c. U_c, as defined before, is composed out of generalized

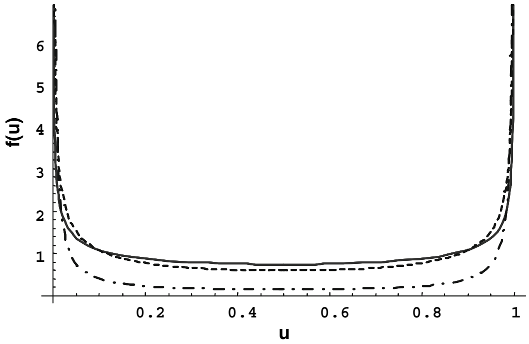


Fig. 1 Pdf (4) for $(\alpha_1, \alpha_2) = (0.5, 0.5)$ and $(\beta_1, \beta_2, \lambda_1, \lambda_2) = (4, 4, 1, 1)$ with $c = 0.8$ (*solid line*), $c = 1$ (*dashed line*) and $c = 4$ (*dot-dashed line*)

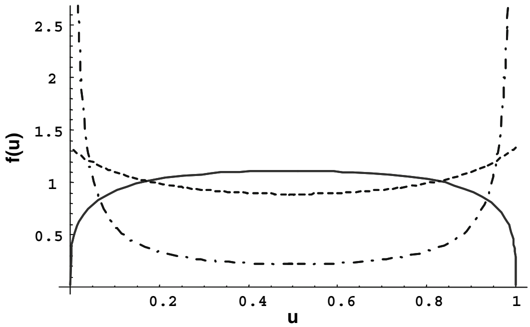


Fig. 2 Pdf (4) for $(\alpha_1, \alpha_2) = (1, 1)$ and $(\beta_1, \beta_2, \lambda_1, \lambda_2) = (4, 4, 1, 1)$ with $c = 0.8$ (*solid line*), $c = 1$ (*dashed line*) and $c = 4$ (*dot-dashed line*)

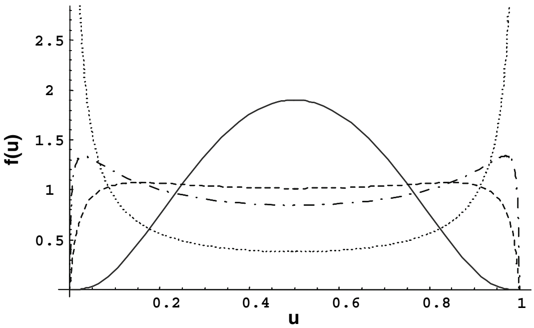


Fig. 3 Pdf (4) for $(\alpha_1, \alpha_2) = (4, 4)$ and $(\beta_1, \beta_2, \lambda_1, \lambda_2) = (4, 4, 1, 1)$ with $c = 0.8$ (*solid line*), $c = 1$ (*dashed line*), $c = 1.8$ (*dot-dashed line*) and $c = 4$ (*dotted line*)

beta-prime variables with more parameters than gamma distributed variables, therefore as expected this parameter-rich density (4) has more flexibility than density (8), for example the bimodal form in Fig. 3 (see also Theorem 2). The effect of different values of λ_i and β_i, $i = 1, 2$, is also illustrated in Figs. 4 and 5, for example the skewed U-shape form in Fig. 4. The estimation of the parameters of the density (4) will be dealt with in a subsequent paper.

Figure 6 shows density (9) when $X_i \sim GBP\left(\alpha_i, \beta, \frac{\lambda}{\beta}\right)$, $i = 1, 2$, for different values of β, as well as density (8) if $\alpha_1 = \alpha_2 = 4$ and $c = 1$. As β increases, the graph of pdf (9) tends to the graph of pdf (8), as proved in Theorem 2.

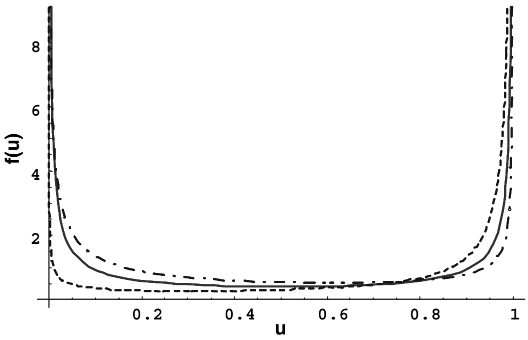


Fig. 4 Pdf (4) for $(\alpha_1, \alpha_2) = (0.5, 0.5)$ and $(\lambda_1, \lambda_2, c) = (1, 1, 1)$ with $(\beta_1, \beta_2) = (0.5, 0.5)$ (*solid line*), $(\beta_1, \beta_2) = (0.5, 4)$ (*dashed line*), and $(\beta_1, \beta_2) = (4, 2)$ (*dot-dashed line*)

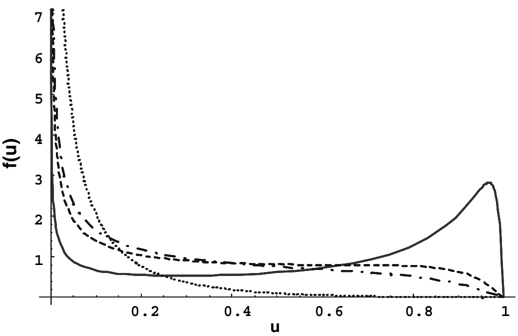


Fig. 5 Pdf (4) for $(\alpha_1, \alpha_2) = (0.5, 4)$ and $(\beta_1, \beta_2, c) = (4, 4, 1)$ with $(\lambda_1, \lambda_2) = (0.1, 4)$ (*solid line*), $(\lambda_1, \lambda_2) = (0.5, 4)$ (*dashed line*), $(\lambda_1, \lambda_2) = (0.8, 4)$ (*dot-dashed line*) and $(\lambda_1, \lambda_2) = (4, 2)$ (*dotted line*)

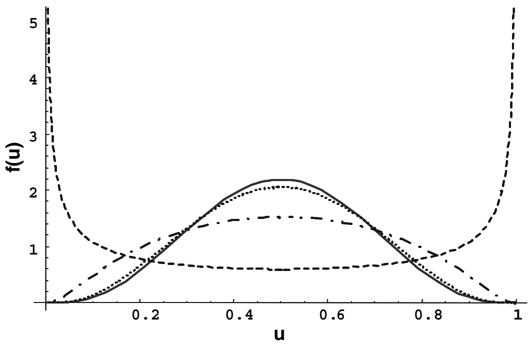


Fig. 6 Pdf (8) (*solid line*); Pdf (9) with $\beta = 0.5$ (*dashed line*), $\beta = 4$ (*dot-dashed line*) and $\beta = 4$ (*dotted line*)

The following applications, the first in Bayesian statistics and the second to economics, show the potential of further uses of this distribution in other applied fields.

Bayesian statistics: Predictive distribution of the ratio of independent sample variances from a normal population

Let us consider $T_1 \sim N(\mu_1, \lambda_1)$, where both the mean μ_1 and the precision (inverse of the variance) λ_1 are unknown. Let (μ_1, λ_1) have as prior the normal-gamma distribution, denoted by $(\mu_1, \lambda_1) \sim Ng(\mu_1, \lambda_1; A, B, \alpha_1, \beta_1)$, i.e., $\mu_1 \sim N(A, B\lambda_1)$ and $\lambda_1 \sim Ga(\alpha_1, \beta_1)$with B being a positive constant. Similarly, let $T_2 \sim N(\mu_2, \lambda_2)$ with $(\mu_2, \lambda_2) \sim Ng(\mu_2, \lambda_2; C, D, \alpha_2, \beta_2)$and let T_1 and T_2 be independent.

Let s_1^2 and s_2^2 be two independent sample variances with sizes n_1 and n_2 respectively, taken from the two populations. A result in Bayesian statistics states that, under normal sampling, the predictive density of $Y_1 = n_1 s_1^2 = \sum_{i=1}^{n} (x_i - \overline{x})^2$ is a Gamma–gamma distribution, i.e., $Gg\left(\frac{1}{2}(n_1 - 1), \alpha_1, 2\beta_1\right)$ (Bernardo and Smith 1994), and is hence independent of A and B. Equivalently, $Y_1 \sim GBP\left(\frac{1}{2}(n_1 - 1), \alpha_1, \frac{1}{2\beta_1}\right)$, as pointed out in Pham-Gia and Turkkan (2002), the Gamma–gamma distribution is a reparametrized form of the generalized beta prime. Similar results hold for $Y_2 = n_2 s_2^2$. Their ratio $U_c = \frac{\left(n_1 s_1^2\right)^c}{\left(n_1 s_1^2\right)^c + \left(n_2 s_2^2\right)^c}$ then has the density given by (4), and for $c = 1$, it is the predictive density of the ratio of two random sample variances, coming respectively from the two populations. If $n_1 = n_2 = n$ and $c = 1/2$, we then have the predictive density of the ratio of two independent sample standard deviations, $V = \frac{s_1}{s_1 + s_2}$, as

$$f(v) = \frac{B(n-1, 2\alpha)}{\left(B\left(\frac{n-1}{2}, \alpha\right)\right)^2} \frac{(1-v)^{\frac{n-3}{2}}}{v^{\frac{n+1}{2}}} \,_2F_1\left(2\alpha, n-1; n-1+2\alpha; \frac{2\nu - 1}{\nu^2}\right),$$
$$0 < v < 1.$$

Economics: Relative value of the unreported income part in a national income distribution

An important subject encountered in economics is the study of the distribution of individual incomes on a national basis, and several common distributions, such as the Gamma, Beta, Lognormal, Pareto, and some more specialized ones, such as the Singh-Maddala, and its generalization, as given in McDonald and Xu (1995), are used as models.

A specific problem considered in Ransom and Cramer (1983) is the presence of disturbances in the reporting of incomes, which lead to errors and under or over-reporting. All these disturbances are considered together there, as forming a normal random variable, resulting in a total income distribution of $X_1 + X_2$ where X_1 is normal and X_2 is the reported income, distributed as either Gamma, Lognormal or Pareto. In a similar way, but using positive skewed distributions, which are more appropriate for their problem, Pham-Gia and Turkkan (1997) considered the case of X_1 representing all unreported individual incomes, including those from the underground economy, and fitted gamma distributions to X_1 and X_2, resulting in a convenient closed form expression for the density of $X_1 + X_2$. The 1988 Canadian income distribution (Pham-Gia and Turkkan 1997) was used as an example, and based on the recognition of the fact that the values reported for large incomes, and affecting the distribution tail, are not reliable, the beta model was also used, as in Ryscavage (1989).

For this problem, let X_1 represent the unreported income, and let both X_1 and X_2 be GBP random variables. Hence, $U = X_1/(X_1 + X_2)$ represents the relative value of the unreported income, with respect to the total income, and its distribution should be very informative for the understanding of how this unreported part could influence the whole distribution. First, let X_2 be represented, adequately, by a GBP variable. For

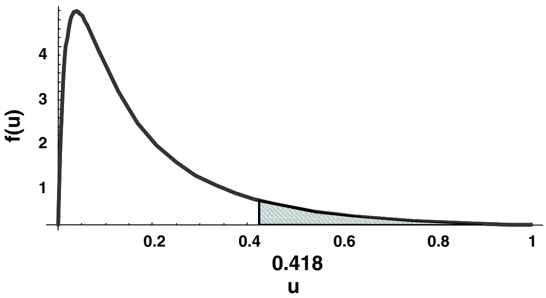


Fig. 7 Pdf (4) for $c = 1$, $\alpha_1 = \alpha_2 = 6.15$, $\beta_1 = \beta_2 = 2.12$, $\lambda_1 = 7.87$, $\lambda_2 = 1.185$

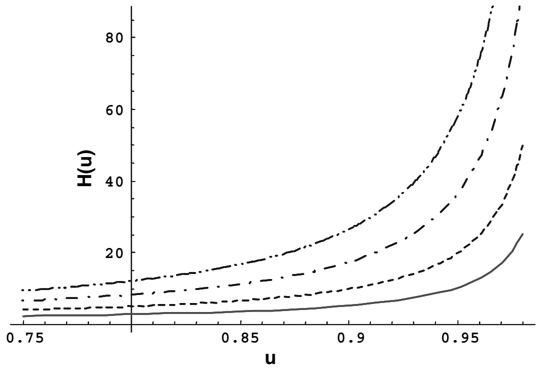


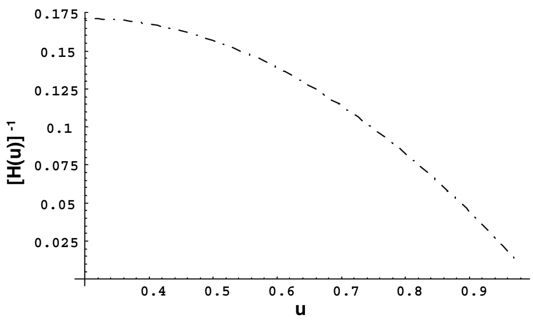
Fig. 8 The hazard function (16) for $\beta = 0.5$ (*solid line*), $\beta = 1$, (*dashed line*), $\beta = 2$ (*dot-dashed line*), $\beta = 4$ (*dot-dot-dashed line*)

example, let $X_2 \sim GBP$ (6.15, 2.12, 1.18) with the values of the parameters obtained by the moments method, as presented in Pham-Gia and Duong (1989). For the unreported income X_1, estimated as 15% of X_2, it can be represented, quite adequately also, by the distribution GBP (6.15, 2.12, 7.87), so that $\mu_{X_1} \approx 0.15\mu_{X_2}$. It should be clearly stated that since very little is known about X_1, and its shape is certainly unknown, this hypothesis is only one among several possible candidates. However, due to the versatility of the GBP distribution, the closed form of the density of U, and the variety of shapes it can have, we have a very convenient approach here, unlike the one used by Ransom and Cramer (1983).

By taking $c = 1$, we have the distribution of U, the relative magnitude of X_1, as given by (4). We can also study this ratio according to different scenarios related to the shape of the density of X_1, by changing the values of the three parameters of the latter, while maintaining its mean fixed. In Fig. 7, we have $P\ (U \geq 0.418) = 0.10$, which means that, under the hypotheses made, it is in 10 % of the cases only that we have the relative value of the unreported income exceeding 41,8% of the total income. Similarly, by taking $X_1 \sim GBP$ (1.23, 2.12, 1.57), for example, we have now, $P\ (U \geq 0.426) = 0.10$, which has a similar interpretation. The ratio $X_1^c / \left(X_1^c + X_2^c\right)$ has a similar meaning as before, but is now related to the power c of the reported, and unreported, incomes and has its density given by (4).

Figure 8 shows the hazard function for $\lambda_1 = \lambda_2, \alpha_1 = 2, \alpha_2 = 4, c = 1$, and $\beta_1 = \beta_2 = \beta$, for different values of β. For the case of our distribution of the 1988

Fig. 9 Mills's ratio for $c = 1$, $\alpha_1 = \alpha_2 = 6.15$, $\beta_1 = \beta_2 = 2.12$, $\lambda_1 = 1.18$, $\lambda_2 = 7.87$



Canadian income the, Mills's ratio $[H(u)]^{-1}$of U is given by Fig. 9, obtained directly from (16), with appropriate values given to the parameters.

Acknowledgments This work is based upon research supported by the National Research Foundation, South Africa.

References

Amaral-Turkman MA, Dunsmore IR (1985) Measures of information in the predictive distribution. In: Bernardo, et al (ed) Bayesian Statistics, vol 2. Elsevier, New York, pp 603–612

Bekker A (1990) Generalising, compounding and characterising as methods to obtain parameter-rich distributions (In Afrikaans). Ph.D. thesis, University of South Africa, Pretoria

Bekker A, Roux JJJ, Mostert PJ (2000) A generalization of the compound Rayleigh distribution: using Bayesian methods on survival data. Communications in Statistics: Theory and Methods 29(7):1419–1433

Bernardo JM, Smith AFM (1994) Bayesian statistics. Wiley, New York

Dyer D (1982) The convolution of generalized F distributions. J Am Stat Ass 77:184–189

Erdélyi A et al (1953) Higher transcendental functions, vol 1. McGraw-Hill, New York

Erdélyi A et al (1954a) Tables of integral transforms, vol 1. McGraw-Hill, New York

Erdélyi A et al (1954b) Tables of integral transforms, vol 2. McGraw-Hill, New York

Gradshteyn IS, Ryzhik IM (2000) Table of integrals, series, and products. Academic, New York

Gupta AK, Nadarajah S (2004) Handbook of Beta distribution and its application. Marcel Dekker, New York

Lebedev NN (1965) Special functions and their applications. Prentice-Hall, London

Malik HJ (1967) Exact distribution of the quotient of independent generalized gamma variables. Can Math Bull 10:463–465

Mathai AM (1993) A handbook of generalized special functions for statistical and physical sciences. Clarendon Press, London

Mathai AM, Moschopoulos PG (1997) A multivariate inverted beta model. Statistica LVII(2):189–197

Mathai AM, Saxena RK (1966) On a generalized hypergeometric distribution. Metrika 11(2):127–132

Mathai AM, Saxena RK (1973) Lecture notes in mathematics: generalized hypergeometric functions with applications in statistics and physical sciences. Springer, Berlin

McDonald JB, Xu YJ (1995) A generalization of the beta distribution with applications. J Econometr 66:133–152

Nadarajah S, Kotz S (2004) A generalized beta distribution. Math Scientist 29:36–41

Pham-Gia T, Duong QP (1989) The generalized beta and F-distributions in statistical modelling. Math Comput Modell 12:1613–1625

Pham-Gia T, Turkkan N (1997) Change in income distribution in the presence of reporting errors. Math Comput Modell 25(7):33–42

Pham-Gia T, Turkkan N (2002) Operations on the generalized F variables and applications. Statistics 36(3):195–209

Ransom MR, Cramer JS (1983) Income distribution functions with disturbances. Eur Econ Rev 22:363–372

Ryscavage P (1989) The underground economy in the US: What we know, what we don't know, and what we are doing to find out. Proceedings of the 47th session of the international statistical institute of Paris, pp 73–92

Sarhan A (1995) Bayesian estimation of service factors of repairable system components. Advances in Modelling and Analysis, C, AMSE Press 46(3):47–53

Springer MD, Thompson WE (1970) The distribution of products of beta, gamma and Gaussian random variables. SIAM J Appl Math 18(4):721–737