Journal of Statistics: Advances in Theory and Applications
Volume 10, Number 2, 2013, Pages 153-187

DIFFERENCES, RATIOS AND PRODUCTS FOR "CONNOR AND MOSIMANN'S GENERALIZED DIRICHLET DISTRIBUTION"

N. A. MOKHLIS[1], M. M. NASSAR[1] and S. A. EL-WARRAD[2]

[1]Department of Mathematics
Faculty of Science
Ain Shams University
Cairo
Egypt
e-mail: m_nassar_1999@yahoo.com

[2]Department of Mathematics
Faculty of Science
Benghazi University
Al-Marj
Libya

Abstract

In this paper, we derive the exact distributions of the differences, ratios and products and the corresponding moments properties of two dependent random variables X and Y following the Connor and Mosimann's generalized Dirichlet distribution. Using the distributions of the differences and the ratios, we obtain the stress-strength reliability function $\Pr(X < Y)$. A simple way for generation of (X, Y) is introduced depending on the marginal distribution of X and the conditional distribution of Y given $X = x$. Moreover, an application of the results to real data is provided.

2010 Mathematics Subject Classification: 60E05, 62E15.

Keywords and phrases: Connor and Mosimann's generalized Dirichlet distribution, differences of random variables, ratios of random variables, products of random variables, hypergeometric function.

Received September 21, 2013

1. Introduction

The linear combinations, ratios and products of two random variables X and Y say have attracted many researchers in the statistics literature, X and Y may be independent or dependent. For independent variables, Pham-Gia and Turkkan [13, 14] derived the exact distribution of $X + Y$ and $X - Y$ of X and Y are independent beta random variables. Also, Pham-Gia [15] has derived distributions for X / Y and $X / (X + Y)$ of X and Y are independent beta random variables. For dependent variables, Nadarajah [10] has derived distributions for $X + Y$, $X / (X + Y)$, and XY for the bivariate Gumbel distribution, followed by Nadarajah and Kotz [12], who derived the distributions for $X + Y$ and $X / (X + Y)$ of Connor and Mosimann's generalized Dirichlet distribution. Also, Al-Ruzaiza and El-Gohary [2] have derived distributions for $X + Y$, $X / (X + Y)$, and XY of inverted bivariate beta distribution. Nadarajah [11] has also derived the reliability of some bivariate beta distribution.

The bivariate beta distribution is one of the basic distributions in statistics, as it attracts useful applications in several areas; for example, in the modelling of the proportions of substances in a mixture, brand shares, i.e., the proportions of brands of some consumer product that are bought by customers (Chatfiled [4]), proportions of the electorate voting for the candidate in a two candidate election (Hoyer and Mayer [9]), and the dependence between two soil strength parameters (A-Grivas and Asaoka [1]). Bivariate beta distributions have also been used extensively as a prior in Bayesian statistics (see, for example, Apostolakis [3]).

In this paper, we derive the exact distributions of $W = X - Y$, $R = X / Y$, and $P = XY$ and their corresponding moments, where X and Y are distributed according to the joint probability density function (pdf) given by

$$h(x,\ y) = \frac{1}{B(a,\ c)B(b,\ d)} x^{a-1} y^{b-1} (1-x)^{c-b-d} (1-x-y)^{d-1}, \tag{1}$$

for $x \geq 0,\ y \geq 0;\ x + y < 1,\ a > 0,\ b > 0,\ d > 0,$ and $c > 0.$

The distribution in (1) is known as the Connor and Mosimann's generalized Dirichlet distribution (see Connor and Mosimann [5]). It has several applications in many areas, including Bayesian statistics, contingency tables, correspondence analysis, environmental sciences, forensic sciences, geochemistry, image analysis, and statistical decision theory (see Gupta and Nadarajah [8] for illustrations of some of these application areas).

The calculations throughout this paper involve some special functions, including the Gauss hypergeometric function

$${}_2F_1(\delta,\ \lambda;\ \lambda + \mu;\ x) = \frac{1}{B(\lambda,\ \mu)} \int_0^1 u^{\lambda-1} (1-u)^{\mu-1} (1-xu)^{-\delta} du, \tag{2}$$

where $\operatorname{Re} \lambda > 0,\ \operatorname{Re} \mu > 0,$ which is given in a series form by

$${}_2F_1(\lambda,\ \mu;\ \delta;\ x) = \sum_{j=0}^{\infty} \frac{(\lambda)_j (\mu)_j}{(\delta)_j} \frac{x^j}{j!}, \tag{3}$$

where $|x| < 1,$ and $(f)_k = f(f+1)\ldots(f+k-1)$ denotes the ascending factorial, the Appell hypergeometric function of the first kind

$$F_1(\lambda,\ \delta,\ \gamma;\ \lambda + \mu;\ u,\ \nu) = \frac{1}{B(\lambda,\ \mu)} \int_0^1 x^{\lambda-1} (1-x)^{\mu-1} (1-ux)^{-\delta} (1-\nu x)^{-\gamma} dx, \tag{4}$$

where $\operatorname{Re} \lambda > 0,\ \operatorname{Re} \mu > 0,$ which is given in a series form by

$$F_1(\lambda,\ \delta,\ \gamma;\ \mu;\ u,\ \nu) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{(\lambda)_{i+j} (\delta)_i (\gamma)_j u^i \nu^j}{(\mu)_{i+j}\, i!\, j!}, \tag{5}$$

where $|u| < 1, |\nu| < 1,$ and

$$B_x(\lambda, \mu) = \int_0^x t^{\lambda-1}(1-t)^{\mu-1}dt, \tag{6}$$

which is given by

$$B_x(\lambda, \mu) = \frac{x^\lambda}{\lambda}\, {}_2F_1(\lambda, 1-\mu; \lambda+1; x). \tag{7}$$

The properties of the above special functions can be found in Gradshteyn and Ryzhik [6] p(286, 1039, 287, 1057, 950).

The rest of the paper is organized as follows: Section 2 deals with the derivation of the probability density functions (pdfs) and the cumulative distribution functions (cdfs) of $W = X - Y$, $T = X / Y$, and $P = XY$. The corresponding moments are discussed in Section 3. Section 4 introduces a simple way for generating observations from the bivariate distribution in (1), and hence observations from W, T, and P. Section 5 provides an application to compositional data of lavas from Skye (see Nadarajah and Kotz [12]).

2. The Probability Density Functions

In this section, we derive the exact probability density functions and the cumulative distribution functions of the differences, ratios, and products of two random variables having Connor and Mosimann's generalized Dirichlet joint distribution, given in (1).

Theorem 1. *If X and Y are jointly distributed random variables with the joint pdf* (1), *then the pdf and cdf of* $W = X - Y$ *are given, respectively, by*

$$g_W(w) = \begin{cases} \frac{\Gamma(a+c)\Gamma(b+d)}{2^a\,\Gamma(b)\Gamma(c)\Gamma(a+d)}(-w)^{b-1}(1+w)^{a+d-1} F_1\left(a, b+d-c, 1-b; a+d; \frac{1+w}{2}, \frac{1+w}{2w}\right), & if\ -1 < w \le 0, \\ \frac{B(a+b-1,d)}{2^{a+c-d-1}B(a,c)B(b,d)}\, {}_2F_1(b+d-c, d; a+b+d-1; -1), & if\ w = 0, \\ \frac{1}{2^b B(a,c)} w^{a-1}(1-w)^{c-1} F_1\left(b, b+d-c, 1-a; b+d; \frac{1}{2}, \frac{w-1}{2w}\right), & if\ 0 < w < 1, \\ 0, & otherwise, \end{cases} \tag{8}$$

where $a + b > 1$; *and*

$$G_W(w)=\begin{cases} 0, & \text{if } w\le -1,\\ \dfrac{\Gamma(a+c)\Gamma(b+d)}{2^a\Gamma(b)\Gamma(c)\Gamma(a+d)}\displaystyle\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{(a)_{m+n}(b+d-c)_m(1-b)_n}{(a+d)_{m+n}m!n!}\left(\frac{1}{2}\right)^{m+n}(-1)^n B_{1+w}(a+d+m+n,b-n), & \text{if } -1<w\le 0,\\ \dfrac{\Gamma(a+c)\Gamma(b+d)}{2^a\Gamma(b)\Gamma(c)\Gamma(a+d)}\displaystyle\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{(a)_{m+n}(b+d-c)_m(1-b)_n}{(a+d)_{m+n}m!n!}\left(\frac{1}{2}\right)^{m+n}(-1)^n B(a+d+m+n,b-n) & \\ +\dfrac{1}{2^b B(a,c)}\displaystyle\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{(b)_{m+n}(b+d-c)_m(1-a)_n}{(b+d)_{m+n}m!n!}\left(\frac{1}{2}\right)^{m+n}(-1)^n B_w(a-n,c+n), & \text{if } 0<w<1,\\ 1, & \text{if } w\ge 1. \end{cases} \tag{9}$$

Proof. Using (1), the joint pdf of $W = X - Y$ and Y is given by

$$f(w,\ y) = \frac{1}{B(a,\ c)B(b,\ d)}(w + y)^{a-1}y^{b-1}(1 - w - y)^{c-b-d}(1 - w - 2y)^{d-1},$$

$$-1 < w < 1,\ -w < y < \frac{1-w}{2}.$$

Thus, the marginal pdf of W is

$$g_W(w) = \int_{-\infty}^{\infty} f(w,\ y)dy, \qquad -1 < w < 1.$$

For $-1 < w \le 0$, the pdf of W is given by

$$g_W(w) = \frac{1}{B(a,\ c)B(b,\ d)}\int_{-w}^{\frac{1-w}{2}} y^{b-1}(w + y)^{a-1}(1 - w - y)^{c-b-d}(1 - w - 2y)^{d-1}dy$$

$$= \frac{1}{B(a,c)B(b,d)}(-w)^{b-1}(1+w)^{d-1}\int_0^1\left(\frac{1+w}{2}\nu\right)^{a-1}\left(1-\frac{1+w}{2}\nu\right)^{c-b-d}$$

$$\times\left(1-\frac{1+w}{2w}\nu\right)^{b-1}(1-\nu)^{d-1}\frac{1+w}{2}d\nu,$$

where the transformation $\nu = \frac{2(w + y)}{1 + w}$ has been applied. Solving the above integral using (4), we get the first part of Equation (8).

Now, for $w = 0$, we have

$$g_W(w) = \frac{1}{B(a, c)B(b, d)} \int_0^{\frac{1}{2}} y^{b+a-2}(1-y)^{c-b-d}(1-2y)^{d-1} dy$$

$$= \frac{1}{B(a, c)B(b, d)} \left(\frac{1}{2}\right)^{a+c-d-1} \int_0^1 \nu^{d-1}(1-\nu)^{a+b-2}(1+\nu)^{-(b+d-c)} d\nu.$$

Using the transformation $\nu = (1-2y)$ and (2), we obtain the second part of Equation (8). As for $0 < w < 1$, we have

$$g_W(w) = \frac{1}{B(a, c)B(b, d)} \int_0^{\frac{1-w}{2}} y^{b-1}(w+y)^{a-1}(1-w-y)^{c-b-d}(1-w-2y)^{d-1} dy$$

$$= \frac{1}{B(a, c)B(b, d)} \int_0^{\frac{1-w}{2}} y^{b-1}\left(w\left(1+\frac{y}{w}\right)\right)^{a-1}\left((1-w)\left[1-\frac{y}{1-w}\right]\right)^{c-b-d}$$

$$\times \left((1-w)\left[1-\frac{2y}{1-w}\right]\right)^{d-1} dy$$

$$= \frac{1}{2^b B(a, c)B(b, d)} w^{a-1}(1-w)^{c-1} \int_0^1 u^{b-1}(1-u)^{d-1}\left(1-\frac{u}{2}\right)^{-(b+d-c)}$$

$$\times \left(1-\frac{w-1}{2w}u\right)^{-(1-a)} du.$$

Using the transformation $u = \frac{2y}{1-w}$ and (4), we get the last part of Equation (8).

Integration of $g_W(w)$ in (8) with respect to W and using (5), leads to Equation (9), which completes the proof.

Remarks. (1) The pdf of $W' = Y - X = -W$ is given by

$$g_{W'}(w') = \begin{cases} \dfrac{1}{2^b B(a,c)}(-w')^{a-1}(1+w')^{c-1} F_1\left(b, b+d-c, 1-a; b+d; \dfrac{1}{2}, \dfrac{1+w'}{2w'}\right), & \text{if } -1 < w' < 0, \\ \dfrac{B(a+b-1,d)}{2^{a+c-d-1} B(a,c)B(b,d)}\, {}_2F_1(b+d-c, d; a+b+d-1; -1), & \text{if } w' = 0, \\ \dfrac{\Gamma(a+c)\Gamma(b+d)}{2^a \Gamma(b)\Gamma(c)\Gamma(a+d)} w'^{b-1}(1-w')^{a+d-1} F_1\left(a, b+d-c, 1-b; a+d; \dfrac{1-w'}{2}, \dfrac{w'-1}{2w'}\right), & \text{if } 0 < w' < 1, \\ 0, & \text{otherwise,} \end{cases}$$

where $a + b > 1$.

(2) The stress-strength reliability function $R = P(X < Y)$ obtained using (9) as

$$P(X < Y) = G_W(0) = \frac{\Gamma(a+c)\Gamma(b+d)}{2^a \Gamma(b)\Gamma(c)\Gamma(a+d)} \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \frac{(a)_{m+n}(b+d-c)_m(1-b)_n}{(a+d)_{m+n}\, m!\, n!}$$

$$\times \left(\frac{1}{2}\right)^{m+n} (-1)^n B(a+b+m+n, b-n). \qquad (10)$$

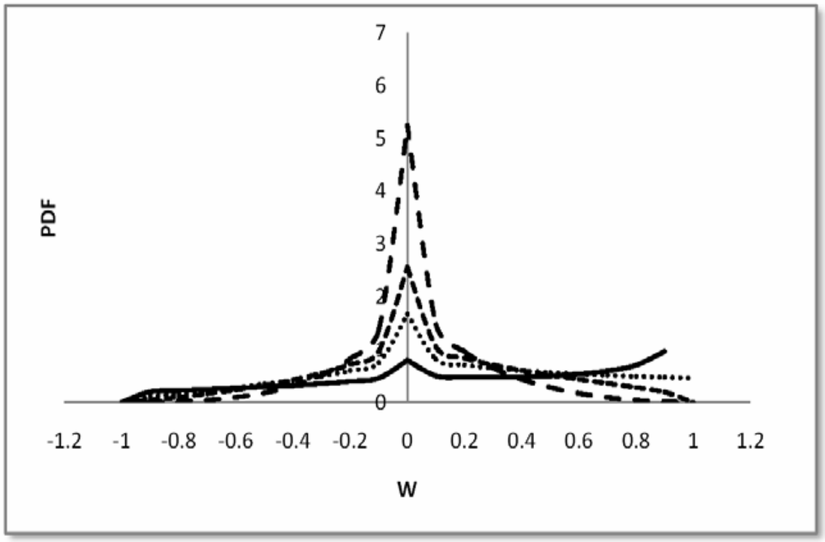
(3) See if $a = b = c = d = 1$, the pdfs of W and W' are, respectively, given by

$$g_W(w) = \begin{cases} \ln 2 - \ln(1-w), & \text{if } -1 < w \le 0, \\ \ln 2, & \text{if } 0 < w < 1, \end{cases} \qquad (11)$$

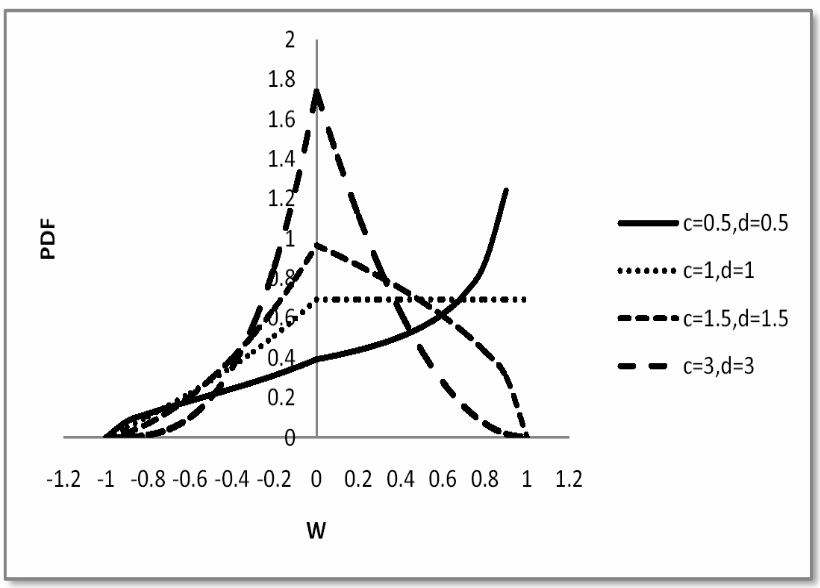
and

$$g_{W'}(w') = \begin{cases} \ln 2, & \text{if } -1 < w' \le 0, \\ \ln 2 - \ln(1+w'), & \text{if } 0 < w' < 1. \end{cases}$$

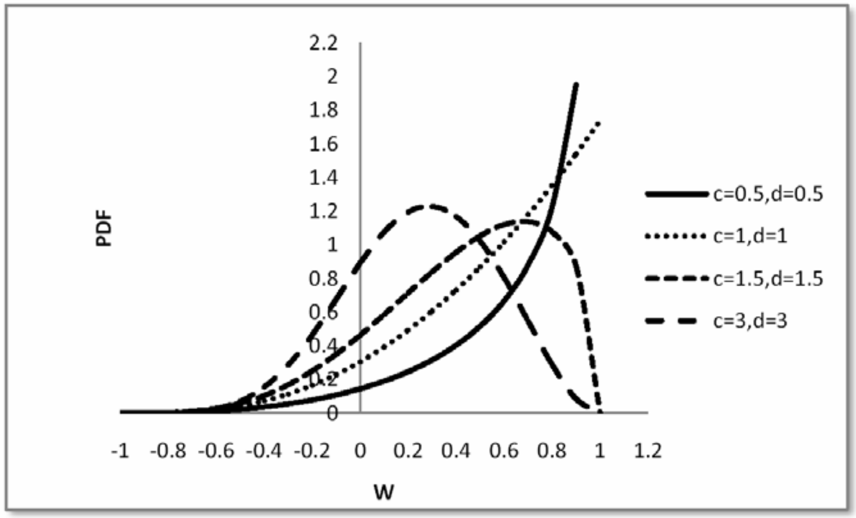
The different plots of Figure 1 illustrate the shapes of the pdf of W for some selected values of a, b, c, and d, where it is obvious that the effect of the parameters is evident.


7
6
5
4
3
2
0
PDF
-1.2
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1
1.2
W

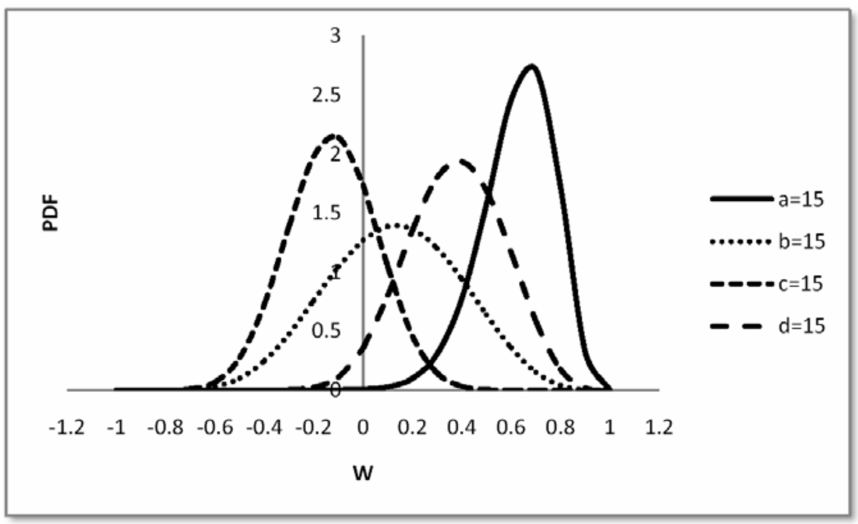
(a)


2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
PDF
-1.2
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1
1.2
W
c=0.5,d=0.5
c=1,d=1
c=1.5,d=1.5
c=3,d=3

(b)



(c)



(d)

Figure 1. Plots of the pdf of $W = X - Y$ for (a): $a = 0.6$, $b = 0.6$; (b): $a = 1$, $b = 1$; (c): $a = 3$, $b = 3$; and (d): when $a = 15$ ($b = 5$, $c = 5$, $d = 5$), when $b = 15$ ($a = 5$, $c = 5$, $d = 5$), when $c = 15$ ($a = 5$, $b = 5$, $d = 5$), when $d = 15$ ($a = 5$, $b = 5$, $c = 5$).

The first plot in Figure 1 indicate that whenever if $a = b < 1$ and $c = d \geq 1,$ then the mode is zero; while, when $a = b = 1$ and $c = d > 1,$ then again the mode is zero in the second plot, also when $a = b = c = d = 1,$ then the curve satisfies Equation (11). If $a = b = 3$ and $c = d \leq 1,$ then the curve is inverse J shaped in the third plot, and if $a = b = c = d \geq 2,$ then the curve is symmetric about mean. Also, in the last plot, when the value of a is large and $b = c = d,$ then the curve is skewed to the left, and when the value of c is large and $a = b = d,$ the curve is skewed to the right, whereas when the value of b is large and $a = c = d,$ the curve is symmetric about mean, and when the value of d is large and $a = b = c,$ then the curve is skewed to the left.

Theorem 2. *If X and Y are jointly distributed random variables with the joint pdf* (1), *then the pdf and cdf of* $T = X / Y$ *are given*, *respectively*, *by*

$$g_T(t) = \frac{\Gamma(a+c)\Gamma(a+b)\Gamma(b+d)}{\Gamma(a)\Gamma(b)\Gamma(c)\Gamma(a+b+d)} \frac{t^{a-1}}{(1+t)^{a+b}} {}_2F_1\left(b+d-c,\, a+b;\, a+b+d;\, \frac{t}{1+t}\right), \tag{12}$$

where $0 < t < \infty;$ *and*

$$G_T(t) = \frac{\Gamma(a+c)\Gamma(a+b)\Gamma(b+d)}{\Gamma(a)\Gamma(b)\Gamma(c)\Gamma(a+b+d)} \sum_{l=0}^{\infty} \frac{(a+b)_l (b+d-c)_l}{(a+b+d)_l\, l!} B_{(t/(1+t))}(a+l,\, b). \tag{13}$$

Proof. Using (1), the joint pdf of $T = X / Y$ and Y is given by

$$f(t,\, y) = \frac{1}{B(a,\, c)B(b,\, d)} (ty)^{a-1} y^{b-1} (1-ty)^{c-b-d} (1-ty-y)^{d-1},$$

$$0 < t < \infty,\, 0 < (1+t)y < 1.$$

Thus, the marginal pdf of T is

$$g_T(t) = \int_{-\infty}^{\infty} y f(t, y) dy$$

$$= \frac{1}{B(a, c)B(b, d)} \frac{t^{a-1}}{(1+t)^{a+b}} \int_0^1 u^{a+b-1}(1-u)^{d-1}\left(1 - \frac{t}{1+t}u\right)^{-(b+d-c)} du,$$

where the transformation $u = (1+t)y$. Using (2), we obtain (11).

Equation (13) is obtained by integration of $g_T(t)$ in (12), with respect to t. This completes the proof of the theorem.

Remarks. (1) The pdf of $T^{-1} = \frac{Y}{X}$ is given by

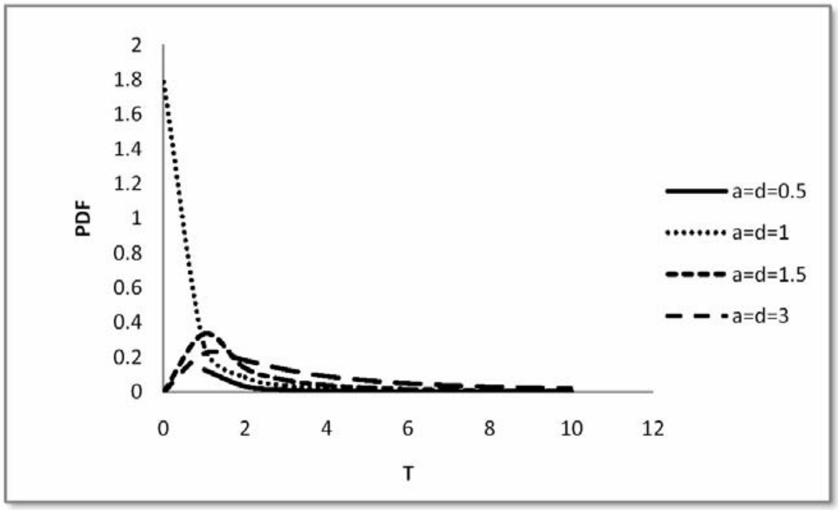
$$g_{T^{-1}}(z) = \frac{\Gamma(a+c)\Gamma(a+b)\Gamma(b+d)}{\Gamma(a)\Gamma(b)\Gamma(c)\Gamma(a+b+d)} \frac{z^{b-1}}{(1+z)^{a+b}} {}_2F_1\left(b+d-c, a+b; a+b+d; \frac{1}{1+z}\right),$$

where $0 < z < \infty$.

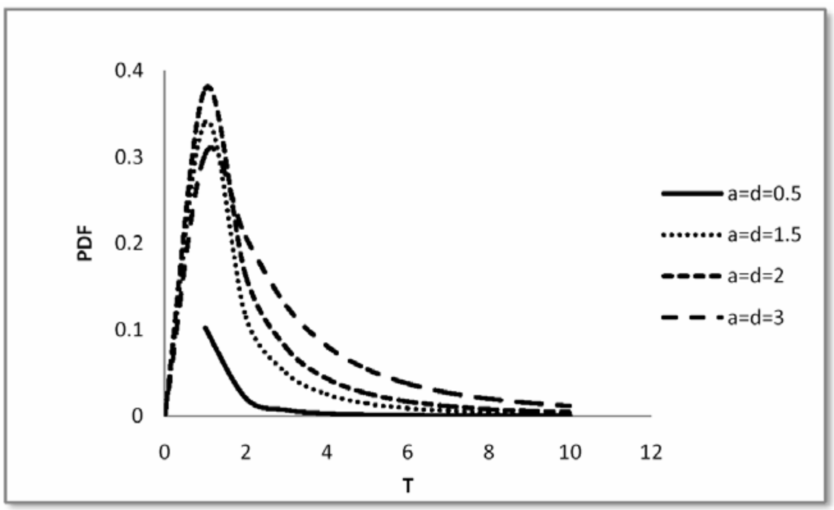
(2) The stress-strength reliability function is given as $R = P(X < Y)$ $= P\left(\frac{X}{Y} < 1\right) = P(T < 1) = G_T(1).$ Using (13), we get

$$G_T(1) = \frac{\Gamma(a+c)\Gamma(a+b)\Gamma(b+d)}{\Gamma(a)\Gamma(b)\Gamma(c)\Gamma(a+b+d)} \sum_{l=0}^{\infty} \frac{(a+b)_l (b+d-c)_l}{(a+b+d)_l \, l!} B_{(1/2)}(a+l, b). \tag{14}$$

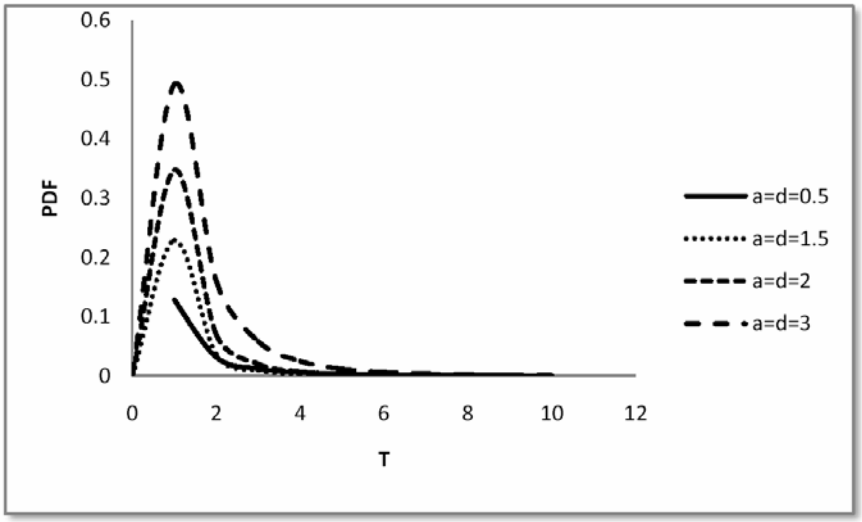
Figure 2 with its various plots illustrates the shapes of the pdf of T for selected values of a, b, c, and d, where the effect of the parameters is quite clear.


PDF
T
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
2
4
6
8
10
12
a=d=0.5
a=d=1
a=d=1.5
a=d=3

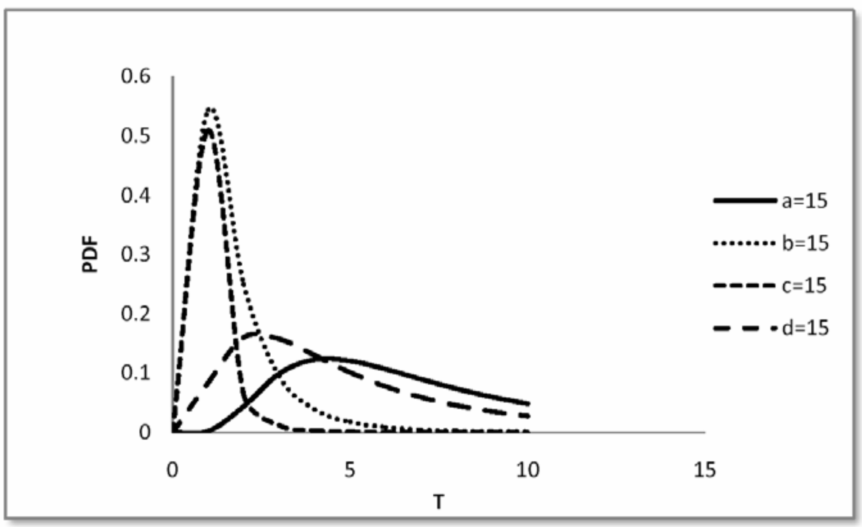
(a)


PDF
T
0.4
0.3
0.2
0.1
0
0
2
4
6
8
10
12
a=d=0.5
a=d=1.5
a=d=2
a=d=3

(b)



(c)



(d)

Figure 2. Plots of the pdf of $T = X/Y$ for (a): $b = 2.5$, $c = 2.5$; (b): $b = 3$, $c = 3$; (c): $b = 5$, $c = 5$; and (d): when $a = 15$ ($b = 5$, $c = 5$, $d = 5$), when $b = 15$ ($a = 5$, $c = 5$, $d = 5$), when $c = 15$ ($a = 5$, $b = 5$, $d = 5$), when $d = 15$ ($a = 5$, $b = 5$, $c = 5$).

From the first plot of Figure 2, it can be seen that if $b = c \geq 2.5$, $a = d = 1$ and $t \to 0$, then pdf is $g_T(t) = b^2/(b+1)$. If $b = c \geq 2.5$, $a = d = 0.5$ and $t \to 0$, then $g_T(t) \to \infty$, and also where $a = d > 1$ and $b = c \geq 2.5$, then the curve is skewed to the right in the second and third plots. As for the fourth plot, when the values of $b = 15$, $a = c = d$ and $c = 15$, $a = b = d$, the curve is again skewed to right, while for values of $a = 15$, $b = c = d = 5$ and $d = 15$, $a = b = c = 5$, the curve away from the x-axis.

Theorem 3. *If X and Y are jointly distributed random variables with the joint pdf* (1), *then the pdf and cdf of* $P = XY$ *are given*, *respectively*, *by*

$$g_P(p) = \frac{\Gamma(a+c)\Gamma(b+d)\Gamma(d)}{2^{a+c-2b-2d}\Gamma(a)\Gamma(b)\Gamma(c)\Gamma(2d)} p^{b-1}\left(1-\sqrt{1-4p}\right)^{a-b-d}\left(1+\sqrt{1-4p}\right)^{c-b-d}$$

$$\times (1-4p)^{d-\frac{1}{2}} F_1\left(d, b+d-a, b+d-c; 2d; 2-\frac{1+\sqrt{1-4p}}{2p}, 2-\frac{1-\sqrt{1-4p}}{2p}\right), \tag{15}$$

where $0 < p < \frac{1}{4}$; *and*

$$G_P(p) = B\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{(d)_{m+n}(b+d-a)_m(b+d-c)_n}{(2d)_{m+n}\, m!\, n!}\left(\frac{1}{2}\right)^{m+n}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^{m+i}$$

$$\times \binom{a-b-d+n}{i}\binom{c-b-d+m}{j}(4)^{(m+n-b)} B_{4_p}\left(b-m-n, d+\frac{1}{2}(m+n+i+j+1)\right), \tag{16}$$

where $B = \dfrac{\Gamma(a+c)\Gamma(b+d)\Gamma(d)}{2^{a+c-2b-2d}\Gamma(a)\Gamma(b)\Gamma(c)\Gamma(2d)}$.

Proof. Using (1), the joint pdf of X and $P = XY$ is given by

$$f(x,\ p) = Kx^{a-2}\left(\frac{p}{x}\right)^{b-1}(1-x)^{c-b-d}\left(1-x-\frac{p}{x}\right)^{d-1}, \quad 0 < p < \frac{1}{4}$$

$$= Kp^{b-1}x^{a-b-d}(1-x)^{c-b-d}(x-p_1)^{d-1}(p_2-x)^{d-1},\ p_1 < x < p_2,$$

where $K = \frac{1}{B(a,\ c)B(b,\ d)}$, $p_1 = \frac{\left(1-\sqrt{1-4p}\right)}{2}$, and $p_2 = \frac{\left(1+\sqrt{1-4p}\right)}{2}$.

Thus, the marginal pdf of P can be written as

$$g_P(p) = Kp^{b-1}\int_{p_1}^{p_2} x^{a-b-d}(1-x)^{c-b-d}(x-p_1)^{d-1}(p_2-x)^{d-1}dx$$

$$= Kp^{b-1}p_1^{a-b-d}(1-p_1)^{c-b-d}(p_2-p_1)^{2d-1}\int_0^1 s^{d-1}(1-s)^{d-1}$$

$$\times\left(1-\left(1-\frac{p_2}{p_1}\right)s\right)^{-(b+d-a)}\left(1-\frac{p_2-p_1}{1-p_1}s\right)^{-(b+d-c)}ds.$$

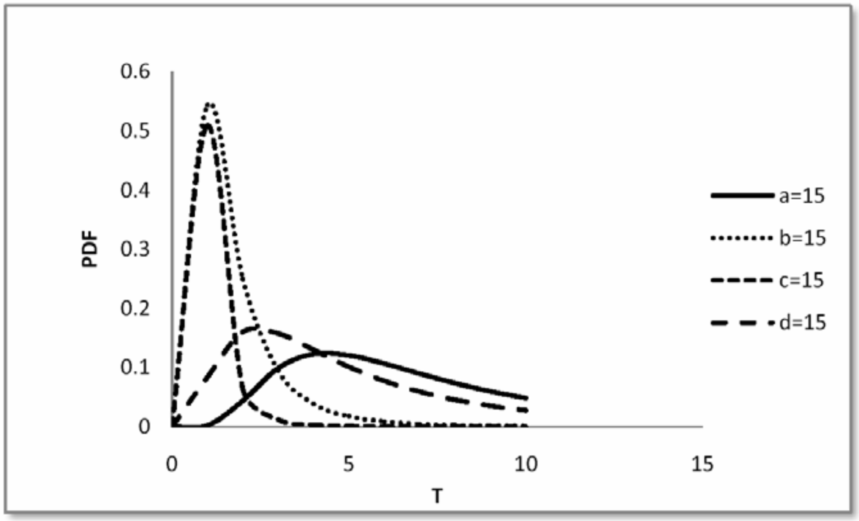
Using the transformation $s = \frac{x-p_1}{p_2-p_1}$ and (4), we have

$$g_P(p) = Kp^{b-1}{p_1}^{a-b-d}(1-p_1)^{c-b-d}(p_2-p_1)^{2d-1}B(d,\ d)$$

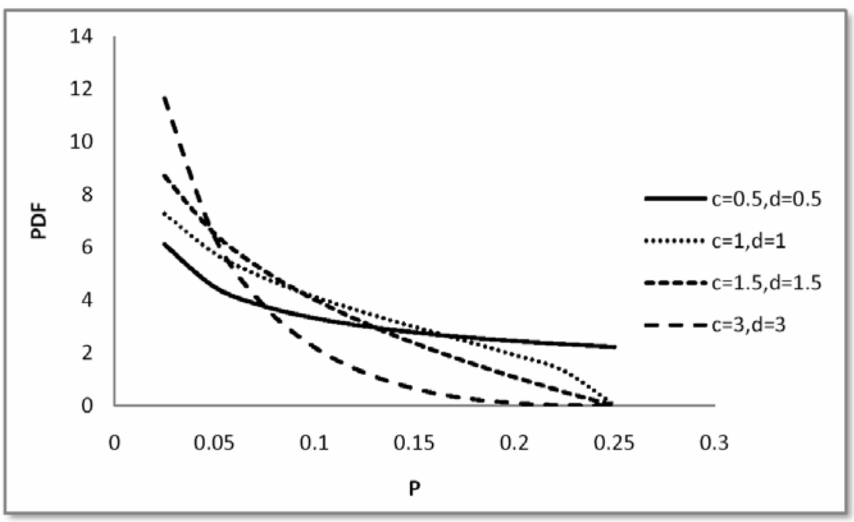
$$\times F_1\left(d,\ b+d-a,\ b+d-c;\ 2d;\ 1-\frac{p_2}{p_1},\ \frac{p_2-p_1}{1-p_1}\right).$$

Hence the pdf can be rewritten as the result given in Equation (13). To obtain the cdf of P, integration of $g_P(p)$ in (15) with respect to p and using (5), leads to the result (16), which completes the proof.

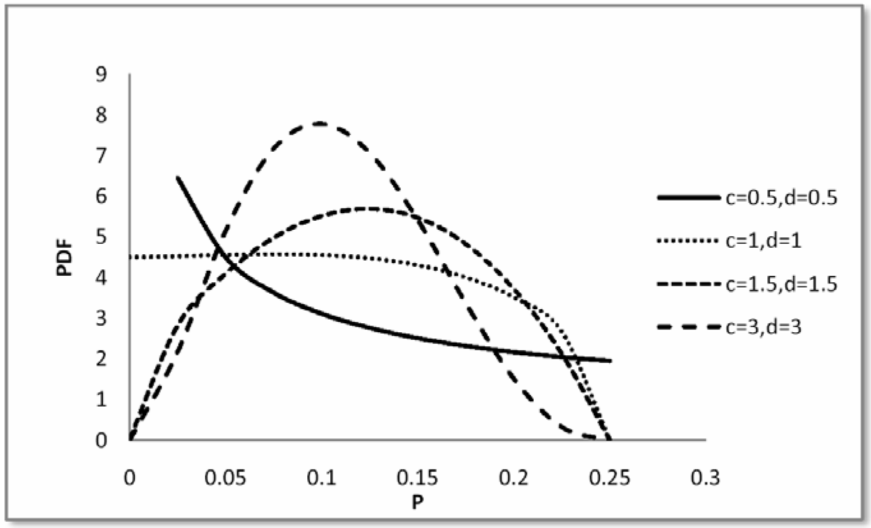
Figure 3 illustrates the shapes of the pdf of P for selected values of a, b, c, and d, and, in this case too, the parameters show a considerable effect on the shapes of the curves.


0.6
0.5
0.4
0.3
0.2
0.1
0
PDF
0
5
10
15
T
a=15
b=15
c=15
d=15

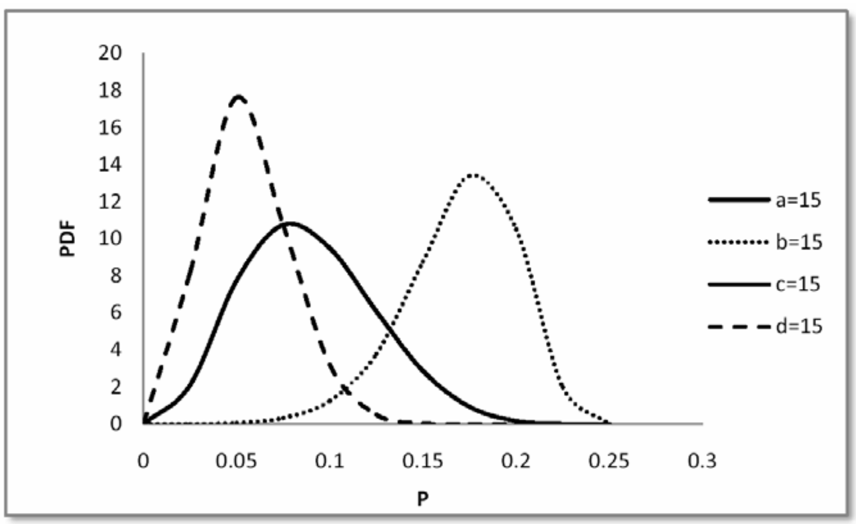
(a)


14
12
10
8
6
4
2
0
PDF
0
0.05
0.1
0.15
0.2
0.25
0.3
P
c=0.5,d=0.5
c=1,d=1
c=1.5,d=1.5
c=3,d=3

(b)



(c)



(d)

Figure 3. Plots of the pdf of $P = XY$ for (a): $a = 0.5$, $b = 0.5$; (b): $a = 1$, $b = 1$; (c): $a = 3$, $b = 3$; and (d): when $a = 15$ ($b = 5$, $c = 5$, $d = 5$), when $b = 15$ ($a = 5$, $c = 5$, $d = 5$), when $c = 15$ ($a = 5$, $b = 5$, $d = 5$), when $d = 15$ ($a = 5$, $b = 5$, $c = 5$).

The first two plots of Figure 3 indicate that for $a = b = 0.5$ and 1, also for increasing values of c and d, the curve is inverse J shaped, decreasing to reach the x-axis, also if $p \to 0$, then $g(p) \to \infty$. Also when $a = b \geq 0.5$, $c = d = 0.5$ and $p = 0.25$, then $g(p) = (\Gamma(b + 0.5))^2 / 2^{(b-2.5)}(\Gamma(b))^2$. If $a = b = c = d > 1$, then the curve is symmetric about mean. Also, in the last plot, similarity is obvious when $a = 15$ and $c = 15$.

Using special properties of the hyper geometric functions, one can derive elementary forms of the pdfs in (12) and (15). This is illustrated in the corollaries below:

Corollary 1. *If X and Y are jointly distributed according to* (1) *and if* $c = (b + d)$, *then T has a beta distribution type II with parameters* (a, b), *given by*

$$g(t) = \frac{1}{B(a, b)} \frac{t^{a-1}}{(1 + t)^{a+b}},$$

where $0 < t < \infty$, *and* T^{-1} *has a beta distribution type II with parameters* (b, a).

Corollary 2. *If X and Y are jointly distributed according to* (1) *and if* $a = c = (b + d)$, *then* $V = 4P$ *has a beta distribution type I with parameters* $(b, d + 0.5)$, *given by*

$$f(\nu) = \frac{1}{B(b, d + 0.5)} \nu^{b-1}(1 - \nu)^{d-\frac{1}{2}},$$

where $0 < \nu < 1$.

3. Moments

For deriving the moments of $W, T,$ and P, we need the following lemma:

Lemma 1. *If X and Y are jointly distributed random variables with the joint pdf in* (1), *then*

$$E\left(X^n Y^m\right) = KB(a + n, c + m)B(b + m, d), \tag{17}$$

for $a + n > 0,\ b + m > 0,\ c + m > 0,$ *and* $K = \dfrac{1}{B(a,\ c)B(b,\ d)}.$

Proof. Knowing that

$$E\left(X^n Y^m\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} X^n Y^m h(x,\ y)dydx, \tag{18}$$

and substituting with (1) into (18), we get

$$E\left(X^n Y^m\right) = K\int_0^1\int_0^{1-x} x^{n+a-1} y^{m+b-1}(1-x)^{c-b-d}(1-x-y)^{d-1}dydx$$

$$= K\int_0^1 x^{n+a-1}(1-x)^{c-b-d}\left(\int_0^{1-x} y^{m+b-1}\left[(1-x)\left(1-\frac{y}{1-x}\right)^{d-1}\right]dy\right)dx.$$

Using the transformation $u = \dfrac{y}{1-x}$, we obtain

$$E\left(X^n Y^m\right) = K\int_0^1 x^{n+a-1}(1-x)^{c-b-1}\left(\int_0^1 [u(1-x)]^{m+b-1}(1-u)^{d-1}(1-x)du\right)dx$$

$$= K\int_0^1 x^{n+a-1}(1-x)^{c+m-1}dx\int_0^1 u^{m+b-1}(1-u)^{d-1}du.$$

Solving the integral, we obtain the result (17). This completes the proof of the lemma.

Theorem 4. *If X and Y are jointly distributed according to the joint pdf* (1), *and* $W = X - Y,$ *then*

$$E(W^m) = K\sum_{l=0}^{m}(-1)^{m-l}\binom{m}{l}B(a+l,\ c+m-l)B(b+m-l,\ d),$$

$$\text{for } m \geq 1,\ b+m > l,\ c+m > l, \tag{19}$$

where K is as defined in Lemma 1.

Proof. Starting with the definition of expectation and using the binomial expansion, we get

$$E(W^m) = E[(X - Y)^m] = \sum_{l=0}^{m} (-1)^{m-l} \binom{m}{l} E\left(X^l Y^{m-l}\right), \quad m \geq 1.$$

Applying (17), we obtain the result (19).

Note that, if $m = 1$, then

$$E(W) = \frac{a(b + d) - cb}{(a + c)(b + d)}.$$

If either $a = b = c = d$ or $a = c, b = d$, then the expectation of W is constant and equal to 0.25. If $a = b, c = d$, then the expectation of W equals the square of the expectation of X. That is,

$$E(W) = \frac{a^2}{(a + c)^2} = (E(X))^2.$$

Also, if $m = 2$, then

$$E\left(W^2\right) = \frac{cb(c + 1)(b + 1) - 2acb(b + d + 1) + a(a + 1)(b + d)(b + d + 1)}{(a + c + 1)(a + c)(b + d + 1)(b + d)},$$

and hence

$$\operatorname{Var}(W) = \frac{cb(c + 1)(b + 1) - 2acb(b + d + 1) + a(a + 1)(b + d)(b + d + 1)}{(a + c + 1)(a + c)(b + d + 1)(b + d)}$$

$$- \left(\frac{a(b + d) - cb}{(a + c)(b + d)}\right)^2.$$

If $a = b = c = d$, then $\operatorname{Var}(W)$ depends on a say and is an increasing function in a.

Theorem 5. *If X and Y are jointly distributed according to the joint pdf* (1), *and $T = X / Y$, then*

$$E\left(T^m\right) = KB(a + m, c - m)B(b - m, d),$$

$$for\ \ m \geq 1,\ b > m,\ c > m, \tag{20}$$

where K is as defined in Lemma 1.

Proof. Writing $E\left(T^m\right) = E\left(X^m Y^{-m}\right)$ and substituting in Equation (17), we obtain the result (20).

Now, if $m = 1,$ then

$$E(T) = \frac{a(b + d - 1)}{(c - 1)(b - 1)},$$

where $b > 1,\ c > 1.$

Also, if $m = 2,$ then

$$E\left(T^2\right) = \frac{a(a + 1)(b + d - 1)(b + d - 2)}{(c - 1)(c - 2)(b - 1)(b - 2)},$$

and hence

$$\operatorname{Var}(T) = \frac{a(a + 1)(b + d - 1)(b + d - 2)}{(c - 1)(c - 2)(b - 1)(b - 2)} - \left(\frac{a(b + d - 1)}{(c - 1)(b - 1)}\right)^2,$$

where $b > 2,\ c > 2.$

If $c = b + d,$ then

$$E(T) = \frac{a}{(b - 1)} \text{ and } \operatorname{Var}(T) = \frac{a(a + b - 1)}{(b - 1)^2(b - 2)},$$

which are the expectation and variance of beta distribution type II with parameters $(a,\ b).$

Theorem 6. *If X and Y are jointly distributed according to the joint pdf* (1), *and* $P = XY,$ *then*

$$E\left(P^m\right) = KB(a + m,\ c + m)B(b + m,\ d),$$

$$for\ \ m \geq 1,\ a + m > 0,\ b + m > 0,\ c + m > 0, \tag{21}$$

where K is as defined in Lemma 1.

Proof. Setting $m = n$ in the relation (17), we get the result (21), thus completing the proof.

Now, if $m = 1$, then

$$E(P) = \frac{acb}{(a + c + 1)(a + c)(b + d)}.$$

If either $a = c = b = d$, or $a = c$, $b = d$, then

$$E(P) = \frac{a}{4(2a + 1)}.$$

Also, if $m = 2$, then

$$E\left(P^2\right) = \frac{acb(a + 1)(c + 1)(b + 1)}{(a + c + 3)(a + c + 2)(a + c + 1)(a + c)(b + d + 1)(b + d)},$$

and hence

$$\operatorname{Var}(P) = \frac{acb(a + 1)(c + 1)(b + 1)}{(a + c + 3)(a + c + 2)(a + c + 1)(a + c)(b + d + 1)(b + d)} - \left(\frac{acb}{(a + c + 1)(a + c)(b + d)}\right)^2.$$

If $a = c = b + d$, then

$$E(4P) = \frac{b}{(b + d + 0.5)} \text{ and } \operatorname{Var}(4P) = \frac{b(d + 0.5)}{(b + d + 0.5)^2(b + d + 1.5)},$$

which are the expectation and variance of beta distribution type I with parameters $(b, d + 0.5)$.

4. Generation of (X, Y)

Notice that the marginal distribution of X and the conditional distribution of Y given $X = x$ are given, respectively, by

$$f(x) = \frac{x^{a-1}(1 - x)^{c-1}}{B(a, c)}, \quad 0 < x < 1, \tag{22}$$

and

$$f(y \mid x) = \frac{(1-x)}{B(b,\, d)} \left(\frac{y}{1-x}\right)^{b-1} \left(1 - \frac{y}{1-x}\right)^{d-1}, \quad 0 < y < 1 - x, \quad 0 < x < 1. \tag{23}$$

From (22) and (23), we see that the distribution of X is beta with parameters $(a,\, c)$, while the distribution of random variable $Z = \frac{Y}{1-x}$ is beta with parameters $(b,\, d)$. Using (22) and (23), we can generate an observation $(X,\, Y)$ from the distribution in (1) as follows:

Step 1. Generate X from $B(a,\, c)$.

Step 2. Generate Z from $B(b,\, d)$.

Step 3. Set $Y = Z(1-x)$.

Step 4. If $X + Y < 1$, then deliver $(X,\, Y)$.

Step 5. If the inequality is violated, reject Y and repeat Steps 2 through 4 again.

5. Application

Here, we provide an application of the results in Sections 2 and 3. We use the data set on compositions of lavas from Skye (see Nadarajah and Kotz [12]). The three variables are: A = proportions of sodium and potassium oxides, F = proportions of iron oxide, and M = proportions of magnesium oxide.

Note that each column in Table 1 adds to 1. Our interest is showing the differences, ratios, and products of proportions of A and F, the proportions of A and M, and the proportions of F and M; and determine the stress-strength reliability function $P(A < F)$, $P(A < M)$, and $P(F < M)$. An obvious model in this situation would be the generalized Dirichlet distribution given by the joint pdf (1).

Table 1. Data on compositions of lavas from Skye

A	F	M
0.52	0.42	0.06
0.52	0.44	0.05
0.47	0.48	0.05
0.45	0.49	0.06
0.4	0.5	0.1
0.37	0.54	0.09
0.27	0.58	0.15
0.27	0.54	0.19
0.23	0.59	0.18
0.22	0.59	0.19
0.21	0.6	0.19
0.25	0.53	0.22
0.24	0.54	0.22
0.22	0.55	0.23
0.22	0.56	0.22
0.2	0.58	0.22
0.16	0.62	0.22
0.17	0.57	0.26
0.14	0.54	0.32
0.13	0.55	0.32
0.13	0.52	0.35
0.14	0.47	0.39
0.24	0.56	0.2

The distribution in (1) reasonably fits the three bivariate data sets of proportions: data set 1 containing the values (A, F), data set 2 containing the values (A, M), and data set 3 containing the values (F, M). Table 2 gives the estimates of $a, b, c,$ and d, which were obtained by using the maximum likelihood method (see Nadarajah and Kotz [12]). Table 3 gives estimated values of the moments $E(W)$, $E(T)$, and $E(P)$; (19), (20), and (21) and sample means, respectively, obtained.

Table 2. Estimated values of a, b, c, and d

Data set	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{d}$
(A, F)	3.827	13.474	10.350	4.497
(A, M)	3.827	4.497	10.350	13.474
(F, M)	53.331	1.991	45.930	2.724

Table 3. Estimated $E(W)$, $E(T)$, $E(P)$ and sample mean

Data set	$E(W)$	$\overline{W}$	$E(T)$	$\overline{T}$	$E(P)$	$\overline{P}$
(A, F)	– 0.2774	– 0.2691	0.5569	0.5201	0.1380	0.14
(A, M)	0.0873	0.0739	1.9864	2.7256	0.0461	0.0415
(F, M)	0.3419	0.3430	4.4497	3.8852	0.1039	0.1061

Figures 4-6, 7-9, and 10-12 show the fitted pdfs of W(8), of T(12), and of P (15), using the estimated parameters, and histograms using the data.

As we see from Table 3, estimated values of the expectations using the estimates of the parameters, is approximately the same as the sample means. This indicates that the distributions in (8), (12), and (15) fit well the data.

The stress-strength reliability functions are given in Table 4 by using Equation (10) of W and Equation (14) of T.

Table 4. Some reliability values

Data set	$P(W < 0) = P(T < 1)$
(A, F)	0.8946
(A, M)	0.2997
(F,M)	0.0040

We see that the results in Table 4 are consistent with the graphs in Figures 4-6 and 7-9. Also, the results in Table 4 are consistent with the results in Table 3.

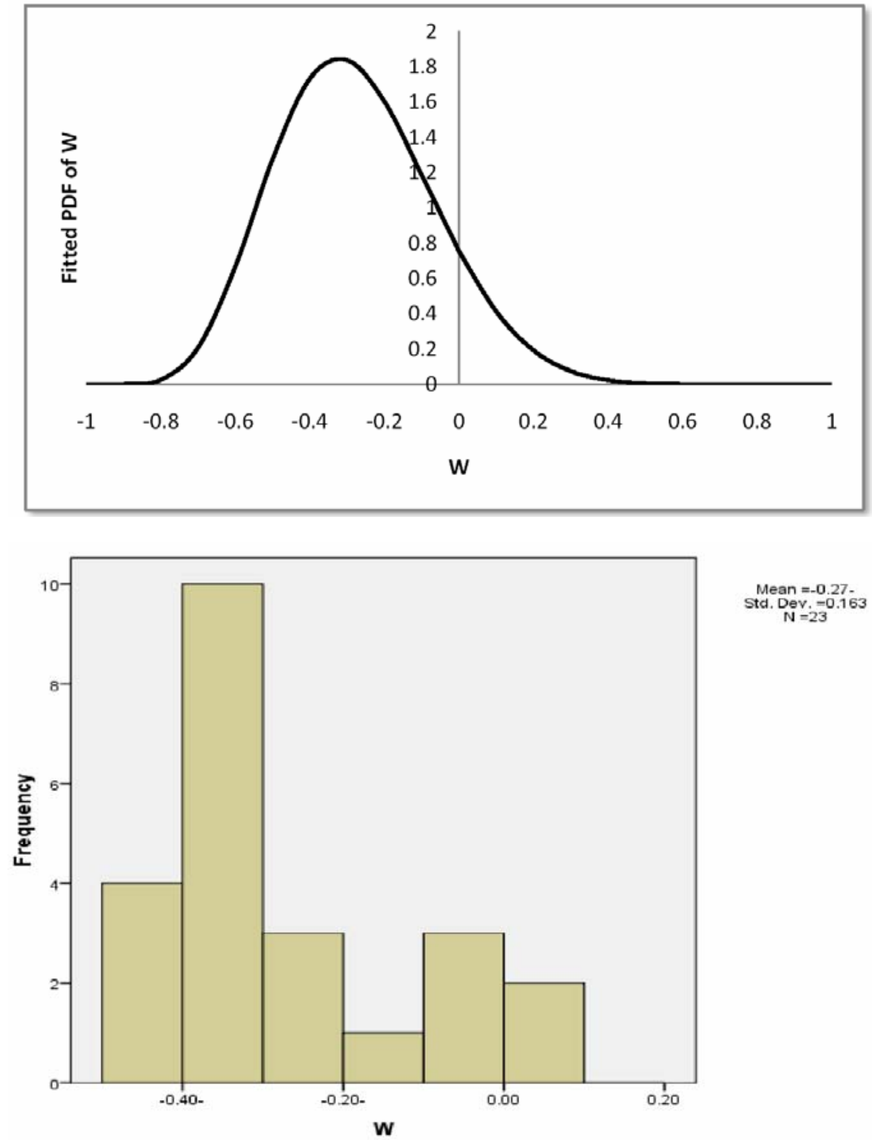


Figure 4. Fitted pdf of $W = X - Y$ given by (8) when X is sodium and potassium oxides and Y is iron oxide.

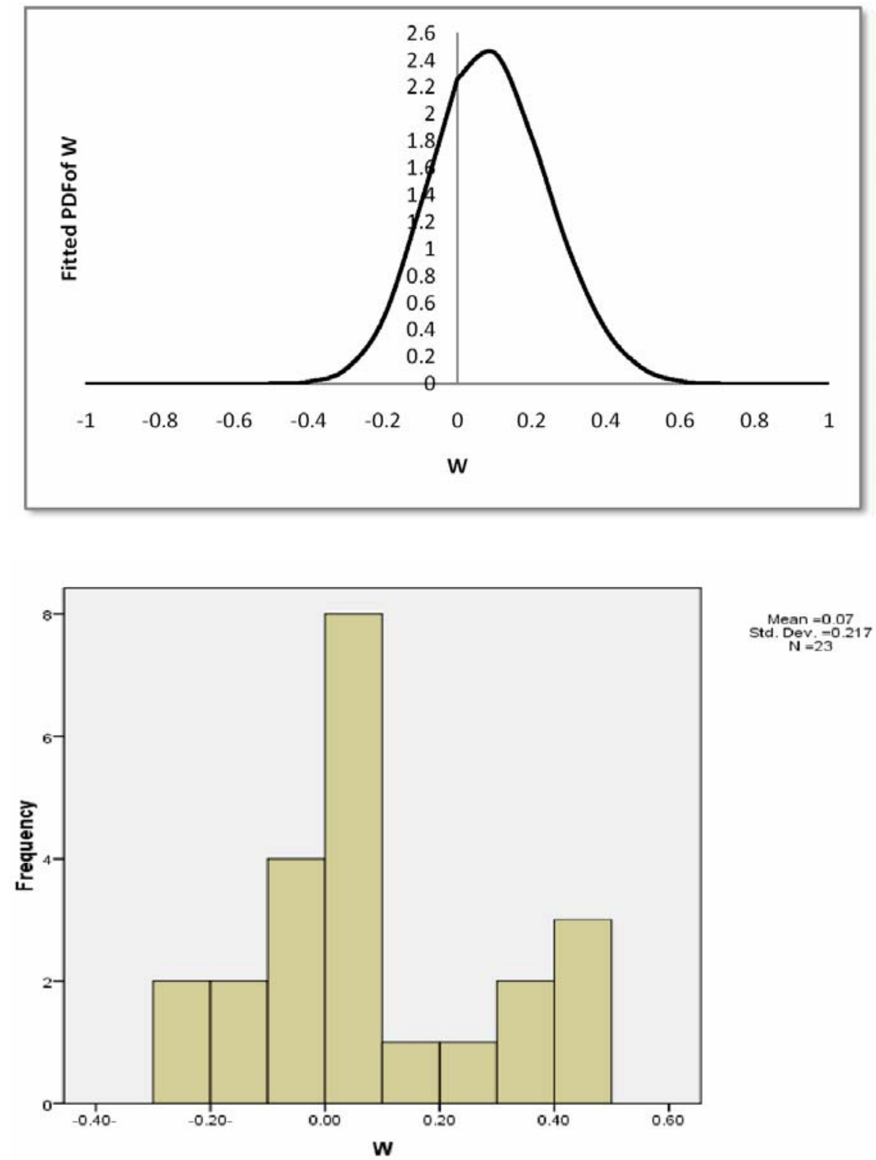


Figure 5. Fitted pdf of $W = X - Y$ given by (8) when X is sodium and potassium oxides and Y is magnesium oxide.

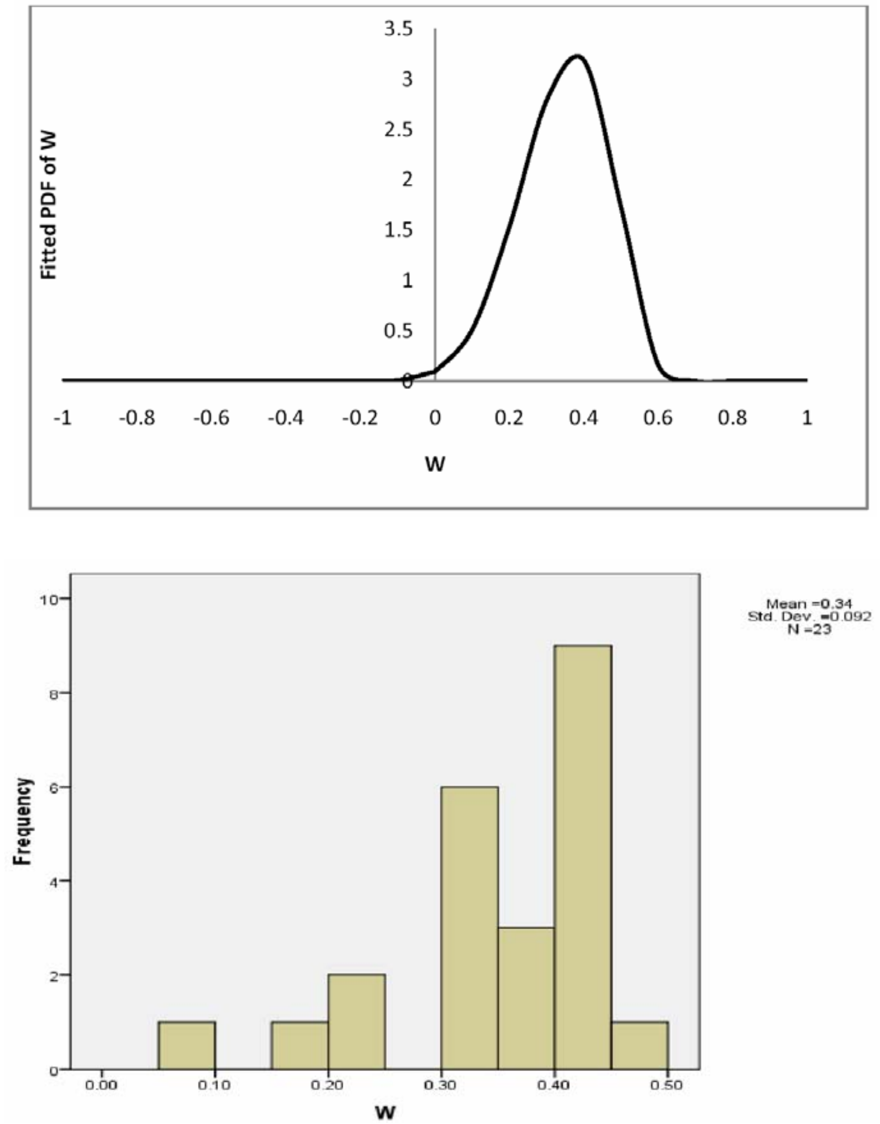
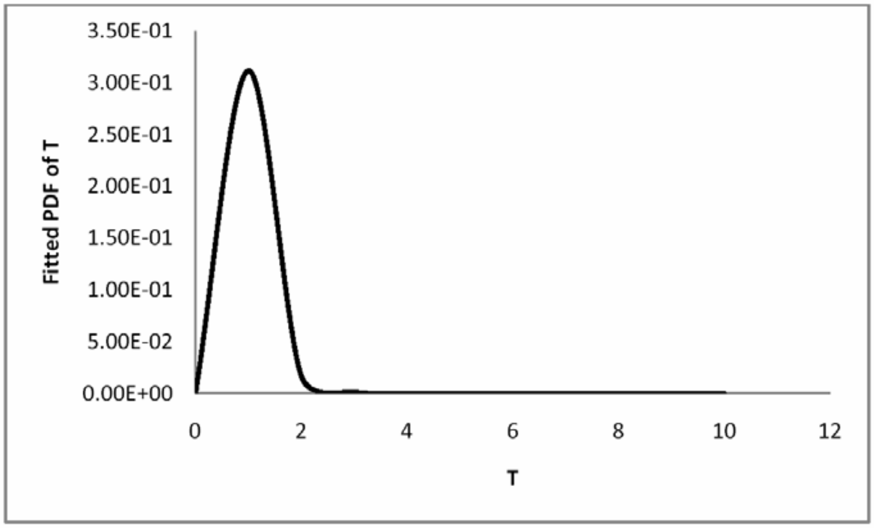


Figure 6. Fitted pdf of $W = X - Y$ given by (8) when X is iron oxide and Y is magnesium oxide.



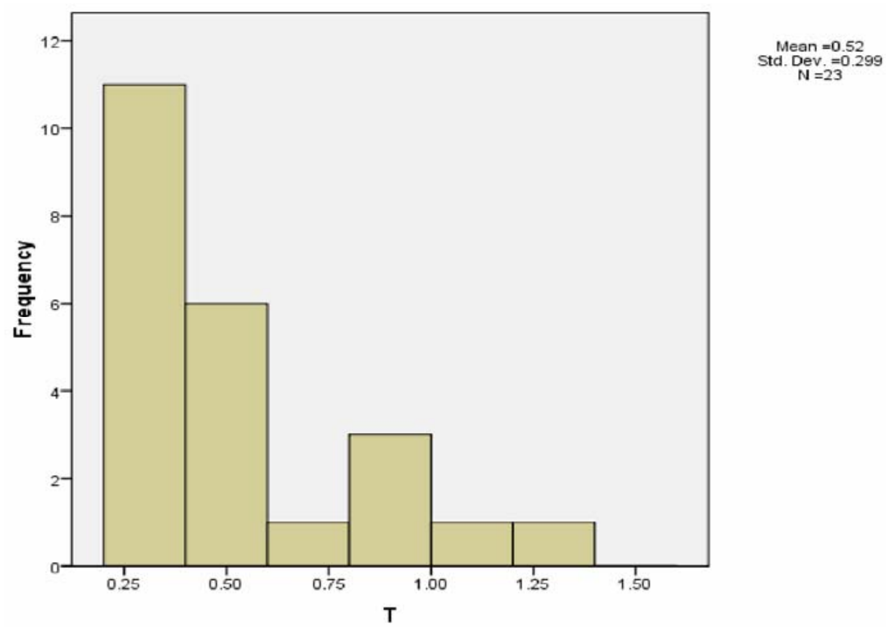


Figure 7. Fitted pdf of $T = X/Y$ given by (12) when X is sodium and potassium oxides and Y is iron oxide.

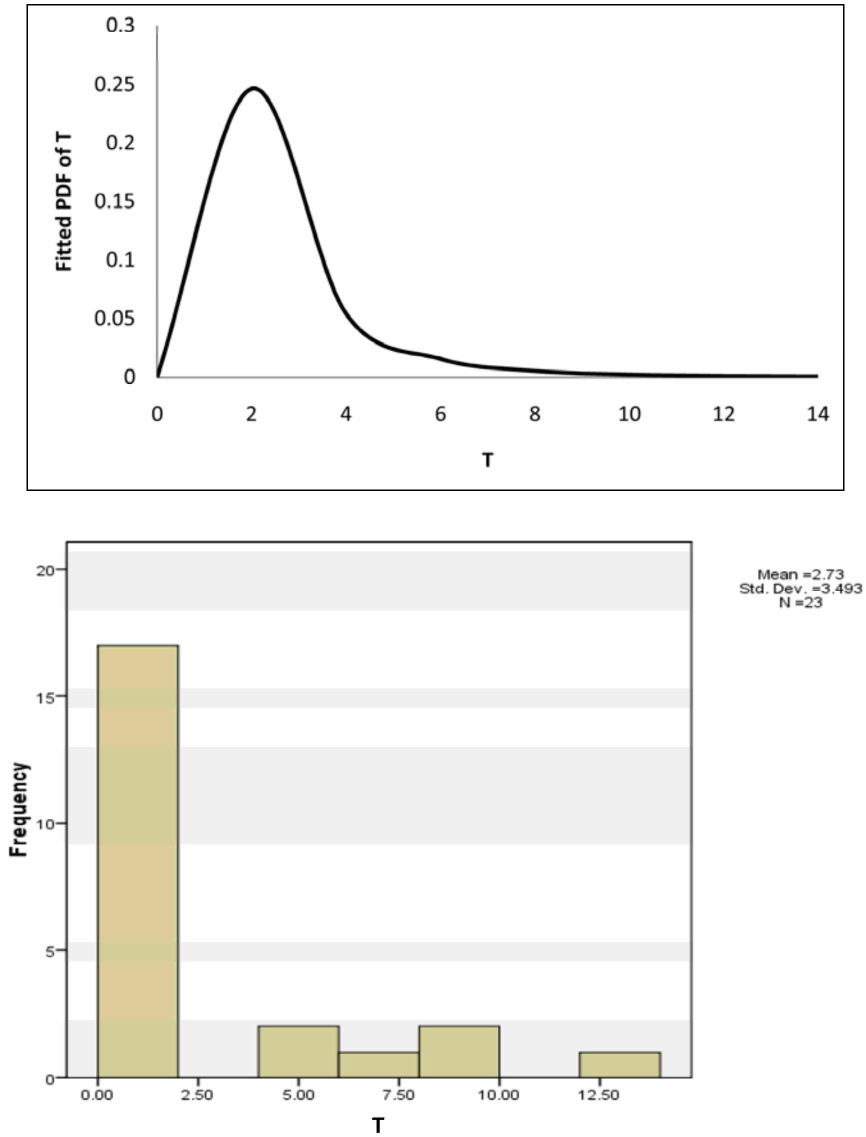


Figure 8. Fitted pdf of $T = X/Y$ given by (12) when X is sodium and potassium oxides and Y is magnesium oxide.

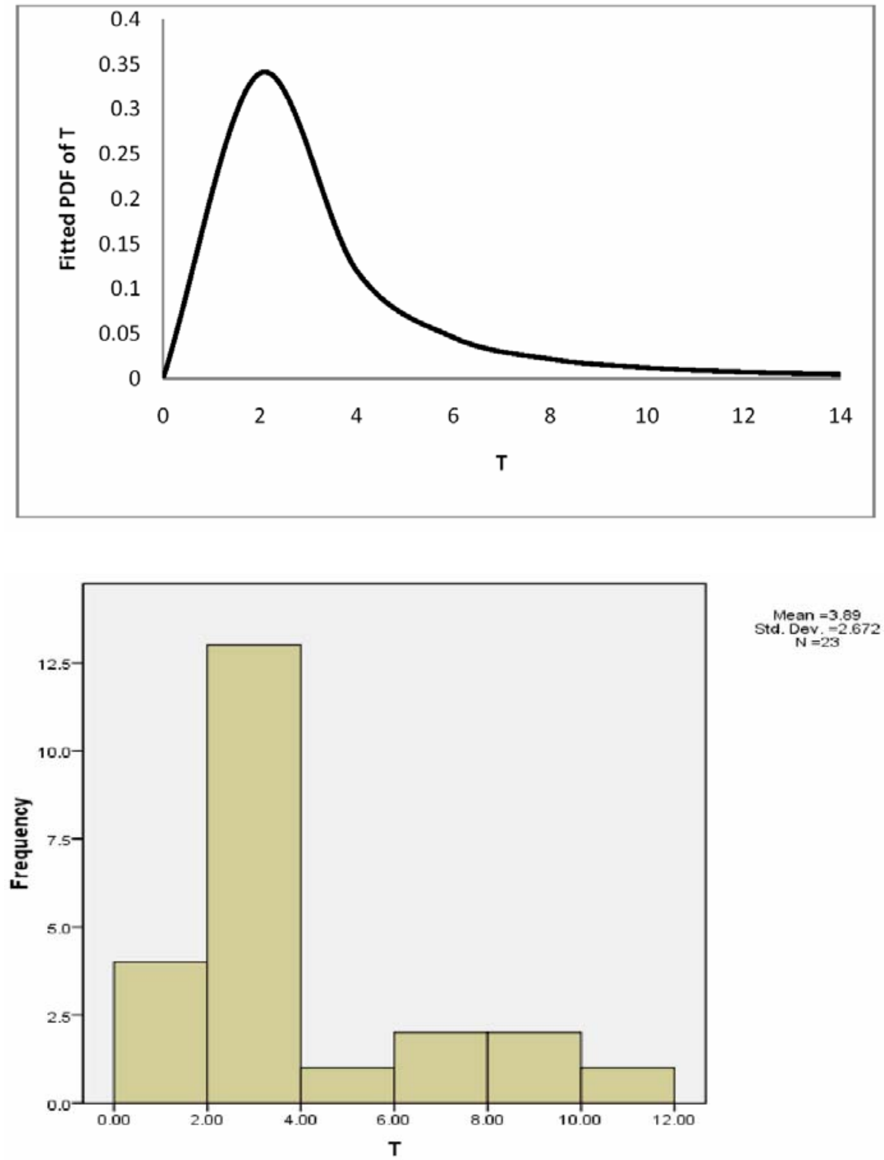


Figure 9. Fitted pdf of $T = X/Y$ given by (12) when X is iron oxide and Y is magnesium oxide.

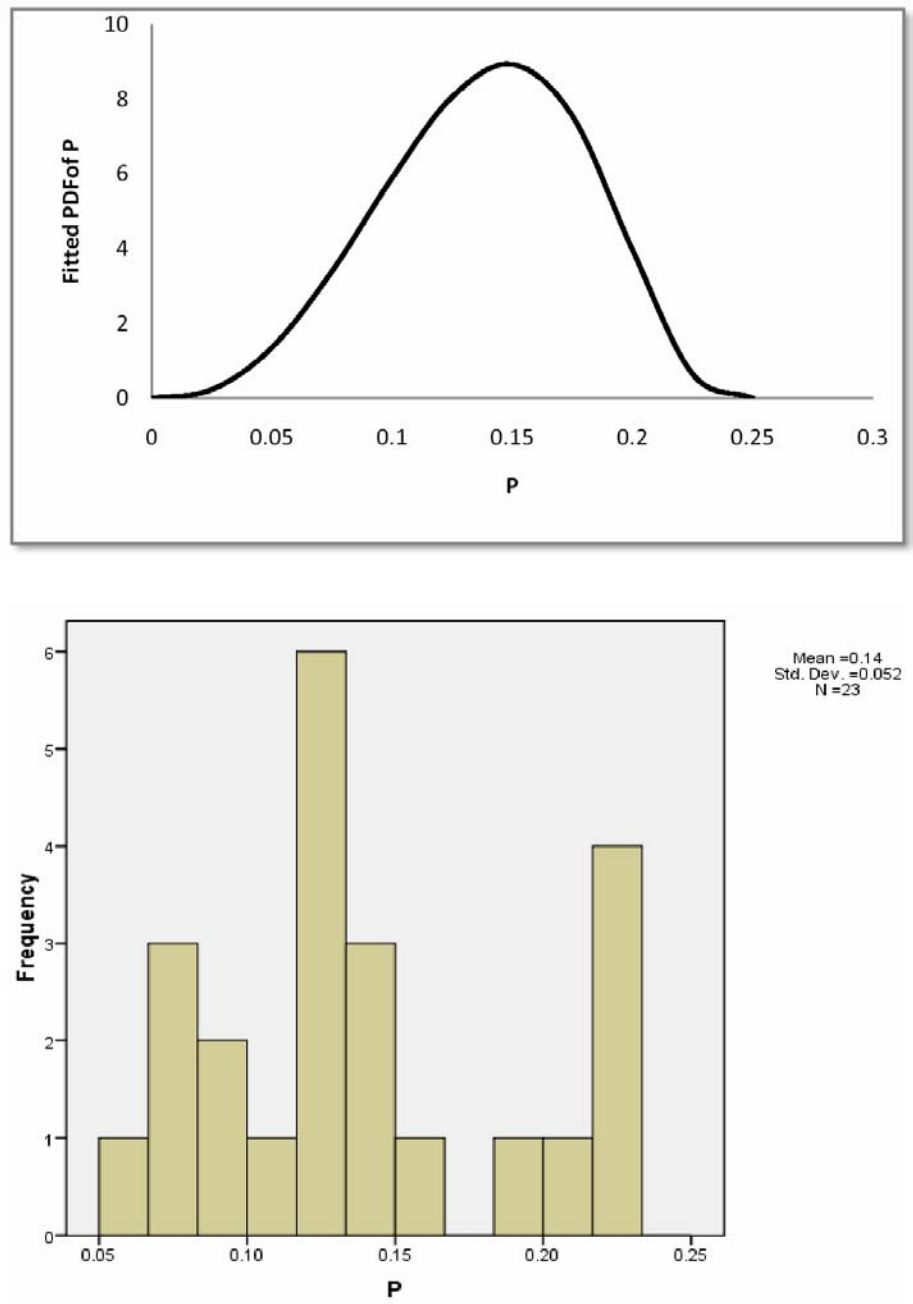
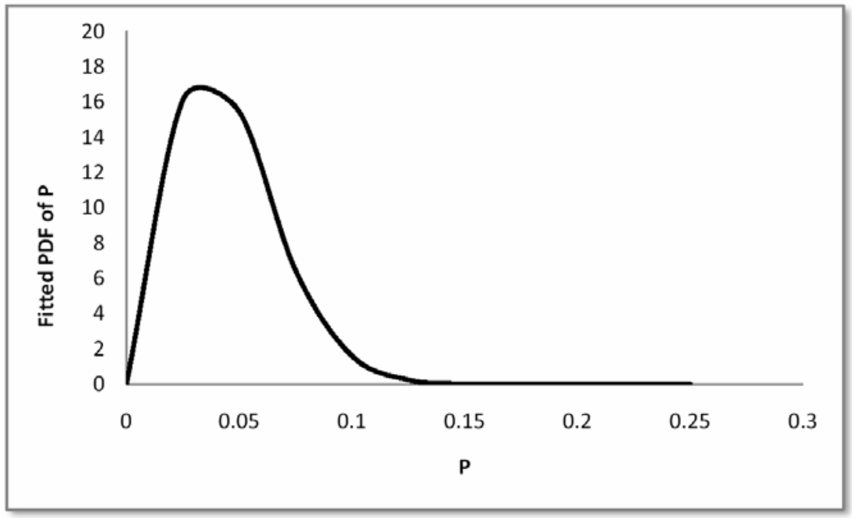


Figure 10. Fitted pdf of $P = XY$ given by (15) when X is sodium and potassium oxides and Y is iron oxide.



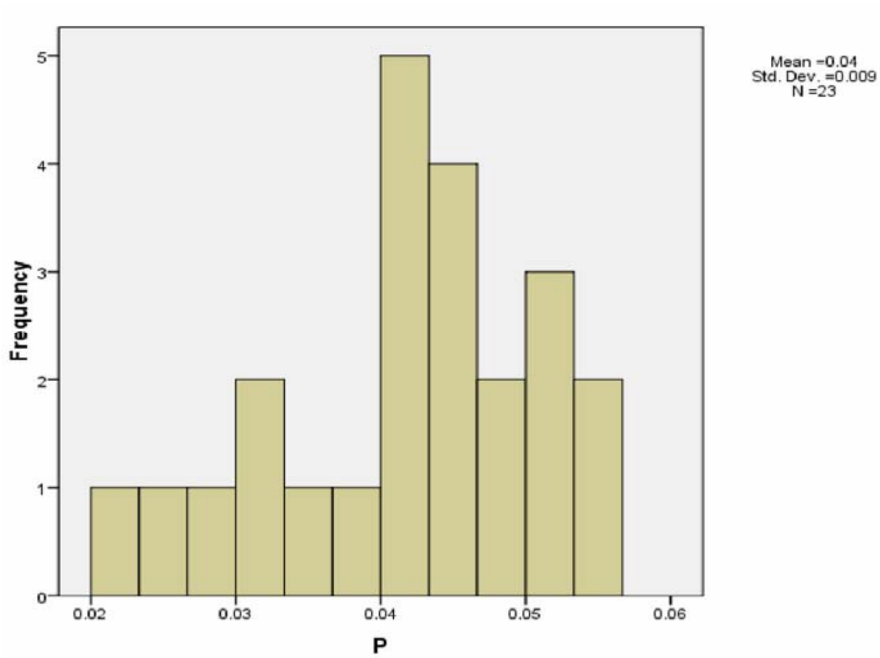
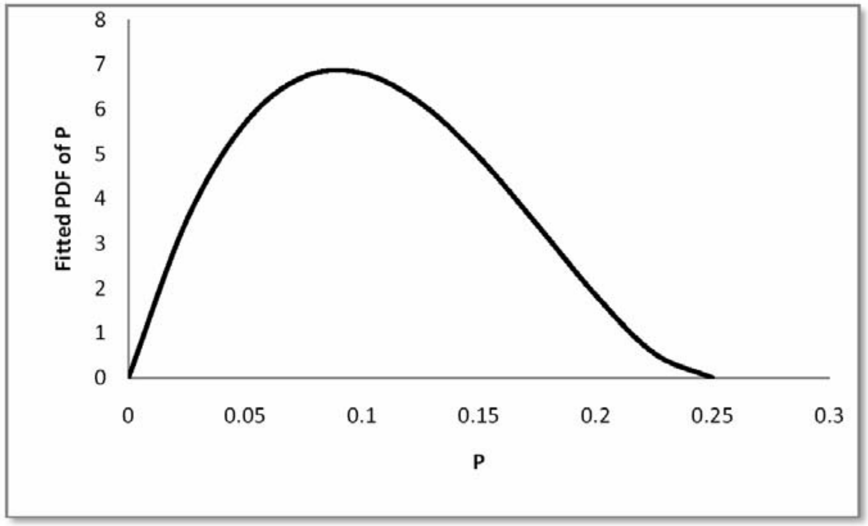


Figure 11. Fitted pdf of $P = XY$ given by (15) when X is sodium and potassium oxides and Y is magnesium oxide.



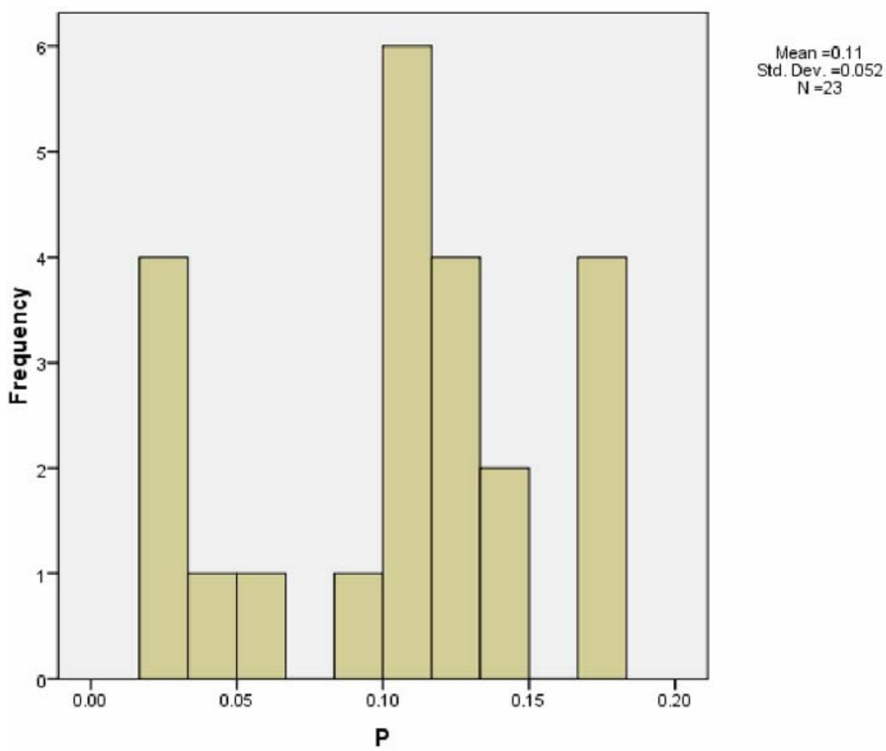


Figure 12. Fitted pdf of $P = XY$ given by (15) when X is iron oxide and Y is magnesium oxide.

Clearly from Figures 4-6, 7-9, and 10-12, we can see that the data fits the given distributions.

References

[1] D. A.-Grivas and A. Asaoka, Slope safety prediction under static and seismic loads, Journal of the Geotechnical Engineering Division Proceedings of the American Society of Civil Engineers 108 (1982), 713-729.

[2] S. Al-Ruzaiza and A. El_Gohary, The distributions of sums, products and ratios of inverted bivariate beta distribution, Applied Mathematical Sciences 2(48) (2008), 2377-2391.

[3] F. J. Apostolakis and P. Moieni, The foundations of models of dependence in probabilistic safety assessment, Reliability Engineering 18 (1987), 177-195.

[4] C. Chatfiled, A marking application of a characterization theorem, In: A Modern Course on Distributions in Scientific Work, Model Building and Model Selection (G. P. Patil, S. Kotz and J. K. Ord, eds.), Reidel, Dordrecht, The Netherland, 2 (1975), 175-185.

[5] R. Connor and J. Mosimann, Concepts of independence for proportions with a generalization of the Dirichlet distribution, Journal of the American Statistical Association 64 (1969), 194-206.

[6] I. Gradshteyn and I. Ryzhik, Table of Integrals, Series, Products, Academic Press, 1980.

[7] H. Grubel, Internationally diversified portfolios: Welfare gains capital flows, American Economic Review 58 (1968).

[8] A. K. Gupta and S. Nadarajah, Handbook of Beta Distribution and its Applications, Marcel Dekker, New York, 2004.

[9] R. W. Hoyer and L. Mayer, The equivalence of various objective functions in a stochastic model of electoral competition, Technical Report, No. 114, Series 2, Department of Statistics, Princeton University, 1976.

[10] S. Nadarajah, Sums, products and ratios for the bivariate Gumbel distribution, Mathematical and Computer Modeling 42 (2005), 499-518.

[11] S. Nadarajah, Reliability for some bivariate beta distributions, Mathematical Problems in Engineering 1 (2005), 101-111.

[12] S. Nadarajah and S. Kotz, Proportions, sums and ratios, Advances in Statistical Analysis 91 (2007), 93-106.

[13] T. Pham-Gia and N. Turkkan, Bayesian analysis of the difference of two proportions, Communications in Statistics-Theory and Method 22 (1993), 1755-1771.

[14] T. Pham-Gia and N. Turkkan, Reliability of a standby system with beta distributed component lives, IEEE Transactions on Reliability 43 (1994), 71-75.

[15] T. Pham-Gia, Distributions of the ratios of independent beta variables and applications, Communications in Statistics-Theory and Method 29 (2000), 2693-2715.

■