

## Distribución Dirichlet-hipergeométrica invertida

Paula A. Bran Cardona  
Universidad del Valle

Recibido Oct. 26, 2009      Aceptado Mar. 11, 2010

### Abstract

The inverted Dirichlet-hypergeometric function distribution (IDH) was introduced by Nagar and Bran (2009) to generalize the inverted Dirichlet distribution (Tiao and Cuttman 1965). In this paper, we derive the marginal, conditional and partial sums densities of the IDH. We also derive the joint moments of it. Furthermore, we show how to derive its density from the Dirichlet distribution. Finally, we use the Fisher's scoring method for the estimation of parameters (Kotz et al 2000).

**Keywords:** Hypergeometric function, beta distribution, moments, likelihood function, Fisher's scoring method

**MSC(2000):** 62E15, 60E05

### Resumen

Nagar y Bran (2009) definieron la distribución Dirichlet-hipergeométrica invertida (IDH), para generalizar la distribución Dirichlet invertida (Tiao y Cuttman 1965). En este artículo, se derivan las densidades marginales, condicionales y de sumas parciales de la IDH. Se obtienen también sus momentos conjuntos. Además, se muestra cómo se puede derivar su densidad a partir de la distribución Dirichlet. Finalmente, se utiliza el método scoring de Fisher para estimar los parámetros de la distribución (Kotz et al 2000).

**Palabras y frases claves:** Función hipergeométrica, distribución beta, momentos, función de verosimilitud, método Fisher scoring

## 1 Introducción

La siguiente notación es tomada de [9], donde se mostró que  $(Y_1, \dots, Y_n)$  tiene distribución Dirichlet-hipergeométrica invertida, denotada por

$$IDH(\nu_1, \dots, \nu_n; \alpha, \beta, \gamma)$$

si su función de densidad es proporcional a

$$\left( \prod_{i=1}^n y_i^{\nu_i-1} \right) \left( 1 + \sum_{i=1}^n y_i \right)^{-(\nu+\gamma)} {}_2F_1 \left( \alpha, \beta; \gamma; \frac{1}{1 + \sum_{i=1}^n y_i} \right), \quad y_i > 0,$$

con constante de proporcionalidad dada por

$$C(\nu_1, \dots, \nu_n; \alpha, \beta, \gamma) = \frac{\Gamma(\nu + \gamma - \alpha)\Gamma(\nu + \gamma - \beta)}{\prod_{i=1}^n \Gamma(\nu_i)\Gamma(\gamma)\Gamma(\nu + \gamma - \alpha - \beta)},$$

donde  $\nu_i > 0$ ,  $i = 1, \dots, n$ ,  $\nu = \sum_{i=1}^n \nu_i$ ,  $\gamma > 0$ ,  $\alpha > 0$ ,  $\beta > 0$ ,  $\nu + \gamma > \alpha + \beta$  y  ${}_2F_1$  es un caso particular de la función hipergeométrica generalizada dada por

$${}_pF_q = \sum_{k=0}^{\infty} \frac{(a_1)_k \dots (a_p)_k}{(b_1)_k \dots (b_p)_k} \frac{z^k}{k!}, \quad (1)$$

donde  $a_i$ ,  $i = 1, 2, \dots, p$ ,  $b_j$ ,  $j = 1, 2, \dots, q$  son números complejos,  $z$  es una variable compleja y el símbolo de Pochhammer  $(a)_n$  está definido por  $(a)_n = a(a+1)\dots(a+n-1)$  para  $n = 1, 2, \dots$  y  $(a)_0 = 1$ . Para más detalles de esta función consultar [7]. Esta distribución es una generalización de la distribución Dirichlet invertida [2], [11] y, en vista de ello, puede abarcar problemas más generales, pues esto implica el uso de un número mayor de parámetros. Por otro lado, es importante destacar que en este artículo, se propone un nuevo miembro de la familia de distribuciones multivariadas de Liouville [6]. Este tipo de distribuciones ha atraído la atención de los investigadores durante los últimos treinta años, motivados principalmente por sus consideraciones teóricas, más que por su campo de aplicación. Aunque estas distribuciones se pueden encontrar en varios contextos, como el análisis bayesiano, estadísticos de orden, teoría de la confiabilidad, procesos estocásticos, entre otros. Algunas aplicaciones y otros aspectos se pueden consultar en Gupta and Song [3], Gupta and Richards [4], Marshall and Olkin [8], Song and Gupta [10]. En este artículo, se presentan algunas de las propiedades de esta distribución multivariada, las cuales no habían sido estudiadas. En la Sección 2, se deriva su función de densidad a partir de la distribución Dirichlet. Además, se muestran propiedades como distribuciones marginales, condicionales y de sumas parciales; y también, algunas representaciones a partir de las distribuciones beta tipo I y beta tipo II (o invertida). Los momentos conjuntos se encuentran en la Sección 3, haciendo uso de la representación que se dio a partir de la distribución Dirichlet. Para finalizar, en la Sección 4 se hace una simulación para estimar los parámetros mediante una optimización numérica de la correspondiente función de verosimilitud, utilizando el método Fisher scoring.

## 2 Propiedades

De acuerdo con las propiedades de la familia de distribuciones de Liouville, se tiene que sus miembros se pueden obtener como el producto de un vector aleatorio con distribución Dirichlet [6] y una variable aleatoria univariada, cuya distribución pertenezca a la familia de distribuciones de Liouville. En este caso en particular, esa distribución univariada es la distribución de la función hipergeométrica invertida<sup>1</sup> (Gupta and Nagar [2]), que se define a continuación.

**Definición 1.** *Se dice que  $X$  tiene distribución de la función hipergeométrica invertida, denotada por  $X \sim IH(\nu, \alpha, \beta, \gamma)$ , si su función de densidad está dada*

<sup>1</sup>Inverted hypergeometric function distribution

por

$$\frac{\Gamma(\nu + \gamma - \alpha)\Gamma(\nu + \gamma - \beta)}{\Gamma(\gamma)\Gamma(\nu)\Gamma(\nu + \gamma - \alpha - \beta)} x^{\nu-1} (1+x)^{-(\nu+\gamma)} {}_2F_1\left(\alpha, \beta; \gamma; \frac{1}{1+x}\right), \quad x > 0.$$

donde  $\nu > 0$ ,  $\gamma > 0$ ,  $\alpha > 0$ ,  $\beta > 0$ ,  $\nu + \gamma > \alpha + \beta$  y  ${}_2F_1$  definida en (1).

Ahora, se presenta una definición alternativa de la distribución Dirichlet dada por Fang, Kotz and Ng [1], que permite derivar la función de densidad de IDH a partir de la Dirichlet.

**Definición 2.** Se dice que  $(X_1, \dots, X_n)$  tiene distribución Dirichlet (tipo I), denotada por  $(X_1, \dots, X_n) \sim D^I(\nu_1, \dots, \nu_{n-1}; \nu_n)$ , si su función de densidad está dada por

$$\frac{\Gamma(\nu)}{\prod_{i=1}^n \Gamma(\nu_i)} \prod_{i=1}^n x_i^{\nu_i-1}, \quad x_i > 0, \quad \sum_{i=1}^{n-1} x_i < 1, \quad x_n = 1 - \sum_{i=1}^{n-1} x_i.$$

donde  $\nu_i > 0$ ,  $i = 1, \dots, n$ ,  $\nu = \sum_{i=1}^n \nu_i$ .

**Teorema 1.** Sea  $(U_1, \dots, U_n) \sim D^I(\nu_1, \dots, \nu_{n-1}; \nu_n)$ , como en la Definición 2, y  $Z \sim IH(\nu, \alpha, \beta, \gamma)$ , independientes, donde  $\nu = \sum_{i=1}^n \nu_i$ . Entonces  $(Y_1, \dots, Y_n) \sim IDH(\nu_1, \dots, \nu_n; \alpha, \beta, \gamma)$ , donde  $Y_i = U_i Z$ , para  $i = 1, \dots, n$ .

*Demostración.* La prueba se obtiene considerando la transformación dada, cuyo jacobiano es  $J = (\sum_{i=1}^n y_i)^{-n+1}$ , y teniendo en cuenta que la función de densidad conjunta de  $Z$  y  $(U_1, \dots, U_n)$  está dada por

$$C(\nu_1, \dots, \nu_n; \alpha, \beta, \gamma) z^{\nu-1} (1+z)^{-\nu-\gamma} {}_2F_1\left(\alpha, \beta; \gamma; \frac{1}{1+z}\right) \prod_{i=1}^n u_i^{\nu_i-1}$$

donde  $z > 0$ ,  $u_i > 0$ ,  $i = 1, \dots, n$ ,  $\sum_{i=1}^{n-1} u_i < 1$ ,  $u_n = 1 - \sum_{i=1}^{n-1} u_i$  □

Una de las propiedades ideales para las distribuciones multivariadas es que las distribuciones marginales correspondan a la distribución del vector aleatorio. Es decir, si  $(U_1, \dots, U_n)$  tiene cierta distribución, entonces sería ideal que  $(U_1, \dots, U_s)$  también siguiera la misma distribución, para  $1 \leq s \leq n$ . Por ejemplo, la distribución Dirichlet tiene esta propiedad, pero la IDH no cuenta con ella. Lo cual se muestra en el siguiente teorema. En adelante,  $\nu = \sum_{i=1}^n \nu_i$ .

**Teorema 2.** Sea  $(Y_1, \dots, Y_n) \sim IDH(\nu_1, \dots, \nu_n; \alpha, \beta, \gamma)$ . Entonces,  $(Y_1, \dots, Y_s)$ , con  $s \leq n$ , tiene función de densidad dada por

$$\frac{\Gamma(\nu + \gamma - \alpha)\Gamma(\nu + \gamma - \beta)\Gamma(\nu^* + \gamma)}{\prod_{i=1}^s \Gamma(\nu_i)\Gamma(\gamma)\Gamma(\nu + \gamma - \alpha - \beta)\Gamma(\nu + \gamma)} \left( \prod_{i=1}^s y_i^{\nu_i-1} \right) \left( 1 + \sum_{i=1}^s y_i \right)^{-\nu^*-\gamma} \\ \times {}_3F_2\left(\alpha, \beta, \nu^* + \gamma; \gamma, \nu + \gamma; \frac{1}{1 + \sum_{i=1}^s y_i}\right)$$

donde  $\nu_i > 0$ ,  $i = 1, \dots, n$ ,  $\nu^* = \sum_{i=1}^s \nu_i$ ,  $\gamma > 0$ ,  $\alpha > 0$ ,  $\beta > 0$ ,  $\nu + \gamma > \alpha + \beta$  y  ${}_3F_2$  definida en (1).

*Demostración.* La densidad marginal de  $(Y_1, \dots, Y_s)$ ,  $s \leq n$ , se obtiene integrando la densidad conjunta de  $(Y_1, \dots, Y_n)$  respecto a  $y_{s+1}, \dots, y_n$ . Una herramienta útil para este cálculo es la integral de Liouville [4].  $\square$

El siguiente corolario muestra que la densidad de una componente  $Y_k$  no es una distribución conocida. Lo que se convierte en una desventaja de esta distribución, y se discutirá en la última sección. Seguido de este corolario, se muestra la distribución condicional de  $(Y_{s+1}, \dots, Y_n)|(Y_1, \dots, Y_s)$ .

**Corolario 1.** Si  $(Y_1, \dots, Y_n) \sim IDH(\nu_1, \dots, \nu_n; \alpha, \beta, \gamma)$ , entonces la función de densidad de  $Y_k$  está dada por

$$\frac{\Gamma(\nu + \gamma - \alpha)\Gamma(\nu + \gamma - \beta)\Gamma(\nu_k + \gamma)}{\Gamma(\nu_k)\Gamma(\gamma)\Gamma(\nu + \gamma - \alpha - \beta)\Gamma(\nu + \gamma)} y_k^{\nu_k - 1} (1 + y_k)^{-\nu_k - \gamma} \times {}_3F_2 \left( \alpha, \beta, \nu_k + \gamma; \gamma, \nu + \gamma; \frac{1}{1 + y_k} \right), \quad y_k > 0. \quad (2)$$

**Corolario 2.** Si  $(Y_1, \dots, Y_n) \sim IDH(\nu_1, \dots, \nu_n; \alpha, \beta, \gamma)$ , entonces la densidad condicional de  $(Y_{s+1}, \dots, Y_n)|(Y_1, \dots, Y_s)$ , para  $1 \leq s \leq n$ , es

$$\frac{\Gamma(\nu + \gamma) \prod_{i=s+1}^n y_i^{\nu_i - 1}}{\prod_{i=s+1}^n \Gamma(\nu_i)\Gamma(\nu^* + \gamma)} \cdot \frac{(1 + \sum_{i=1}^s y_i)^{\nu^* + \gamma}}{(1 + \sum_{i=1}^n y_i)^{\nu + \gamma}} \times \frac{{}_2F_1 \left( \alpha, \beta; \gamma; (1 + \sum_{i=1}^n y_i)^{-1} \right)}{{}_3F_2 \left( \alpha, \beta, \nu^* + \gamma; \gamma, \nu + \gamma; (1 + \sum_{i=1}^s y_i)^{-1} \right)},$$

donde  $y_i > 0$ ,  $\nu_i > 0$ ,  $\gamma > 0$  para  $i = 1, \dots, n$ ,  $\nu^* = \sum_{i=1}^s \nu_i$ .

Otro de los resultados importantes de este artículo es la distribución de las sumas parciales. Por ejemplo, se sabe que si  $(X_1, X_2) \sim D^{II}(\nu_1, \nu_2; \nu_3)$ , entonces  $X_1/X_2$  y  $X_1/(X_1 + X_2)$  son independientes de  $X_1 + X_2$ , donde  $D^{II}$  denota la distribución Dirichlet tipo II (o invertida). También,  $X_1/X_2 \sim B^{II}(\nu_1, \nu_2)$ ,  $X_1/(X_1 + X_2) \sim B^I(\nu_1, \nu_2)$ , y  $X_1 + X_2 \sim B^{II}(\nu_1 + \nu_2, \nu_3)$ , donde  $B^I$  y  $B^{II}$  denotan las distribuciones beta tipo I y beta tipo II (o invertida), respectivamente. En el siguiente teorema se derivan resultados similares para la distribución en cuestión.

**Teorema 3.** Sean  $(Y_1, \dots, Y_n) \sim IDH(\nu_1, \dots, \nu_n; \alpha, \beta, \gamma)$  y  $n_1, \dots, n_\ell$  enteros positivos tales que  $\sum_{i=1}^\ell n_i = n$ . Además, sean

$$\nu_{(i)} = \sum_{j=n_{i-1}^*+1}^{n_i^*} \nu_j, \quad n_0^* = 0, \quad n_i^* = \sum_{j=1}^i n_j, \quad i = 1, \dots, \ell.$$

Defina

$$Y_{(i)} = \sum_{j=n_{i-1}^*+1}^{n_i^*} Y_j, i = 1, \dots, \ell \quad y \quad Z_j = \frac{Y_j}{Y_{(i)}}, j = n_{i-1}^* + 1, \dots, n_i^* - 1.$$

Entonces,

- (i)  $(Y_{(1)}, \dots, Y_{(\ell)})$  y  $(Z_{n_{i-1}^*+1}, \dots, Z_{n_i^*-1}), i = 1, \dots, \ell$ , son independientes.
- (ii)  $(Z_{n_{i-1}^*+1}, \dots, Z_{n_i^*-1}) \sim D^I(\nu_{n_{i-1}^*+1}, \dots, \nu_{n_i^*-1}; \nu_{n_i^*}), i = 1, \dots, \ell$ .
- (iii)  $(Y_{(1)}, \dots, Y_{(\ell)}) \sim IDH(\nu_{(1)}, \dots, \nu_{(\ell)}; \alpha, \beta, \gamma)$ .

*Demostración.* El resultado se obtiene teniendo en cuenta la transformación y su jacobiano, que está dado por  $J = \prod_{i=1}^{\ell} y_{(i)}^{n_i^*-1}$ .  $\square$

Debido a la notación utilizada en el teorema anterior, es importante resaltar su utilidad, que consiste en mostrar la distribución, por ejemplo para  $n = 3$ , de expresiones de la forma  $\frac{Y_1 + Y_2}{Y_1 + Y_2 + Y_3}, \frac{Y_3}{Y_1 + Y_2}, \frac{Y_1 + Y_2}{Y_1}$ .

Los siguientes corolarios son consecuencia del teorema anterior.

**Corolario 3.** Sea  $(Y_1, \dots, Y_n) \sim IDH(\nu_1, \dots, \nu_n; \alpha, \beta, \gamma)$ . Defina  $Z_j = Y_j/Z$ , para  $j = 1, \dots, n-1$ , y  $Z = \sum_{i=1}^n Y_i$ . Entonces

$$Z \sim IH(\nu, \alpha, \beta, \gamma) \quad y \quad (Z_1, \dots, Z_{n-1}) \sim D^I(\nu_1, \dots, \nu_{n-1}; \nu_n)$$

son independientes.

**Corolario 4.** Si  $(Y_1, \dots, Y_n) \sim IDH(\nu_1, \dots, \nu_n; \alpha, \beta, \gamma)$  entonces

$$\frac{\sum_{i=1}^s Y_i}{\sum_{i=1}^n Y_i} \sim B^I(\nu^*, \nu - \nu^*), \quad s < n.$$

A continuación, se muestra cómo aplicando una transformación (muy conocida en estadística) se llega a la IDH.

Sea  $(Y_1, \dots, Y_n) \sim IDH(\nu_1, \dots, \nu_n; \alpha, \beta, \gamma)$

y defina  $X_i = \frac{Y_i}{1 + \sum_{j=1}^s Y_j}$ , para  $i = s+1, \dots, n$ . Entonces,  $(X_{s+1}, \dots, X_n) \sim$

$IDH(\nu_{s+1}, \dots, \nu_n; \alpha, \beta, \nu^* + \gamma)$ . Para llegar a esta conclusión, se escribe la función de densidad conjunta de  $Y_1, \dots, Y_s, X_{s+1}, \dots, X_n$  como

$$C(\nu_1, \dots, \nu_n; \alpha, \beta, \gamma) \left( \prod_{i=1}^s y_i^{\nu_i-1} \right) \left( 1 + \sum_{i=1}^s y_i \right)^{-\nu^*-\gamma} \left( \prod_{i=s+1}^n x_i^{\nu_i-1} \right) \\ \times \left( 1 + \sum_{i=s+1}^n x_i \right)^{-\nu-\gamma} {}_2F_1 \left( \alpha, \beta; \gamma; \frac{1}{(1 + \sum_{i=1}^s y_i)(1 + \sum_{i=s+1}^n x_i)} \right),$$

y se integra con respecto a  $Y_1, \dots, Y_s$  para obtener la función de densidad conjunta de  $X_{s+1}, \dots, X_n$ .

Como ya se vió con anterioridad, es posible derivar la IDH a partir de la distribución Dirichlet. Ahora, se verá como se puede representar mediante las distribuciones beta tipo I y beta tipo II. Sean  $U_1, \dots, U_{n-1}$  variables aleatorias independientes, con  $U_i \sim B^I(\sum_{j=1}^i \nu_j, \nu_{i+1})$  y  $Z \sim IH(\nu, \alpha, \beta, \gamma)$ . Entonces,

$$(Y_1, \dots, Y_n) = \left( \prod_{j=1}^{n-1} U_j, \dots, (1 - U_{i-1}) \prod_{j=i}^{n-1} U_j, \dots, 1 - U_{n-1} \right) Z \sim IDH(\nu_1, \dots, \nu_n; \alpha, \beta, \gamma).$$

Esto se muestra utilizando la transformación dada y su jacobiano

$$J = \prod_{i=1}^{n-1} \left( \sum_{j=1}^{i+1} y_j \right)^{-1}.$$

También se puede representar en términos de la distribución beta tipo II. Sean  $V_1, \dots, V_{n-1}$  variables aleatorias independientes, con  $V_i \sim B^{II}(\nu_{i+1}, \sum_{j=1}^i \nu_j)$  y  $Z \sim IH(\nu, \alpha, \beta, \gamma)$ . Entonces,

$$(Y_1, \dots, Y_n) = \left( \frac{1}{\prod_{j=1}^{n-1} (1 + V_i)}, \dots, \frac{V_{i-1}}{\prod_{j=i-1}^{n-1} (1 + V_j)}, \dots, \frac{V_{n-1}}{1 + V_{n-1}} \right) Z \sim IDH(\nu_1, \dots, \nu_n; \alpha, \beta, \gamma).$$

Otra representación es posible considerando  $W_1, \dots, W_{n-1}$  variables aleatorias independientes, con  $W_i \sim B^I(\nu_i, \sum_{j=i+1}^n \nu_j)$  y  $Z \sim IH(\nu, \alpha, \beta, \gamma)$ . Entonces se tiene que

$$(Y_1, \dots, Y_n) = \left( W_1, \dots, W_i \prod_{j=1}^{i-1} (1 - W_j), \dots, \prod_{j=1}^{n-1} (1 - W_j) \right) Z \sim IDH(\nu_1, \dots, \nu_n; \alpha, \beta, \gamma).$$

Análogo al anterior resultado, se utiliza la transformación y su jacobiano, dado por,  $J = \prod_{i=1}^{n-1} \left( \sum_{j=i}^n y_j \right)^{-1}$ . Ahora, si se consideran las variables aleatorias independientes,  $T_1, \dots, T_{n-1}$ , con  $T_i \sim B^{II}(\nu_i, \sum_{j=i+1}^n \nu_j)$  y  $Z \sim IH(\nu, \alpha, \beta, \gamma)$ , entonces

$$(Y_1, \dots, Y_n) = \left( \frac{T_1}{1 + T_1}, \dots, \frac{T_i}{\prod_{j=1}^i (1 + T_j)}, \dots, \frac{1}{\prod_{j=1}^{n-1} (1 + T_j)} \right) Z \sim IDH(\nu_1, \dots, \nu_n; \alpha, \beta, \gamma).$$

En general, se pueden obtener varias representaciones de la distribución Dirichlet-hipergeométrica invertida teniendo en cuenta lo siguiente:

1. Si  $X \sim B^I(a, b)$  entonces  $Y = \frac{X}{1-X} \sim B^{II}(a, b)$
2. Si  $X \sim B^{II}(a, b)$  entonces  $Y = \frac{1}{X} \sim B^{II}(b, a)$

### 3 Momentos

De la representación  $Y_i = U_i Z$ ,  $i = 1, \dots, n$ , dada en el Teorema 1, donde  $U_i$  es una componente Dirichlet y  $Z$  tiene distribución de la función hipergeométrica invertida, se obtienen los momentos conjuntos, que son los que proporcionan la media y la varianza de  $Y_i$ , así como la covarianza entre  $Y_i$  y  $Y_j$ . Además, recuerde que  $Z$  y  $(U_1, \dots, U_n)$  son independientes. A continuación se dan las expresiones para hallar los momentos de estas distribuciones:

$$E\left(\prod_{i=1}^n U_i^{r_i}\right) = \prod_{j=1}^n \frac{\nu_j(\nu_j + 1) \dots (\nu_j + r_j - 1)}{\nu(\nu + 1) \dots (\nu + r - 1)}, \quad (3)$$

donde  $r = \sum_{i=1}^n r_i$  y  $\nu = \sum_{i=1}^n \nu_i$ .

$$E(Z^r) = K \cdot \frac{\Gamma(\gamma - r)\Gamma(\nu + r)}{\Gamma(\nu + \gamma)} {}_3F_2(\gamma - r, \alpha, \beta; \gamma, \nu + \gamma; 1), \quad (4)$$

donde  $K = \frac{\Gamma(\gamma + \nu - \alpha)\Gamma(\gamma + \nu - \beta)}{\Gamma(\gamma)\Gamma(\nu)\Gamma(\gamma + \nu - \alpha - \beta)}$ .

Se sabe que  $(Y_1, \dots, Y_n) \sim IDH(\nu_1, \dots, \nu_n; \alpha, \beta, \gamma)$  entonces

$$E\left(\prod_{i=1}^n Y_i^{r_i}\right) = E\left(\prod_{i=1}^n U_i^{r_i}\right) E(Z^r)$$

Por lo tanto, si  $r_i = 1, 2$  y  $r_j = 0$  para  $j \neq i$ , se tiene que,

$$E(Y_i) = \frac{\nu_i \Gamma(\gamma + \nu - \alpha) \Gamma(\gamma + \nu - \beta)}{(\gamma - 1) \Gamma(\gamma + \nu) \Gamma(\gamma + \nu - \alpha - \beta)} {}_3F_2(\gamma - 1, \alpha, \beta; \gamma, \nu + \gamma; 1)$$

$$E(Y_i^2) = \frac{\nu_i(\nu_i + 1) \Gamma(\gamma + \nu - \alpha) \Gamma(\gamma + \nu - \beta)}{(\gamma - 1)(\gamma - 2) \Gamma(\gamma + \nu) \Gamma(\gamma + \nu - \alpha - \beta)} {}_3F_2(\gamma - 2, \alpha, \beta; \gamma, \nu + \gamma; 1)$$

Para no utilizar expresiones complicadas, se hace uso de la definición de varianza para escribir la varianza de  $Y_i$ ,

$$\text{Var}(Y_i) = E(Y_i^2) - [E(Y_i)]^2.$$

Tomando los coeficientes adecuados en (3) y (4), se puede encontrar una expresión para la covarianza entre las variables  $Y_i$  y  $Y_j$ ,

$$\begin{aligned} \text{Cov}(Y_i, Y_j) &= E(Y_i Y_j) - E(Y_i)E(Y_j) \\ &= E(Z^2)E(U_i U_j) - [E(Z)^2]E(U_i)E(U_j) \end{aligned}$$

Después de algunas manipulaciones algebraicas, se puede escribir que,

$$\text{Var}(Y_i) = \frac{\nu_i}{\nu^2(\nu+1)} [\nu_i\nu\text{Var}(Z) + \nu\text{E}(Z^2) - \nu_i(\text{E}(Z))^2]$$

$$\text{Cov}(Y_i, Y_j) = \frac{\nu_i\nu_j}{\nu^2(\nu+1)} [\nu\text{Var}(Z) - (\text{E}(Z))^2]$$

$\text{Cov}(Y_i, Y_j) > 0$  si  $\nu\text{Var}(Z) > (\text{E}(Z))^2$ . En caso contrario, esta covarianza es negativa.

#### 4 Estimación

Para estimar los parámetros de la distribución Dirichlet-hipergeométrica invertida se utilizó el método de máxima verosimilitud, el cual escoge como estimador de los parámetros aquel valor que hace máxima la probabilidad de que el modelo a estimar genere la muestra observada. En la práctica, el procedimiento es maximizar la función de verosimilitud,  $L(\theta)$ , o  $\ln L(\theta)$ , donde  $\theta$  es el vector de parámetros de la distribución. En primer lugar, considere  $n$  variables aleatorias  $Y_1, \dots, Y_n$  que tienen distribución conjunta Dirichlet-hipergeométrica invertida, y  $m$  observaciones de cada una de esas  $n$  variables. Sea  $Y_{ij}$  la  $i$ -ésima observación de la  $j$ -ésima variable y  $y_{ij}$  un valor fijo de  $Y_{ij}$ . Así, el logaritmo natural de la función de verosimilitud es

$$\ln L(\theta) = m \left( \ln C + \sum_{j=1}^n (\nu_j - 1) \ln G_j - (\nu + \gamma) \ln G \right) + H, \quad (5)$$

donde

$$\begin{aligned} \ln C &= \ln \Gamma(\nu + \gamma - \alpha) + \ln \Gamma(\nu + \gamma - \beta) \\ &\quad - \ln \Gamma(\nu + \gamma - \alpha - \beta) - \ln \Gamma(\gamma) - \sum_{j=1}^n \ln \Gamma(\nu_j) \end{aligned}$$

además

$$\begin{aligned} G_j &= \left( \prod_{i=1}^m y_{ij} \right)^{1/m} \\ G &= \left( \prod_{j=1}^n \left( 1 + \sum_{i=1}^m y_{ij} \right) \right)^{1/m} \\ H &= \sum_{i=1}^m \ln_2 F_1 \left( \alpha, \beta; \gamma; \frac{1}{1 + \sum_{j=1}^n y_{ij}} \right) \end{aligned}$$



Esta función de verosimilitud se maximizó numéricamente, utilizando el método de Fisher scoring [6]. En aras de buscar simplicidad y tratar de solucionar el problema, se dejan  $\alpha$ ,  $\beta$  y  $\gamma$  como valores fijos, para obtener las siguientes expresiones:

$$\begin{aligned}\frac{\partial \ln L(\theta)}{\partial \nu_j} &= m[\psi(\nu + \gamma - \alpha) + \psi(\nu + \gamma - \beta) - \psi(\nu + \gamma - \alpha - \beta) - \psi(\nu_j) + \\ &\quad \ln G_j - (\nu + \gamma) \ln G] \\ \frac{\partial^2 \ln L(\theta)}{\partial \nu_j^2} &= m[\psi'(\nu + \gamma - \alpha) + \psi'(\nu + \gamma - \beta) - \psi'(\nu + \gamma - \alpha - \beta) - \psi'(\nu_j)] \\ \frac{\partial^2 \ln L(\theta)}{\partial \nu_j \partial \nu_k} &= m[\psi'(\nu + \gamma - \alpha) + \psi'(\nu + \gamma - \beta) - \psi'(\nu + \gamma - \alpha - \beta)], \quad j \neq k\end{aligned}$$

donde  $\psi$  es la función digamma,  $\psi(x) = (\ln \Gamma(x))'$ , y  $\psi'$  es la función trigamma, que es la derivada de la función digamma.

El método Fisher scoring se usa por iteraciones, calculando el siguiente vector:

$$\begin{bmatrix} \hat{\nu}_1 \\ \vdots \\ \hat{\nu}_n \end{bmatrix}_i = \begin{bmatrix} \hat{\nu}_1 \\ \vdots \\ \hat{\nu}_n \end{bmatrix}_{i-1} + \begin{bmatrix} \text{Var}(\hat{\nu}_1) & \dots & \text{Cov}(\hat{\nu}_1, \hat{\nu}_n) \\ \vdots & \ddots & \vdots \\ \text{Cov}(\hat{\nu}_n, \hat{\nu}_1) & \dots & \text{Var}(\hat{\nu}_n) \end{bmatrix}_{i-1} \begin{bmatrix} g_1(\hat{\nu}_1) \\ \vdots \\ g_n(\hat{\nu}_n) \end{bmatrix}_{i-1}$$

donde  $\hat{\nu}'_{[0]} = [\hat{\nu}_{1(0)}, \dots, \hat{\nu}_{n(0)}]$  son los estimadores iniciales y  $g_j(\hat{\nu}_j) = \frac{\partial \ln L(\theta)}{\partial \nu_j}$ .

La matriz de varianzas y covarianzas se calcula como

$$\mathbf{V} = \mathbf{D}^{-1} + \delta \mathbf{a} \mathbf{a}', \quad \mathbf{D} = \text{diag}[m\psi'(\nu_1), \dots, m\psi'(\nu_n)],$$

donde

$$\mathbf{a}' = \left[ \frac{1}{\psi'(\nu_1)}, \dots, \frac{1}{\psi'(\nu_n)} \right], \quad \delta = \frac{-\lambda}{1 + \frac{\lambda}{m} \sum_{j=1}^n \frac{1}{\psi'(\nu_j)}}$$

con

$$\lambda = m[\psi'(\nu_1 + \nu_2 + \gamma - \alpha - \beta) - \psi'(\nu_1 + \nu_2 + \gamma - \alpha) - \psi'(\nu_1 + \nu_2 + \gamma - \beta)].$$

Para la prueba de convergencia se utiliza la forma cuadrática  $S = \mathbf{g}' \mathbf{V} \mathbf{g}$ , donde  $\mathbf{g}$  es el vector cuyas componentes son  $g_j(\hat{\nu}_j)$ . Kendall y Stuart [5] mostraron que  $S$  es una variable aleatoria con distribución chi-cuadrado con  $n$  grados de libertad. Por lo tanto, puede utilizarse como valor crítico en la prueba de convergencia. La iteración continua hasta que  $S$  sea menor que  $\chi_c^2(n)$ , para un valor  $\epsilon$  fijo en la cola inferior de la distribución chi-cuadrado.

Los resultados de aplicar este procedimiento en Mupad Pro 4.0, en el caso bivariado, para una muestra de tamaño  $m = 25$ , con  $\epsilon = 0,001$  y valores iniciales para los parámetros  $\nu_1 = 3$ ,  $\nu_2 = 2$ ,  $\alpha = 0,5$ ,  $\beta = 3$ ,  $\gamma = 3$ , se muestran a

continuación. La salida del algoritmo muestra el número de iteración,  $\nu_1$ ,  $\nu_2$  y el valor crítico  $S$  utilizado para la prueba de convergencia:

```

Iteración, v1, v2, S", 1, 3.57, 3.12, 17.87
Iteración, v1, v2, S", 2, 4.05, 3.95, 5.49
Iteración, v1, v2, S", 3, 4.38, 4.37, 1.27
Iteración, v1, v2, S", 4, 4.57, 4.57, 0.285
Iteración, v1, v2, S", 5, 4.66, 4.66, 0.0631
Iteración, v1, v2, S", 6, 4.71, 4.71, 0.0138
Iteración, v1, v2, S", 7, 4.73, 4.73, 0.003
Iteración, v1, v2, S", 8, 4.74, 4.74, 6.49e-4
"Termina el proceso"

```

El algoritmo se detiene cuando encuentra un valor para  $S$  menor que  $\chi_{0,001}^2(2)$ . Por tanto, para la muestra que se generó en este caso, el algoritmo converge en la iteración número ocho.

Ahora se describe brevemente, cómo se generó la muestra. En primer lugar, es importante recordar que si un vector  $(Y_1, \dots, Y_n)$  tiene distribución IDH, entonces la distribución marginal de  $Y_k$ , con  $k = 1, \dots, n$ , está dada por (2), la cual no es una distribución conocida. Por lo que no es posible encontrar un software que genere observaciones de una variable aleatoria con esa distribución. Para solucionar este problema, se hizo una aproximación, utilizando la distribución beta tipo I de dos parámetros.

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}, \quad 0 < x < 1. \quad (6)$$

Sea  $u_k = (1+y_k)^{-1}$ , si  $y_k > 0$  entonces  $0 < u_k < 1$ , para  $k = 1, \dots, n$ . Por esta razón, se elige la distribución beta para hacer la aproximación. Los parámetros se estimaron utilizando el método de los momentos, teniendo en cuenta que el primer y el segundo momento de la distribución beta son  $\frac{a}{a+b}$  y  $\frac{a(a+1)}{(a+b)(a+b+1)}$ , respectivamente. De este modo se obtiene el siguiente sistema de ecuaciones:

$$\begin{aligned} E(U_k) &= \frac{a}{a+b} \\ E(U_k^2) &= \frac{a(a+1)}{(a+b)(a+b+1)}, \end{aligned}$$

cuya solución está dada por

$$\begin{aligned} a &= E(U_k) \frac{E(U_k) - E(U_k^2)}{E(U_k^2) - [E(U_k)]^2} \\ b &= [1 - E(U_k)] \frac{E(U_k) - E(U_k^2)}{E(U_k^2) - [E(U_k)]^2} \end{aligned}$$

Ahora, se calcula  $E(U_k^t) = E[(1+Y_k)^{-t}]$ , para cualquier real  $t$  como,

$$\frac{(\gamma)_t \Gamma(\nu + \gamma - \alpha) \Gamma(\nu + \gamma - \beta)}{(\gamma + \nu_k)_t \Gamma(\nu + \gamma - \alpha - \beta) \Gamma(\nu + \gamma)} {}_4F_3(\alpha, \beta, \nu_k + \gamma, \gamma + t; \gamma, \nu + \gamma, \gamma + t + \nu_k; 1),$$

donde  $(\gamma)_t = \gamma(\gamma + 1)\dots(\gamma + t - 1)$ . Para  $t = 1, 2$  se obtienen expresiones para  $a$  y  $b$ , que son los parámetros de la distribución beta que se utilizará como aproximación.

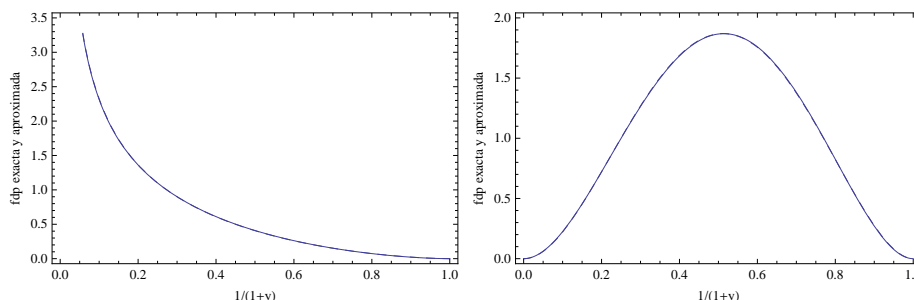
La calidad de la aproximación se verificó gráficamente, es decir, se compararon ambas densidades: la exacta y la aproximada de la variable  $U_k = (1 + Y_k)^{-1}$ , asignando valores a los parámetros  $(\nu_1, \nu_2, \alpha, \beta, \gamma)$ . La distribución exacta se encuentra aplicando la técnica de la transformación a la densidad dada en 2 como

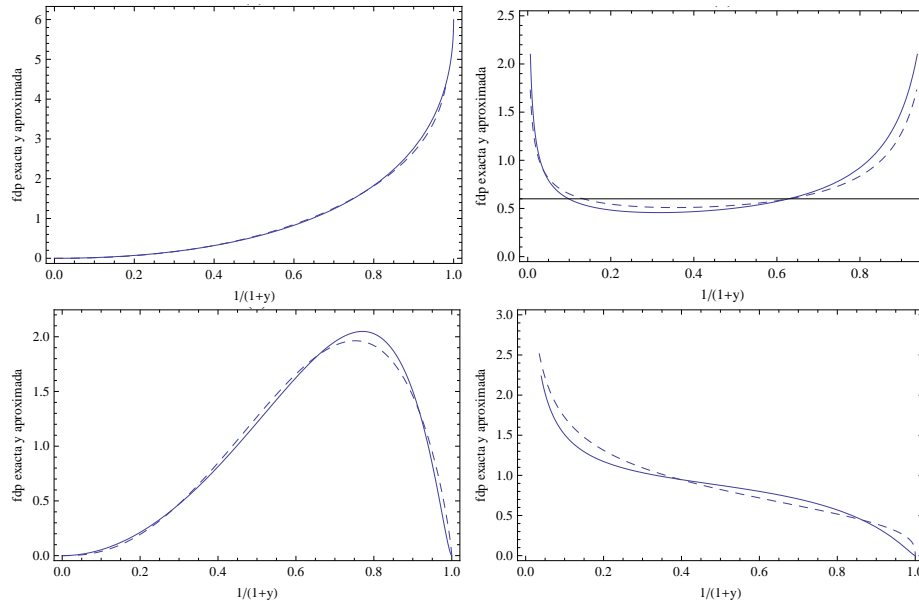
$$\frac{\Gamma(\nu + \gamma - \alpha)\Gamma(\nu + \gamma - \beta)\Gamma(\nu_k + \gamma)}{\Gamma(\nu_k)\Gamma(\gamma)\Gamma(\nu + \gamma - \alpha - \beta)\Gamma(\nu + \gamma)} \times u_k^{\gamma-1}(1 - u_k)^{\nu_k-1} {}_3F_2(\alpha, \beta, \nu_k + \gamma; \gamma, \nu + \gamma; u_k),$$

con  $0 < u_k < 1$ , y la aproximada está dada por (6). Se asignaron valores a los parámetros menores que 1, iguales a 1 y mayores que 1, y se consideraron todas las posibles combinaciones, teniendo en cuenta que  $\nu_1 + \nu_2 + \gamma < \alpha + \beta$ . Unos pocos resultados de este procedimiento se muestran en la siguiente tabla, la cual contiene los estimadores de  $a$  y  $b$ , dependiendo de los valores de los parámetros. Luego se muestran los gráficos que ilustran ambas distribuciones, exacta y aproximada; la curva sólida denota la función de densidad exacta y la curva punteada denota la aproximada.

Observando los gráficos, se puede concluir que la aproximación realizada se ajusta a la distribución exacta, para los diferentes valores dados de los parámetros  $\nu_1, \nu_2, \alpha, \beta$  y  $\gamma$ . Note que la comparación de ambas distribuciones, exacta y aproximada, se realizó gráficamente.

Gráfico	$\nu_1$	$\nu_2$	$\alpha$	$\beta$	$\gamma$	$a$	$b$
1	0.5	3	0.5	0.5	0.5	0.5033	2.6028
2	0.5	3	0.5	1	3	3.0306	2.9342
3	1	1	0.5	3	3	3.2098	0.8200
4	1	1	1	1	0.5	0.6261	0.3170
5	3	3	3	3	3	3.4712	1.8154
6	3	3	3	1	0.5	0.6629	1.3338





### Agradecimientos

La investigadora agradece al Departamento de Matemáticas de la Facultad de Ciencias Naturales y Exactas de la Universidad del Valle, por la financiación a este proyecto registrado con el código CI 7797. También, al profesor titular Daya K. Nagar del Departamento de Matemáticas de la Facultad de Ciencias Exactas y Naturales de la Universidad de Antioquia, por sugerir el tema.

### Referencias

- [1] K. T. Fang, S. Kotz and K. W. Ng, *Symmetric Multivariate and Related Distributions*, Chapman and Hall, London, 1989.
- [2] A. K. Gupta and D. K. Nagar, *Matrix Variate Distributions*, Chapman & Hall/CRC, Boca Raton, Fl, 2000.
- [3] A. K. Gupta and D. Song, Generalized Liouville distribution, *Computters & Mathematics with Applications*, **32**(1996), no. 2, 103–109.
- [4] Rameshwar D. Gupta and Donald St. P. Richards, The History of the Dirichlet and Liouville distributions, *International Statistical Review*, **69** (2001), no. 3, 433-446.
- [5] M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics, Vol. II*, Charles Griffin and Co., London, 1963.
- [6] S. Kotz, N. Balakrishnan and N. L. Johnson, *Continuous Multivariate Distributions, Vol. 1*, Second edition, Wiley-Interscience, New York, 2000.

- [7] Y. L. Luke, *The Special Functions and Their Approximations, Vol. 1*, Academic Press, New York, 1969.
- [8] Albert W. Marshall and Ingram Olkin, *Inequalities: theory of majorization and its applications. Mathematics in Science and Engineering, 143*, Academic Press, Inc. [Harcourt Brace Jovanovich Publishers], New York, 1979.
- [9] Daya K. Nagar, Paula A. Bran-Cardona and Arjun K. Gupta, Multivariate generalization of the hypergeometric function type I distribution. *Acta Applicandae Mathematicae*, **105**(2009), no. 1, 111–122.
- [10] D. Song and A. K. Gupta, Properties of generalized Liouville distributions, *Random Operators and Stochastic Equations*, **5**(1997), no. 4, 337–348.
- [11] George G. Tiao and Irwin Cuttman, The inverted Dirichlet distribution with applications, *American Statistical Association Journal*, **60**(1965), 793–805.

*Dirección del autor*

Paula A. Bran Cardona — Departamento de Matemáticas, Universidad del Valle, Cali-Colombia

e-mail: paula.bran@correounivalle.edu.co