Revista Integración
Escuela de Matemáticas
Universidad Industrial de Santander
Vol. 28, No. 2, 2010, pág. 101–109

Distribución exacta y aproximada del producto de variables Kummer-beta

Fabio Humberto Sepúlveda Murillo*

Instituto Tecnológico Metropolitano (ITM), Facultad de Ciencias, Medellín, Colombia.

Resumen. Se presenta la distribución exacta para el producto $Y = X_1X_2$, cuando X_1 y X_2 son variables aleatorias correlacionadas con distribución conjunta Kummer-beta. Se propone una distribución aproximada mostrando el desempeño de su ajuste.
Palabras claves: Distribución Kummer-beta, función hipergeométrica confluente, función hipergeométrica de Gauss, momentos.
MSC2000: Primaria 62H15; secundaria 62E15.

Exact and approximate distributions of the product of Kummer-beta variables

Abstract. We present the exact distribution for the product $Y = X_1X_2$, where X_1 and X_2 are random variables with combined Kummer-beta distribution. We propose an approximate distribution exhibiting the performance of its adjustment.
Keywords: Kummer-beta distribution, confluent hypergeometric function, Gauss hypergeometric function, moments.

1. Introducción

La distribución de combinaciones lineales, del cociente y del producto de variables aleatorias surgen en una gran variedad de contextos probabilísticos y estadísticos. Una buena exposición de aplicaciones y otros aspectos interesantes de estas distribuciones pueden encontrarse en Sornette [10] y Witkosvsky [11].

La distribución del producto de dos variables aleatorias independientes, X_1 y X_2, ha sido estudiada por varios autores, por ejemplo, Nagar y Zarrazola [4],

*Autor para correspondencia: *E-mail*: `fahum65@hotmail.com`
Recibido: 26 de octubre de 2010, aceptado: 10 de diciembre de 2010.

Pham-Gia y Turkkan [6], Fan [1], Podolski [7], Rafhie y Rohrer [8], Sakamoto [9], entre otros. Sin embargo, existen pocas investigaciones sobre distribuciones del producto de dos variables aleatorias X_1 y X_2 correlacionadas. El único trabajo conocido es de los autores Garg, Katta y Grupta [2] para la familia Dirichlet.

La familia de distribuciones Kummer-beta fue propuesta inicialmente por Ng y Kotz [5]. Se dice que una variable aleatoria X tiene distribución Kummer-beta, si la función de densidad de probabilidad es proporcional a

$$x^{\alpha-1}(1-x)^{\beta-1}\exp(-\lambda x)({}_1F_1(\alpha;\alpha+\beta;-\lambda))^{-1}, \quad \text{donde } 0<x<1.$$

Se dice que el vector aleatorio $\mathbf{X}=(X_1,X_2)$ tiene una distribución conjunta Kummer-beta, la cual es denotada por $(X_1,X_2)\sim \mathbf{KB}(\alpha_1,\alpha_2,\beta,\lambda)$, si su función de densidad de probabilidad conjunta es dada por:

$$f(x_2,x_2)=\frac{\Gamma(\alpha_1+\alpha_2+\beta)x_1^{\alpha_1-1}x_2^{\alpha_2-1}(1-x_1-x_2)^{\beta-1}\exp(-\lambda(x_1+x_2))}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\beta){}_1F_1(\alpha_1+\alpha_2;\alpha_1+\alpha_2+\beta;-\lambda)}, \quad (1)$$

donde $0<x_1<1$, $0<x_2<1$, $x_1+x_2<1$, α_1, $\alpha_2>0$, $\beta>0$, $-\infty<\lambda<\infty$ y ${}_1F_1$ es la función hipergeométrica confluente.

El objetivo del presente estudio es determinar la distribución del producto $Y=X_1X_2$ cuando X_1 y X_2 tienen una distribución conjunta Kummer-beta. El artículo se divide en cuatro secciones, incluyendo la presente introducción. En la segunda sección se presentan las definiciones y algunos resultados necesarios para obtener la función de densidad de probabilidad exacta del producto $Y=X_1X_2$ y sus respectivos momentos, cuando X_1 y X_2 tienen una distribución conjunta Kummer-beta, la cual es presentada en la tercera sección. Finalmente, en la cuarta sección se da una distribución aproximada para la distribución de probabilidad exacta de $Y=X_1X_2$ y se muestra gráficamente el comportamiento de su ajuste.

2. Definiciones y resultados preliminares

Definición 2.1. La función hipergeométrica generalizada de argumento escalar está definida por

$${}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q;x)=\sum_{k=0}^{\infty}\frac{(a_1)_k\cdots(a_p)_k}{(b_1)_k\cdots(b_q)_k}\frac{x^k}{k!}, \quad (2)$$

donde a_i, $i=1,\ldots,p$; b_j , $j=1,\ldots,q$ son números complejos con restricciones, x es una variable compleja y el símbolo de Pochammer $(c)_n$ es definido por $(c)_n=c(c+1)\cdots(c+n-1)$ para $n=1,2,\ldots,$ y $(c)_0=1$.

Definición 2.2. Las representaciones integrales de la función hipergeométrica confluente y la función hipergeométrica de Gauss están dadas por

$$ {}_1F_1(a;c;x) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(c-a)} \int_0^1 t^{\alpha-1}(1-t)^{c-a-1} \exp(xt)dt \tag{3} $$

y

$$ {}_2F_1(a,b;c;x) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(c-a)} \int_0^1 t^{\alpha-1}(1-t)^{c-a-1}(1-xt)^{-b}dt, \tag{4} $$

donde $\text{Re}(c) > \text{Re}(a) > 0$ y $|\arg(1-x)| < \pi$, respectivamente.

Nótese que la función hipergeométrica confluente ${}_1F_1$ satisface la relación de Kummer dada por

$$ {}_1F_1(a;c;-x) = \exp(-x){}_1F_1(c-a;c;x). \tag{5} $$

Definición 2.3. Sea $f(\cdot)$ una función continua y $\alpha_i > 0$, $i = 1, \ldots, n$. La integral

$$ D_n(\alpha_1, \ldots, \alpha_n; f) = \underset{\substack{\sum_{i=1}^n x_i < 1 \\ x_i > 0,\ i=1,\ldots,n}}{\int \cdots \int} \prod_{i=1}^n x_i^{\alpha_i - 1} f\left(\sum_{i=1}^n x_i\right) \prod_{i=1}^n dx_i \tag{6} $$

es conocida como la integral de Liouville-Dirichlet.

Sustituyendo en la ecuación (6) a $y_i = x_i/x$, $i = 1, \ldots, n-1$, y $x = \sum_{i=1}^n x_i$ con el jacobiano $J(x_1, \ldots, x_n - 1, x_n \to y_1, \ldots, y_n - 1, x) = x^{n-1}$, es fácil ver que

$$ D_n(\alpha_1, \ldots, \alpha_n; f) = \frac{\prod_{i=1}^n \Gamma(\alpha_i)}{\Gamma\left(\sum_{i=1}^n \alpha_i\right)} \int_0^1 x^{\sum_{i=1}^n \alpha_i - 1} f(x) dx. \tag{7} $$

A pesar de que se pueden encontrar en [3], presentamos aquí las definiciones de las distribuciones Beta tipo 1 y Beta tipo 2.

Definición 2.4. Una variable aleatoria X tiene una distribución Beta tipo 1 con parámetros (a, b), denotada por $X \sim B_1(a, b)$, si su función de densidad de probabilidad es

$$ f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}, \qquad 0 < x < 1. $$

Definición 2.5. Una variable aleatoria X tiene una distribución Beta tipo 2 con parámetros (a, b), denotada por $X \sim B_2(a, b)$, si su función de densidad de probabilidad es

$$ f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{-(a+b)}, \qquad x > 0. $$

3. Distribución exacta del producto $Y = X_1X_2$

El producto de dos variables aleatorias X_1, X_2 definidas sobre un mismo espacio de probabilidades $(\Omega, \mathcal{A}, P)$, es definido como $(X_1X_2)(\omega) = X_1(\omega)X_2(\omega)$ para cada $\omega \in \Omega$. En el siguiente teorema se presenta la función de densidad de probabilidad del producto $Y = X_1X_2$, cuando (X_1, X_2) tienen una distribución conjunta Kummer-beta.

Teorema 3.1. *Si* $(X_1, X_2) \sim \mathbf{KB}(\alpha_1, \alpha_2, \beta, \lambda)$, *entonces la función de densidad de probabilidad de* $Y = X_1X_2$ *está dada por*

$$\begin{aligned} f(y) = {} & \frac{\Gamma(\alpha_1+\alpha_2+\beta)(1-4y)^{\beta-1/2}(1+\sqrt{1-4y})^{\alpha_1-\alpha_2-\beta}y^{\alpha_2-1}}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\beta)_1F_1(\beta;\alpha_1+\alpha_2+\beta;\lambda)} \\ & \times \sum_{k=0}^{\infty} \frac{(2\lambda)^k(1-4y)^k}{(1+\sqrt{1-4y})^k k!} \, \frac{\Gamma(\beta+k)^2}{\Gamma(2\beta+2k)} \, {}_2F_1(\beta+k, \alpha_2+\beta+k-\alpha_1; 2\beta+2k; z), \end{aligned} \tag{8}$$

donde

$$z = \frac{2\sqrt{1+4y}}{1+\sqrt{1+4y}}, \qquad 0 < y < 1/4, \ -\infty < \lambda < \infty, \ \lambda \neq 0.$$

Demostración. Sea $x_1 = x_1$ y $x_2 = y/x_1$, cuyo jacobiano es $J(x_1, x_2 \to x_1, y) = 1/x_1$. Entonces la función de densidad de probabilidad conjunta de x_1 y y es

$$f(x_1, y) = kx_1^{\alpha_1-\alpha_2-\beta}y^{\alpha_2-1}(x_1 - x_1^2 - y)^{\beta-1}\exp(-\lambda(x_1 + y/x_1)), \tag{9}$$

donde

$$k = \frac{\Gamma(\alpha_1+\alpha_2+\beta)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\beta)\,{}_1F_1(\alpha_1+\alpha_2;\alpha_1+\alpha_2+\beta;-\lambda)}, \qquad x_1 + \frac{y}{x_1} < 1.$$

Utilizando la relación de Kummer (5) podemos escribir la expresión (9) de la forma

$$f(x_1, y) = k_1x_1^{\alpha_1-\alpha_2-\beta}y^{\alpha_2-1}(x_1 - x_1^2 - y)^{\beta-1}\exp\left(\lambda\left(\frac{x_1 - x_1^2 - y}{x_1}\right)\right), \tag{10}$$

donde

$$k_1 = \frac{\Gamma(\alpha_1+\alpha_2+\beta)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\beta)\,{}_1F_1(\beta;\alpha_1+\alpha_2+\beta;\lambda)}.$$

Como $x_1 - x_1^2 - y > 0$, la solución está en el intervalo (p_1, p_2), donde $p_1 = \frac{1-\sqrt{1-4y}}{2}$ y $p_2 = \frac{1+\sqrt{1-4y}}{2}$. Integrando la expresión (10) con respecto a x_1, obtenemos la distribución marginal de y:

$$f(y) = k_1 y^{\alpha_2-1} \int_{p_1}^{p_2} x_1^{\alpha_1-\alpha_2-\beta}[(x_1-p_1)(p_2-x_1)]^{\beta-1} \exp\left(\frac{\lambda(x_1-p_1)(p_2-x_1)}{x_1}\right) dx_1.$$

Expandiendo la función exponencial en términos de series de potencia, se tiene que

$$\begin{aligned} f(y) &= k_1(p_2-p_1)^{2\beta-1} p_2^{\alpha_1-\alpha_2-\beta} y^{\alpha_2-1} \int_0^1 t^{\beta-1}(1-t)^{\beta-1}(1-zt)^{\alpha_1-\alpha_2-\beta} \times \\ &\quad \times \sum_{k=0}^{\infty} \frac{(\lambda t(1-t))^k (p_2-p_1)^{2k}}{(p_2(1-zt))^k k!} \\ &= k_1(p_2-p_1)^{2\beta-1} p_2^{\alpha_1-\alpha_2-\beta} y^{\alpha_2-1} \sum_{k=0}^{\infty} \frac{\lambda^k (p_2-p_1)^{2k}}{p_2^k k!} \times \\ &\quad \times \int_0^k t^{\beta+k-1}(1-t)^{\beta+k-1}(1-zt)^{\alpha_1-\alpha_2-\beta-k} dy \end{aligned}$$

Finalmente, utilizando la representación integral (4) y reemplazando k_1, p_1, p_2 y z, completamos la demostración deseada. ☑

3.1. Momentos

Primero derivamos el r-ésimo momento para una variable con distribución Kummer-beta, $X \sim \mathbf{KB}(\alpha, \beta, \lambda)$ y los momentos conjuntos de X_1 y X_2, como base para derivar los momentos de $Y = X_1 X_2$.

Teorema 3.2. *Si* $X \sim \mathbf{KB}(\alpha, \beta, \lambda)$*, entonces*

$$E(X^r) = \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+r)\, {}_1F_1(\beta; \alpha+\beta+r; \lambda)}{\Gamma(\alpha)\Gamma(\alpha+\beta+r)\, {}_1F_1(\beta; \alpha+\beta; \lambda)}. \tag{11}$$

Demostración. De la función de densidad de X tenemos que

$$E(X^r) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)\, {}_1F_1(\alpha; \alpha+\beta; -\lambda)} \int_0^1 x^{\alpha+r-1}(1-x)^{\beta-1} \exp(-\lambda x)\, dx.$$

De forma simple, evaluamos la integral usando (3) y luego simplificamos la expresión aplicando el resultado (5). ☑

Teorema 3.3. *Si* $(X_1, X_2) \sim \mathbf{KB}(\alpha_1, \alpha_2, \beta, \lambda)$*, entonces*

$$E(X_1^{r_1} X_2^{r_2}) = \frac{\Gamma(\alpha_1+r_1)\Gamma(\alpha_2+r_2)\Gamma(\alpha_1+\alpha_2+\beta)\, {}_1F_1(\beta; \alpha_1+\alpha_2+\beta+r_1+r_2; \lambda)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_1+\alpha_2+\beta+r_1+r_2)\, {}_1F_1(\beta; \alpha_1+\alpha_2+\beta; \lambda)}.$$

Demostración. De la función de densidad conjunta de (X_1, X_2) tenemos que

$$E(X_1^{r_1} X_2^{r_2}) = \frac{\Gamma(\alpha_1+\alpha_2+\beta)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\beta)\,{}_1F_1(\alpha_1+\alpha_2;\alpha_1+\alpha_2+\beta;-\lambda)} \times \int\cdots\int_{\substack{x_1+x_2<1\\ x_1,\ x_2>0}} \prod_{i=1}^{2} x_i^{\alpha_1+\alpha_2-1}\left(1-\sum_{i=1}^{2} x_i\right)\prod_{i=1}^{2} dx_i.$$

Este teorema se obtiene de forma directa, aplicando la definición de la integral de Liouville-Dirichlet (6) y el Teorema 3.2. ☑

Teorema 3.4. *Si* $(X_1, X_2) \sim \mathbf{KB}(\alpha_1, \alpha_2, \beta, \lambda)$*, entonces*

$$E(Y^h) = \frac{\Gamma(\alpha_1+h)\Gamma(\alpha_2+h)\Gamma(\alpha_1+\alpha_2+\beta)\,{}_1F_1(\beta;\alpha_1+\alpha_2+\beta+2h;\lambda)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_1+\alpha_2+\beta+2h)\,{}_1F_1(\beta;\alpha_1+\alpha_2+\beta;\lambda)}. \tag{12}$$

Demostración. La prueba es inmediata usando el Teorema 3.3. ☑

Usando propiedades de la función gamma la expresión (12) se puede escribir como

$$E(Y^h) = \frac{\alpha_1(\alpha_2+1)\cdots(\alpha_1+h-1)\alpha_2(\alpha_2+1)\cdots(\alpha_2+h-1)}{(\alpha_1+\alpha_2+\beta)(\alpha_1+\alpha_2+\beta+1)\,{}_1F_1(\beta;\alpha_1+\alpha_2+\beta;\lambda)} \times \times \frac{{}_1F_1(\beta;\alpha_1+\alpha_2+\beta+2h;\lambda)}{{}_1F_1(\beta;\alpha_1+\alpha_2+\beta;\lambda)}.$$

En particular, si $h = 1$ se tiene que

$$E(Y) = \frac{\alpha_1\alpha_2\,{}_1F_1(\beta;\alpha_1+\alpha_2+\beta+2;\lambda)}{(\alpha_1+\alpha_2+\beta)(\alpha_1+\alpha_2+\beta+1)\,{}_1F_1(\beta;\alpha_1+\alpha_2+\beta;\lambda)}, \tag{13}$$

y cuando $h = 2$,

$$E(Y^2) = \frac{\alpha_1(\alpha_1+1)\alpha_2(\alpha_2+1)\,{}_1F_1(\beta;\alpha_1+\alpha_2+\beta+4;\lambda)({}_1F_1(\beta;\alpha_1+\alpha_2+\beta;\lambda))^{-1}}{(\alpha_1+\alpha_2+\beta)(\alpha_1+\alpha_2+\beta+1)(\alpha_1+\alpha_2+\beta+2)(\alpha_1+\alpha_2+\beta+3)}. \tag{14}$$

4. *Aproximación a la distribución exacta*

Sabemos que si $1/(1+Y) \sim B_1(a,b)$, entonces $Y \sim B_2(b,a)$, es decir, Y tiene función de densidad de probabilidad dada por

$$f(y) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\, y^{b-1}(1+y)^{-(a+b)}, \quad y>0,\ a,b>0, \tag{15}$$

donde a y b son parámetros estimados utilizando el método de los momentos.

Por otro lado, si igualamos los dos primeros momentos de Y, ecuaciones (13) y (14) con los de la distribución Beta tipo 2, tenemos

$$E(Y) = \frac{b}{a-1}$$

y

$$E(Y^2) = \frac{b(b+1)}{(a-1)(a-2)}.$$

Resolviendo simultáneamente las dos expresiones anteriores para hallar los dos parámetros a y b se tiene que

$$a = 1 + \frac{E(Y^2) + E(Y)}{E(Y^2) - E^2(Y)} \tag{16}$$

y

$$b = E(Y)\,\frac{E(Y^2) + E(Y)}{E(Y^2) - E^2(Y)}. \tag{17}$$

Los dos momentos $E(Y)$ y $E(Y^2)$ se calculan de (13) y (14), para valores dados de los parámetros α_1, α_2, β y λ.

Para visualizar el ajuste de la aproximación seleccionamos 8 valores para los parámetros α_1, α_2, β y λ, y posteriormente calculamos los respectivos valores de a y b usando (16) y (17).

Gráfico	α_1	α_2	β	λ	a	b
1	1	0,5	3	-1	25,5092	0,6051
2	0,5	0,5	0,5	1	14,9758	0,7458
3	0,5	1	1	1	18,0085	0,7487
4	0,5	1	3	1	28,3630	0,4489
5	0,5	0,5	0,5	2	14,3430	0,5342
6	3	0,5	1	2	19,0668	0,9254
7	2	0,5	0,5	3	16,8336	0,9616
8	0,5	0,5	0,5	0,5	15,4542	0,8691

Tabla 1. Valores para los parámetros.

Los valores para cada parámetro se muestran en la Tabla 1. Luego comparamos gráficamente las gráficas de las funciones de densidad de probabilidad exacta y aproximada de Y dadas por (8) y (15), respectivamente.

Observando los diferentes gráficos no se notan grandes desviaciones que generen dudas sobre el buen ajuste que tiene la distribución aproximada a la distribución exacta.

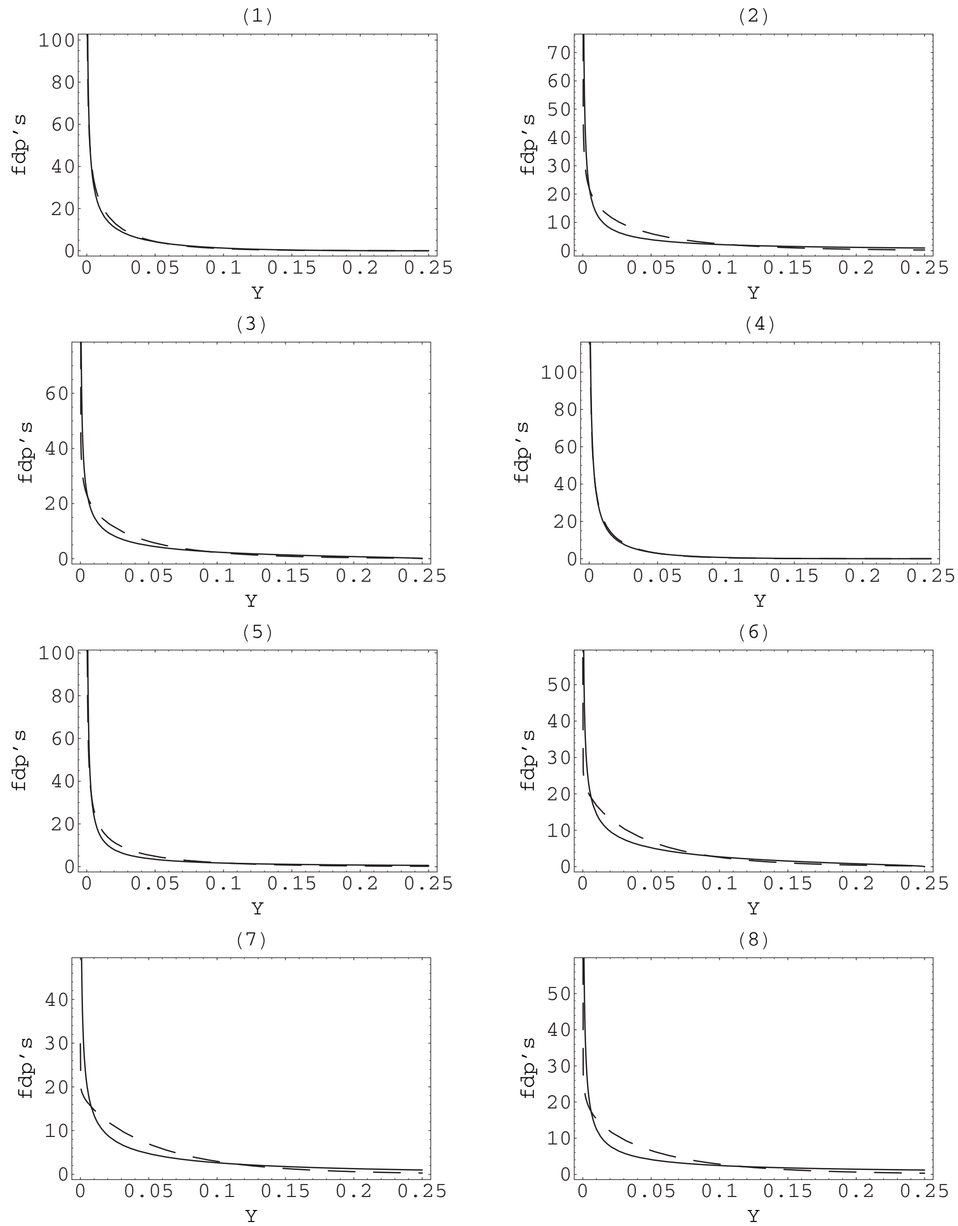

Figura 1. La fdp exacta (curva sólida) y la fdp aproximada (curva punteada) de $Y = X_1X_2$.

Agradecimiento

Este trabajo fue soportado por el Comité para el Desarrollo de la Investigación del Instituto Tecnológico Metropolitano de Medellín, ITM.

Referencias

[1] Fan D. Y., "The distribution of the product of independent beta variables", *Communications Statistics-Theory Methods*, 20 (1991), 4043–4052.

[2] Garg M., Katta V., Gupta M.K., "Thde distribution of the products of powers of generalized Dirochlet components", *Kyungpook Math. J.*, 42 (2002), 429–436.

[3] Johnson N. L., Kotz S., and Balakrishnan N., *Continuous Univariate Distributions*, Vol. 2, Second Edition, John Wiley & Sons, New York, 1995.

[4] Nagar D. K., and Zarrazola E., "Distributions of the product and the quotient of independent Kummer-beta variables", *Scientiae Mathematicae Japonicae*, 61 (2005), 109–117.

[5] Ng K. W., and Kotz S., *Kummer-gamma and Kummer-beta univariate and multivariate distributions*, Research Report, no. 84, Department of Statistics, The University of Hong Kong, Hong Kong, 1995.

[6] Pham-Gia T., and Turkkan N., "The product and quotient of general beta distributions", *Statistical Paper*, 43 (2002), 537–550.

[7] Podolski H., "The distribution of a product of n independent random variables with generalized gamma distribution", *Demonstration Mathematica*, 4 (1972), 119–123.

[8] Rathie P. N., and Rohrer H. G., "The exact distribution of products of independent random variables", *Metron*, 45 (1987), 235–245.

[9] Sakamoto H., "On the distributions of the product and quotient of the independent and uniformly distributed random variables", *Tohoku Mathematical Journal*, 49 (1943), 243–260.

[10] Sornette D., "Multiplicative processes and power laws", *Physical Review E*, 57 (1998), 4811–4813.

[11] Witkovsky V., "Computing the distribution of a linear combination of inverted gamma variables", *Kybernetika*, 37 (2001), 79–90.