SUMS, PRODUCTS, AND RATIOS FOR THE GENERALIZED BIVARIATE PARETO DISTRIBUTION

SARALEES NADARAJAH AND MARIANO RUIZ ESPEJO

Abstract

We derive the exact distributions of R = X + Y, U = XY and W = X/(X + Y) and the corresponding moment properties when X and Y follow the generalized bivariate Pareto distribution. The expressions turn out to involve special functions.

1. Introduction

Since the 1930s, the statistics literature has seen many developments in the theory and applications of linear combinations and ratios of random variables. Some of these include:

- Ratios of normal random variables appear as sampling distributions in single equation models, in simultaneous equations models, as posterior distributions for parameters of regression models and as modeling distributions, especially in economics when demand models involve the indirect utility function (details in Yatchew, 1986).
- Weighted sums of uniform random variables—in addition to the well known application to the generation of random variables—have applications in stochastic processes which in many cases can be modeled by these weighted sums. In computer vision algorithms these weighted sums play a pivotal role (Kamgar-Parsi et al., 1995). An earlier application of the linear combinations of uniform random variables is given in connection with the distribution of errors in nth tabular differences Δ^n (Lowan and Laderman, 1939).
- Ratio of linear combinations of chi-squared random variables are part of von Neumann's (1941) test statistics (mean square successive difference divided by the variance). These ratios appear in various two-stage tests (Toyoda and Ohtani, 1986). They are also used in tests on structural coefficients of a multivariate linear functional relationship model (details in Chaubey and Nur Enayet Talukder (1983) and Provost and Rudiuk (1994)).
- Sums of independent gamma random variables have applications in queuing theory problems such as determination of the total waiting time and in

civil engineering problems such as determination of the total excess water flow into a dam. They also appear in test statistics used to determine the confidence limits for the coefficient of variation of fiber diameters (Linhart (1965) and Jackson (1969)) and in connection with the inference about the mean of the two-parameter gamma distribution (Grice and Bain, 1980).

- Linear combinations of inverted gamma random variables are used for testing hypotheses and interval estimation based on generalized *p*-values, specifically for the Behrens-Fisher problem and variance components in balanced mixed linear models (Witkovský, 2001).
- As to the Beta distributions their linear combinations occur in calculations of the power of a number of tests in ANOVA (Monti and Sen, 1976) among other applications. More generally, the linear combinations are used for detecting changes in the location of the distribution of a sequence of observations in quality control problems (Lai, 1974). Pham-Gia and Turkkan (1993, 1994, 1998, 2002) and Pham-Gia (2000) provided applications of sums and ratios to availability, Bayesian quality control and reliability.
- Linear combinations of the form $T = a_1 t_{f_1} + a_2 t_{f_2}$, where t_f denotes the Student t random variable based on f degrees of freedom, represents the Behrens-Fisher statistic and—as early as the middle of the twentieth century—Stein (1945) and Chapman (1950) developed a two-stage sampling procedure involving the T to test whether the ratio of two normal random variables is equal to a specified constant.
- Weighted sums of the Poisson parameters are used in medical applications for directly standardized mortality rates (Dobson et al., 1991).

In this paper, we consider the distributions of R = X + Y, U = XY and W = X/(X + Y) when X and Y are correlated Pareto random variables with the joint pdf given by

(1)
$$f(x,y) = \frac{K(xy)^{p-1}}{(\alpha + \beta x + \lambda y + \delta xy)^{p+q}}$$

for x > 0, y > 0, p > 0 and q > 0, where $K = K(\alpha, \beta, \lambda, \delta, p, q)$ denotes the normalizing constant. This distribution is known as the generalized bivariate Pareto distribution. As often with the Pareto distributions, this distribution has applications in reliability studies. Inaba and Shirahata (1986) fitted this distribution to data on white blood counts and survival times of patients who died of acute myelogenous leukemia (Gross and Clark, 1975), comparing it with the bivariate normal distribution.

The aim of this paper is to derive explicit expressions for the pdfs and moments of R = X + Y, U = XY and W = X/(X + Y). The calculations involve the Gauss hypergeometric function defined by

$$G(a,b;c;x) = \sum_{k=0}^{\infty} \frac{(a)_k(b)_k}{(c)_k} \frac{x^k}{k!},$$

where $(e)_k = e(e+1)\cdots(e+k-1)$ denotes the ascending factorial. We also need the following important lemmas.

LEMMA 1. For $0 < \alpha < \rho$,

$$\int_0^\infty \frac{x^{\alpha-1}}{(x+z)^{\rho}} dx = z^{\alpha-\rho} B(\alpha, \rho - \alpha),$$

where

$$B(a,b) = \int_0^1 w^{a-1} (1-w)^{b-1} dw$$

for a > 0 and b > 0 is the beta function.

LEMMA 2. For p > 0 and q > 0,

$$\int_{a}^{b} (x-a)^{p-1} (b-x)^{q-1} (cx+d)^{r} dx$$

$$= (b-a)^{p+q-1} (ac+d)^{r} B(p,q) G\left(p,-r;p+q;\frac{c(a-b)}{ac+d}\right).$$

Lemma 3. For $0 < \alpha < \rho + \lambda$,

$$\int_0^\infty x^{\alpha-1}(x+y)^{-\rho}(x+z)^{-\lambda}\,dx = z^{-\lambda}y^{\alpha-\rho}B(\alpha,\rho+\lambda-\alpha)G\bigg(\alpha,\lambda;\rho+\lambda;1-\frac{y}{z}\bigg).$$

Lemma 4. For
$$a > 0$$
, $b^2 < ac$ and $0 ,
$$\int_0^\infty \frac{x^{p-1}}{(ax^2 + 2bx + c)^\rho} \, dx = a^{-p/2} c^{p/2 - \rho} B(p, 2\rho - p) G\left(\frac{p}{2}, \rho - \frac{p}{2}; \rho + \frac{1}{2}; 1 - \frac{b^2}{ac}\right).$$$

LEMMA 5. For a > b and $0 < \alpha < 2\rho$,

$$\int_0^\infty \frac{x^{\alpha - 1}}{\{(x+a)^2 - b^2\}^{\rho}} dx$$

$$= B(\alpha, 2\rho - \alpha)(a^2 - b^2)^{\alpha/2 - \rho} G\left(\alpha, 2\rho - \alpha; \rho + \frac{1}{2}; \frac{1}{2}\left(1 - \frac{a}{\sqrt{a^2 - b^2}}\right)\right).$$

LEMMA 6. For $0 < |\eta| < \pi$ and $1 < \alpha < 2\rho$,

$$\begin{split} &\int_0^\infty \frac{x^{\alpha-1}}{(x^2+y^2+2xy\cos\eta)^\rho} \, dx \\ &= \frac{1}{2y^{2\rho-\alpha}} \left\{ B\left(\frac{\alpha}{2}, \rho - \frac{\alpha}{2}\right) G\left(\frac{\alpha}{2}, \rho - \frac{\alpha}{2}; \frac{1}{2}; \cos^2\eta\right) \\ &- (\alpha-1)|\cos\eta| B\left(\frac{\alpha-1}{2}, \rho - \frac{\alpha-1}{2}\right) G\left(\frac{\alpha-1}{2}, \rho - \frac{\alpha-1}{2}; \frac{3}{2}; \cos^2\eta\right) \right\}. \end{split}$$

The properties of the Gauss hypergeometric function can be found in Prudnikov et al. (1986) and Gradshteyn and Ryzhik (2000).

2. PDFS

Theorems 1 to 5 derive the pdfs of R = X + Y, U = XY and W = X/(X + Y) when X and Y are distributed according to (1).

Theorem 1. If X and Y are jointly distributed according to (1) and if $\beta = \lambda$ then

(2)
$$f_R(r) = K2^{1-2p}B(p, 1/2)\frac{r^{2p-1}}{(\alpha + \beta r)^{p+q}}G\left(p, p+q; p+\frac{1}{2}; -\frac{\delta r^2}{4(\alpha + \beta r)}\right)$$

for $0 < r < \infty$.

Proof. From (1), the joint pdf of (R, W) = (X + Y, X/R) becomes

(3)
$$f(r,w) = \frac{Kr^{2p-1}\{w(1-w)\}^{p-1}}{\{\alpha + \beta rw + \lambda r(1-w) + \delta r^2 w(1-w)\}^{p+q}}.$$

Thus, the pdf of R can be written as

(4)
$$f_R(r) = Kr^{2p-1} \int_0^1 \frac{\{w(1-w)\}^{p-1}}{\{\alpha + \beta rw + \lambda r(1-w) + \delta r^2 w(1-w)\}^{p+q}} dw$$
$$= Kr^{2p-1} \int_0^{1/4} \frac{u^{p-1} (1/4-u)^{-1/2}}{(\alpha + \beta r + \delta r^2 u)^{p+q}} du$$

after substituting u = w(1 - w). The result of the theorem following applying Lemma 2 to calculate the integral in (4).

THEOREM 2. If X and Y are jointly distributed according to (1) then

(5)
$$f_{U}(u) = K(\beta \lambda)^{-(p+q)/2} B(p+q, p+q) u^{p-1} \times G\left(\frac{p+q}{2}, \frac{p+q}{2}; p+q+\frac{1}{2}; 1 - \frac{(\alpha + \delta u)^{2}}{4\beta \lambda}\right)$$

for $(\alpha + \delta u)^2 < 4\beta\lambda$, and

(6)
$$f_{U}(u) = K(\beta \lambda)^{-(p+q)/2} B(p+q, p+q) u^{(p-q)/2-1} \times G\left(p+q, p+q; p+q+\frac{1}{2}; \frac{1}{2} \left(1 - \frac{\alpha + \delta u}{2\sqrt{\beta \lambda u}}\right)\right)$$

for $(\alpha + \delta u)^2 > 4\beta \lambda u$.

Proof. From (1), the joint pdf of (X, U) = (X, XY) becomes

(7)
$$f(x,u) = \frac{Ku^{p-1}x^{p+q-1}}{\{\beta x^2 + (\alpha + \delta u)x + \lambda u\}^{p+q}}.$$

If $(\alpha + \delta u)^2 > 4\beta \lambda u$ then one can rewrite (7) as

$$f(x,u) = \frac{Ku^{p-1}x^{p+q-1}}{\beta^{p+q}\{(x+a)^2 - b^2\}^{p+q}},$$

where $a=(\alpha+\delta u)/(2\beta)$ and $b^2=a^2-\lambda u/\beta$. Integrating with respect to x using Lemma 5 one obtains the form in (6). On the other hand, if $(\alpha+\delta u)^2<4\beta\lambda$ then integrating (7) with respect to x using Lemma 4 one obtains the form in (5).

THEOREM 3. If X and Y are jointly distributed according to (1) then

(8)
$$f_W(w) = K\alpha^{-q}\delta^{-p}B(2p, 2q)\{w(1-w)\}^{-1}G\left(p, q; p+q+\frac{1}{2}; 1-\frac{\{\beta w+\lambda(1-w)\}^2}{4\alpha\delta w(1-w)}\right)$$

for $\{\beta w + \lambda(1-w)\}^2 < 4\alpha \delta w(1-w)$, and

(9)
$$f_{W}(w) = K\alpha^{-q}\delta^{q}B(2p, 2q)\{w(1-w)\}^{p+q-1} \times G\left(2p, 2q; p+q+\frac{1}{2}; \frac{1}{2}\left(1-\frac{\beta w+\lambda(1-w)}{2\sqrt{\alpha\delta w(1-w)}}\right)\right)$$

for
$$\{(\beta - \lambda)^2 + 4\delta\alpha\}^2 w^2 + \{2(\beta - \lambda) - 4\alpha\delta\}w + \lambda^2 > 0.$$

Proof. If $\{(\beta - \lambda)^2 + 4\delta\alpha\}^2 w^2 + \{2(\beta - \lambda) - 4\alpha\delta\}w + \lambda^2 > 0$ then one can rewrite (3) as

$$f(r,w) = \frac{Kr^{2p-1}}{\delta^{p+q} \{w(1-w)\}^{q+1} \{(r+a)^2 - b^2\}^{p+q}},$$

where $a = \{\beta w + \lambda(1-w)\}/\{2\delta w(1-w)\}$ and $b^2 = a^2 - \alpha/\{\delta w(1-w)\}$. Integrating with respect to r using Lemma 5 one obtains the form in (9). On the other hand, if $\{\beta w + \lambda(1-w)\}^2 < 4\alpha\delta w(1-w)$ then integrating (3) with respect to r using Lemma 4 one obtains the form in (8).

Alternative and equivalent formulas for (5)–(6) and (8)–(9) are given in the following theorems.

THEOREM 4. If X and Y are jointly distributed according to (1) then

 \blacktriangle

$$f_{U}(u) = \frac{Ku^{(p-q)/2-1}}{2(\beta\lambda)^{(p+q)/2}} \left[B\left(\frac{p+q}{2}, \frac{p+q}{2}\right) G\left(\frac{p+q}{2}, \frac{p+q}{2}; \frac{1}{2}; \frac{(\alpha+\delta u)^{2}}{4\beta\lambda u}\right) - \frac{(p+q-1)(\alpha+\delta u)}{2\sqrt{\beta\lambda u}} B\left(\frac{p+q-1}{2}, \frac{p+q+1}{2}\right) \times G\left(\frac{p+q-1}{2}, \frac{p+q+1}{2}; \frac{3}{2}; \frac{(\alpha+\delta u)^{2}}{4\beta\lambda u}\right) \right]$$

for $0 < u < \infty$ and $|\alpha + \delta u| \le 2\sqrt{\beta \lambda u}$.

Proof. Using (7), one can write

$$f_U(u) = \frac{Ku^{p-1}}{\beta^{p+q}} \int_0^\infty \frac{x^{p+q-1}}{\{x^2 + 2xy\cos\eta + y^2\}^{p+q}} dx,$$

where $y = \sqrt{\lambda u/\beta}$ and $2\cos\eta = {\alpha + \delta u}/\sqrt{\beta\lambda u}$. Application of Lemma 6 to calculate this integral yields the result of the theorem.

THEOREM 5. If X and Y are jointly distributed according to (1) then

$$f_{W}(w) = \frac{K}{2\alpha^{q}\delta^{p}w(1-w)} \left[B(p,q)G\left(p,q;\frac{1}{2};\frac{\{\beta w + \lambda(1-w)\}^{2}}{4\alpha\delta w(1-w)}\right) - (2p-1)B\left(p-\frac{1}{2},q+\frac{1}{2}\right) \frac{\beta w + \lambda(1-w)}{2\alpha\sqrt{\delta w(1-w)}} \times G\left(p-\frac{1}{2},q+\frac{1}{2};\frac{3}{2};\frac{\{\beta w + \lambda(1-w)\}^{2}}{4\alpha\delta w(1-w)}\right) \right]$$

for 0 < w < 1 and $|\beta w + \lambda(1 - w)| \le 2\sqrt{\alpha \delta w(1 - w)}$.

Proof. Using (3), one can write

$$f_W(w) = \frac{K}{\delta^{p+q} \{ w(1-w) \}^{q+1}} \int_0^\infty \frac{r^{2p-1}}{\{ r^2 + 2ry \cos \eta + y^2 \}^{p+q}} dr,$$

where $y = \sqrt{\alpha}/\sqrt{\delta w(1-w)}$ and $2\cos\eta = \{\beta w + \lambda(1-w)\}/\sqrt{\alpha\delta w(1-w)}$. Application of Lemma 6 to calculate this integral yields the result of the theorem.

Figures 1 to 3 illustrate the shape of the pdfs (2), (5)-(6) and (8)-(9) for selected values of p and q. Each plot contains four curves corresponding to selected values of q. The effect of the parameters is evident.

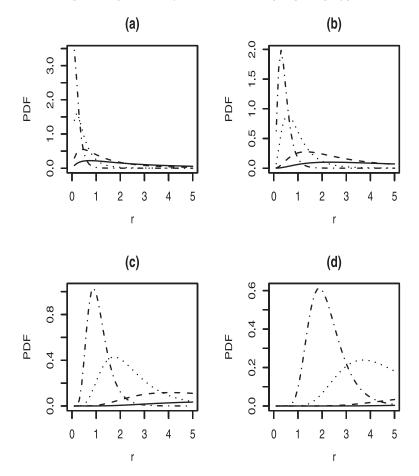


FIGURE 1. Plots of the pdf of (2) for $\alpha=1$, $\beta=1$, $\lambda=1$, $\delta=1$ and (a): p=1; (b): p=2; (c): p=5; and, (d): p=10. The four curves in each plot are: the solid curve (q=1), the curve of lines (q=2), the curve of dots (q=5), and the curve of lines and dots (q=10).

3. Moments

Here, we derive the moments of R = X + Y and U = XY when X and Y are distributed according to (1). We need the following lemma.

Lemma 7. If X and Y are jointly distributed according to (1) then $E(X^mY^n) = \frac{KB(n+p,q-n)B(m+p,q-m)\alpha^{m+n+p-q}}{\beta^{m+p}\lambda^{n+p}}$

$$\times G\left(m+p,n+p;p+q;1-\frac{\alpha\delta}{\beta\lambda}\right)$$

for $m \ge 1$, $n \ge 1$, q > m and q > n.

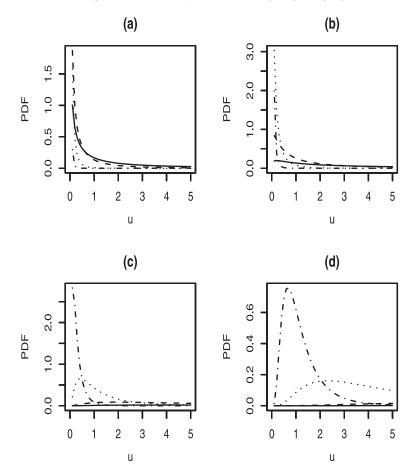


FIGURE 2. Plots of the pdf of (5)–(6) for $\alpha=1$, $\beta=1$, $\lambda=1$, $\delta=1$ and (a): p=1; (b): p=2; (c): p=5; and, (d): p=10. The four curves in each plot are: the solid curve (q=1), the curve of lines (q=2), the curve of dots (q=5), and the curve of lines and dots (q=10).

Proof. The result follows by the arguments

$$\begin{split} E(X^m Y^n) &= \int_0^\infty \int_0^\infty \frac{K x^{m+p-1} y^{n+p-1}}{(\alpha + \beta x + \lambda y + \delta x y)^{p+q}} \, dy dx \\ &= K \int_0^\infty \frac{x^{m+p-1}}{(\lambda + \delta x)^{p+q}} \int_0^\infty \frac{y^{n+p-1}}{\{y + (\alpha + \beta x)/(\lambda + \delta x)\}^{p+q}} \, dy dx \\ &= K \int_0^\infty \frac{x^{m+p-1}}{(\lambda + \delta x)^{p+q}} \left(\frac{\alpha + \beta x}{\lambda + \delta x}\right)^{n-q} B(n+p,q-n) \, dx \end{split}$$

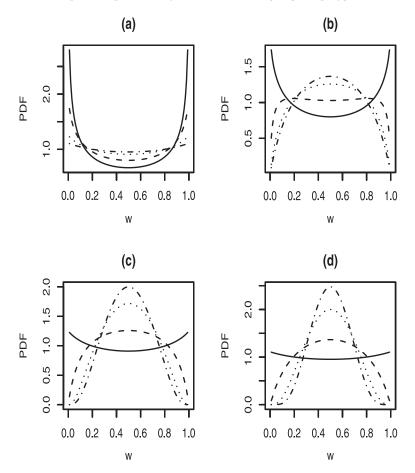


FIGURE 3. Plots of the pdf of (8)–(9) for $\alpha=1$, $\beta=1$, $\lambda=1$, $\delta=1$ and (a): p=1; (b): p=2; (c): p=5; and, (d): p=10. The four curves in each plot are: the solid curve (q=1), the curve of lines (q=2), the curve of dots (q=5), and the curve of lines and dots (q=10).

$$\begin{split} &=KB(n+p,q-n)\int_{0}^{\infty}\frac{x^{m+p-1}(\alpha+\beta x)^{n-q}}{(\lambda+\delta x)^{n+p}}\,dx\\ &=\frac{KB(n+p,q-n)B(m+p,q-m)\alpha^{m+n+p-q}}{\beta^{m+p}\lambda^{n+p}}\\ &\quad\times G\bigg(m+p,n+p;p+q;1-\frac{\alpha\delta}{\beta\lambda}\bigg), \end{split}$$

where we have applied Lemmas 1 and 3 (in the third and the fifth steps, respectively).

The moments of R = X + Y and U = XY are now simple consequences of this lemma as illustrated in Theorems 6 and 7.

THEOREM 6. If X and Y are jointly distributed according to (1) then

$$E(R^n) = K\alpha^{n+p-q} \sum_{k=0}^n \binom{n}{k} \frac{B(k+p,q-k)B(n-k+p,q-n+k)}{\beta^{n-k+p}\lambda^{k+p}}$$
$$\times G\left(n-k+p,k+p;p+q;1-\frac{\alpha\delta}{\beta\lambda}\right)$$

for $n \ge 1$ and q > n.

Proof. the result follows by writing

$$E((X + Y)^{n}) = \sum_{k=0}^{n} {n \choose k} E(X^{n-k} Y^{k})$$

and applying Lemma 7 to each expectation in the sum.

THEOREM 7. If X and Y are jointly distributed according to (1) then

$$E(U^n) = \frac{KB(n+p,q-n)B(n+p,q-n)\alpha^{2n+p-q}}{\beta^{n+p}\lambda^{n+p}}G\left(n+p,n+p;p+q;1-\frac{\alpha\delta}{\beta\lambda}\right)$$

for $n \ge 1$ and q > n.

Proof. follows by writing $E(U^n) = E(X^n Y^n)$ and applying Lemma 7 with m = n.

Acknowledgments. The authors would like to thank the referee and the managing editor for carefully reading the paper and for their great help in improving the paper.

REFERENCES

- D. G. CHAPMAN, Some two-sample tests, Annals of Mathematical Statistics 21 (1950), 601-606.
- Y. P. CHAUBEY AND A. B. M. NUR ENAYET TALUKDER, Exact moments of a ratio of two positive quadratic forms in normal variables, Communications in Statistics, Theory and Methods 12 (1983), 675–679.
- A. J. DOBSON, K. KULASMAA AND J. SCHERER, Confidence intervals for weighted sums of Poisson parameters, Statistics in Medicine 10 (1991), 457–462.
- I. S. Gradshteyn and I. M. Ryzhik, Table of integrals, series, and products, 6th ed., Academic Press, San Diego, 2000.

- J. V. GRICE AND L. J. BAIN, Inferences concerning the mean of the gamma distribution, Journal of the American Statistical Association 75 (1980), 929-933.
- A. J. GROSS AND V. A. CLARK, Survival distributions: reliability applications in the biomedical sciences, John Wiley and Sons, New York, 1975.
- T. INABA AND S. SHIRAHATA, Measures of dependence in normal models and exponential models by information gain, Biometrika 73 (1986), 345–352.
- O. A. Y. Jackson, Fitting a gamma or log-normal distribution to fibre-diameter measurements on wool tops, Applied Statistics 18 (1969), 70–75.
- B. KAMGAR-PARSI, B. KAMGAR-PARSI AND M. BROSH, Distribution and moments of weighted sum of uniform random variables with applications in reducing Monte Carlo simulations, Journal of Statistical Computation and Simulation 52 (1995), 399–414.
- S. KOTZ, N. BALAKRISHNAN AND N. L. JOHNSON, Continuous multivariate distributions 1, Models and applications, 2nd ed., John Wiley and Sons, New York, 2000.
- T. L. LAI, Control charts based on weighted sums, Annals of Statistics 2 (1974), 134-147.
- H. LINHART, Approximate confidence limits for the coefficient of variation of gamma distributions, Biometrics 21 (1965), 733–738.
- A. N. LOWAN AND J. LADERMAN, On the distribution of errors in nth tabular differences, Annals of Mathematical Statistics 10 (1939), 360–364.
- H. J. Malik, The distribution of the product of two noncentral beta variates, Naval Research Logistics Quarterly 17 (1970), 327–330.
- N. N. MIKHAIL AND D. S. TRACY, The exact non-null distribution of Wilk's Λ criterion in the bivariate collinear case, Canadian Mathematical Bulletin 17 (1975), 757–758.
- K. L. Monti and P. K. Sen, The locally optimal combination of independent test statistics, Journal of the American Statistical Association 71 (1976), 903–911.
- J. VON NEUMANN, Distribution of the ratio of the mean square successive difference to the variance, Annals of Mathematical Statistics 12 (1941), 367-395.
- T. PHAM-GIA, Distributions of the ratios of independent beta variables and applications, Communications in Statistics, Theory and Methods 29 (2000), 2693–2715.
- T. Pham-Gia and N. Turkkan, Bayesian analysis of the difference of two proportions, Communications in Statistics, Theory and Methods 22 (1993), 1755–1771.
- T. Pham-Gia and N. Turkkan, Reliability of a standby system with beta component lifelength, IEEE transactions on reliability, 1994, 71–75.
- T. PHAM-GIA AND N. TURKKAN, Distribution of the linear combination of two general beta variables and applications, Communications in Statistics, Theory and Methods 27 (1998), 1851– 1869
- T. Pham-Gia and N. Turkkan, The product and quotient of general beta distributions, Statistical Papers 43 (2002), 537–550.
- S. B. PROVOST AND E. M. RUDIUK, The exact density function of the ratio of two dependent linear combinations of chi-square variables, Annals of the Institute of Statistical Mathematics 46 (1994), 557–571.
- A. P. Prudnikov, Y. A. Brychkov and O. I. Marichev, Integrals and series 1, 2 and 3, Gordon and Breach Science Publishers, Amsterdam, 1986.
- C. STEIN, A two-sample test for a linear hypothesis whose power is independent of the variance, Annals of Mathematical Statistics 16 (1945), 243–258.
- T. TOYODA AND K. OHTANI, Testing equality between sets of coefficients after a preliminary test for equality of disturbance variances in two linear regressions, Journal of Econometrics 31 (1986), 67–80.
- V. WITKOVSKÝ, Computing the distribution of a linear combination of inverted gamma variables, Kybernetika 37 (2001), 79–90.

A. J. Yatchew, Multivariate distributions involving ratios of normal variables, Communications in Statistics, Theory and Methods 15 (1986), 1905–1926.

DEPARTMENT OF STATISTICS UNIVERSITY OF NEBRASKA LINCOLN NEBRASKA 68583 USA

DEPARTAMENTO DE MATEMÁTICAS FUNDAMENTALES UNED

Madrid

Spain