Hacettepe Journal of Mathematics and Statistics
Volume 44 (4) (2015), 933 – 948

Multivariate generalization of the Gauss hypergeometric distribution

Daya K. Nagar* , Danilo Bedoya-Valencia† and Saralees Nadarajah‡

Abstract

The Gauss hypergeometric distribution with the density proportional to $x^{\alpha-1}(1-x)^{\beta-1}(1+\xi x)^{-\gamma}$, $0 < x < 1$ arises in connection with the prior distribution of the parameter ρ $(0 < \rho < 1)$ representing traffic intensity in a $M/M/1$ queue system. In this article, we define and study a multivariate generalization of this distribution and derive some of its properties like marginal densities, joint moments, and factorizations. A data application is given.

2000 AMS Classification: Primary 33E99; Secondary 62H99.

Keywords: Beta function; Dirichlet function; Gamma function; Gauss hypergeometric function; Liouville integral; Multivariate; Transformation.

Received 15/11/2013 : Accepted 01/09/2014 Doi : 10.15672/HJMS.2014277478

1. Introduction

A random variable X is said to have a Gauss hypergeometric distribution with parameters $\alpha > 0$, $\beta > 0$, $-\infty < \gamma < \infty$ and $\xi > -1$, denoted by $X \sim \mathrm{GH}(\alpha, \beta, \gamma, \xi)$, if its probability density function (p.d.f.) is given by

$$(1.1) \qquad f_{\mathrm{GH}}(x; \alpha, \beta, \gamma, \xi) = C(\alpha, \beta, \gamma, \xi)\, \frac{x^{\alpha-1}(1-x)^{\beta-1}}{(1+\xi x)^{\gamma}}, \quad 0 < x < 1,$$

where the normalizing constant $C(\alpha, \beta, \gamma, \xi)$ is given by

$$(1.2) \qquad C(\alpha, \beta, \gamma, \xi) = [B(\alpha, \beta)\ {}_2F_1(\gamma, \alpha; \alpha + \beta; -\xi)]^{-1},$$

*Instituto de Matemáticas, Universidad de Antioquia, Calle 67, No. 53-108, Medellín, Colombia.
Email: dayaknagar@yahoo.com
†Instituto de Matemáticas, Universidad de Antioquia, Calle 67, No. 53-108, Medellín, Colombia.
Email:danilo.bv@gmail.com
‡School of Mathematics, University of Manchester, Manchester, M13 9PL, UK.
Email:Saralees.Nadarajah@manchester.ac.uk

with $B(\alpha,\beta)$ being the beta function is defined by

$$B(\alpha,\beta)=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)},$$

and ${}_2F_1$ is the Gauss hypergeometric function (Luke [13]). Note that the Gauss hypergeometric function ${}_2F_1$ in (1.2) can be expanded in series form if $-1<\xi<1$. If $\xi>1$, then the function can be suitably transformed such that the absolute value of its argument is less than one, see (2.5).

The above distribution was suggested by Armero and Bayarri [1] in connection with the prior distribution of the parameter ρ, $0<\rho<1$ representing the traffic intensity in a $M/M/1$ queueing system. A brief introduction of this distribution is given in the encyclopedic work of Johnson, Kotz and Balakrishnan [10, p. 253]. In the context of Bayesian analysis of unreported Poisson count data, while deriving the marginal posterior distribution of the reporting probability p, Fader and Hardie [5] have shown that $q=1-p$ has a Gauss hypergeometric distribution. The Gauss hypergeometric distribution has also been used by Dauxois [4] to introduce conjugate priors in the Bayesian inference for linear growth birth and death processes. Sarabia and Castillo [21] have pointed out that this distribution is conjugate prior for the binomial distribution.

When either γ or ξ equals to zero, the Gauss hypergeometric p.d.f. reduces to a beta type 1 p.d.f. given by (Johnson, Kotz and Balakrishnan [10]),

$$f_{\mathrm{B1}}(x;\alpha,\beta)=\frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)},\quad 0<x<1. \tag{1.3}$$

Further, for $\gamma=\alpha+\beta$ and $\xi=1$ the Gauss hypergeometric distribution simplifies to a beta type 3 distribution given by the p.d.f. (Cardeño, Nagar and Sánchez [3], Sánchez and Nagar [20]),

$$f_{\mathrm{B3}}(x;\alpha,\beta)=\frac{2^{\alpha}x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)(1+x)^{\alpha+\beta}},\quad 0<x<1. \tag{1.4}$$

The matrix variate generalizations of beta type 1 and beta type 3 distributions have been defined and studied extensively. For example, see Gupta and Nagar [6, 7]. For $\gamma=\alpha+\beta$ and $\xi=-(1-\lambda)$ the GH distribution slides to a three parameter generalized beta type 1 distribution (Libby and Novic [12], Pham-Gia and Duong [19], Nadarajah [15], Nagar and Rada-Mora [17]) defined by the p.d.f.

$$f_{\mathrm{GB1}}(x;\alpha,\beta;\lambda)=\frac{\lambda^{\alpha}x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)[1-(1-\lambda)x]^{\alpha+\beta}},\quad 0<x<1, \tag{1.5}$$

where $\alpha>0$ and $\beta>0$.

In this article, we propose a multivariate generalization of the Gauss hypergeometric distribution which is a new members of the Liouville family of distributions. We define the multivariate generalization of (1.1) and study some of its properties such as marginal p.d.f.s, joint moments, variance and covariances. We also derive the distribution of partial sums of random variables jointly distributed as multivariate Gauss hypergeometric and several results on factorizations in terms of known distributions. Finally, a data application of the multivariate Gauss hypergeometric p.d.f. is illustrated.

Multivariate Liouville family of distributions was proposed by Marshall and Olkin [14]. Sivazlian [22] introduced Liouville distributions as generalizations of gamma and Dirichlet distributions. The Dirichlet and Liouville distributions arise in a variety of context including Bayesian analysis, modeling of multivariate data, order statistics, limit laws, multivariate analysis, reliability theory and stochastic processes. These distributions have been widely used in geology, biology, chemistry, forensic science, and statistical genetics. A comprehensive account of some applications and other aspects of these distributions can be found in Gupta and Song [9], Gupta and Richards [8], Marshall and Olkin [14], Nagar, Bran-Cardona and Gupta [16], Nagar and Sepúlveda-Murillo [18], and Song and Gupta [23].

2. Preliminaries

In this section we give definitions and results that will be used in subsequent sections. Throughout this work we will use the Pochhammer symbol $(a)_n$ defined by $(a)_n = a(a+1)\cdots(a+n-1) = (a)_{n-1}(a+n-1)$ for $n = 1, 2, \ldots,$ and $(a)_0 = 1$.

The generalized hypergeometric function of scalar argument is defined by

$$ {}_pF_q\left(a_1, \ldots, a_p; b_1, \ldots, b_q; x\right) = \sum_{k=0}^{\infty} \frac{(a_1)_k \cdots (a_p)_k}{(b_1)_k \cdots (b_q)_k} \frac{x^k}{k!}, \tag{2.1} $$

where a_i, $i = 1, \ldots, p$; b_j, $j = 1, \ldots, q$ are complex numbers with suitable restrictions and x is a complex variable.

Conditions for the convergence of the series in (2.1) are available in the literature, see Luke [13]. From (2.1) it is easy to see that

$$ {}_2F_1(a, b; c; x) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} \frac{x^k}{k!}, \quad |x| < 1. \tag{2.2} $$

The integral representation of the Gauss hypergeometric function is

$$ {}_2F_1(a, b; c; x) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(c-a)} \int_0^1 t^{a-1}(1-t)^{c-a-1}(1-xt)^{-b}\, \mathrm{d}t, \tag{2.3} $$

where $\mathrm{Re}(c) > \mathrm{Re}(a) > 0$ and $|\arg(1-x)| < \pi$. Note that, the series expansion for ${}_2F_1$ given in (2.2) can be obtained by expanding $(1-xt)^{-b}$, $|xt| < 1$, in (2.3) and integrating t. Substituting $x = 1$ in (2.3) and integrating, we obtain

$$ {}_2F_1(a, b; c; 1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}, \quad \mathrm{Re}(c-a-b) > 0, \tag{2.4} $$

$c \neq 0, -1, -2, \ldots$. The Gauss' hypergeometric function ${}_2F_1$ satisfies the following relations

$$ \begin{aligned} {}_2F_1(a, b; c; x) &= (1-x)^{-b}\, {}_2F_1(c-a, b; c; -x(1-x)^{-1}) \\ &= (1-x)^{c-a-b}\, {}_2F_1(c-a, c-b; c; x). \end{aligned} \tag{2.5} $$

Let f be a continuous function and $\mathrm{Re}(\alpha_i) > 0$, $i = 1, \ldots, n$, the integral

$$ D_n(\alpha_1, \ldots, \alpha_n; f) = \underset{\substack{x_1>0,\ldots,x_n>0 \\ \sum_{i=1}^n x_i<1}}{\int \cdots \int} \prod_{i=1}^n x_i^{\alpha_i - 1} f\left(\sum_{i=1}^n x_i\right) \prod_{i=1}^n \mathrm{d}x_i, $$

is known as the Liouville-Dirichlet integral. Making the substitution $y_i = x_i/x$, $i = 1, \ldots, n-1$ and $x = \sum_{i=1}^n x_i$ with the Jacobian $J(x_1, \ldots, x_n \to y_1, \ldots, y_{n-1}, x) = x^{n-1}$ and integrating, we obtain

$$ D_n(\alpha_1, \ldots, \alpha_n; f) = \frac{\prod_{i=1}^n \Gamma(\alpha_i)}{\Gamma\left(\sum_{i=1}^n \alpha_i\right)} \int_0^1 x^{\sum_{i=1}^n \alpha_i - 1} f(x)\, \mathrm{d}x. \tag{2.6} $$

In particular, for $f(x) = (1-x)^{\beta-1}/(1+\xi x)^{\gamma}$, we obtain

$$ D_n\left(\alpha_1, \ldots, \alpha_n; f\right) = \frac{\prod_{i=1}^n \Gamma(\alpha_i)\Gamma(\beta)}{\Gamma\left(\sum_{i=1}^n \alpha_i + \beta\right)}\, {}_2F_1\left(\gamma, \sum_{i=1}^n \alpha_i; \sum_{i=1}^n \alpha_i + \beta; -\xi\right). \tag{2.7} $$

3. The Density Function

We propose a multivariate generalization of the Gauss hypergeometric distribution as follows.

3.1. Definition. The random variables $X_1, \dots, X_n$ are said to have a multivariate Gauss hypergeometric distribution with parameters $\alpha_i > 0$, $i = 1, \dots, n$, $\beta > 0$, $-\infty < \gamma < \infty$ and $\xi > -1$, denoted as $(X_1, \dots, X_n) \sim \mathrm{GH}(\alpha_1, \dots, \alpha_n, \beta, \gamma, \xi)$, if their joint p.d.f. is

$$C(\alpha_1, \dots, \alpha_n, \beta, \gamma, \xi)\ \frac{\prod_{i=1}^n x_i^{\alpha_i - 1} \left(1 - \sum_{i=1}^n x_i\right)^{\beta - 1}}{\left(1 + \xi \sum_{i=1}^n x_i\right)^{\gamma}}, \tag{3.1}$$

$$x_i > 0,\ i = 1, \dots, n,\ \sum_{i=1}^n x_i < 1,$$

where $C(\alpha_1, \dots, \alpha_n, \beta, \gamma, \xi)$ is the normalizing constant.

Since, the integration of the p.d.f. (3.1) over its support set is one, we have

$$C(\alpha_1, \dots, \alpha_n, \beta, \gamma, \xi) \underset{\substack{x_1 > 0, \dots, x_n > 0 \\ \sum_{i=1}^n x_i < 1}}{\int \cdots \int} \frac{\prod_{i=1}^n x_i^{\alpha_i - 1} \left(1 - \sum_{i=1}^n x_i\right)^{\beta - 1}}{\left(1 + \xi \sum_{i=1}^n x_i\right)^{\gamma}}\, \mathrm{d}x_1 \cdots \mathrm{d}x_n = 1,$$

and, using (2.7), we obtain the expression for $C(\alpha_1, \dots, \alpha_n, \beta, \gamma, \xi)$ as

$$[C(\alpha_1, \dots, \alpha_n, \beta, \gamma, \xi)]^{-1} = \frac{\prod_{i=1}^n \Gamma(\alpha_i) \Gamma(\beta)}{\Gamma\left(\sum_{i=1}^n \alpha_i + \beta\right)}\ {}_2F_1\left(\sum_{i=1}^n \alpha_i, \gamma; \sum_{i=1}^n \alpha_i + \beta; -\xi\right). \tag{3.2}$$

For specific values of the parameters, we obtain several known multivariate distributions. For $\xi = 0$ or $\gamma = 0$, the p.d.f. (3.1), takes the form of a Dirichlet type 1 p.d.f., $(X_1, \dots, X_n) \sim \mathrm{D1}(\alpha_1, \dots, \alpha_n; \beta)$, given by

$$\frac{\Gamma\left(\sum_{i=1}^n \alpha_i + \beta\right)}{\prod_{i=1}^n \Gamma(\alpha_i) \Gamma(\beta)} \prod_{i=1}^n x_i^{\alpha_i - 1} \left(1 - \sum_{i=1}^n x_i\right)^{\beta - 1}, \quad x_i > 0,\ i = 1, \dots, n,\ \sum_{i=1}^n x_i < 1.$$

For $\gamma = \sum_{i=1}^n \alpha_i + \beta$ and $\xi = 1$, the p.d.f. (3.1) reduces to a Dirichlet type 3 p.d.f., $(X_1, \dots, X_n) \sim \mathrm{D3}(\alpha_1, \dots, \alpha_n; \beta)$, stated as

$$\frac{2^{\sum_{i=1}^n \alpha_i} \Gamma\left(\sum_{i=1}^n \alpha_i + \beta\right)}{\prod_{i=1}^n \Gamma(\alpha_i) \Gamma(\beta)} \frac{\prod_{i=1}^n x_i^{\alpha_i - 1} \left(1 - \sum_{i=1}^n x_i\right)^{\beta - 1}}{\left(1 + \sum_{i=1}^n x_i\right)^{\sum_{i=1}^n \alpha_i + \beta}},$$

$$x_i > 0,\ i = 1, \dots, n,\ \sum_{i=1}^n x_i < 1.$$

The Dirichlet type 1 and Dirichlet type 3 distributions have been studied extensively in the literature. For example, see, Kotz, Balakrishnana and Johnson [11] and Cardeño, Nagar and Sánchez [3].

In Bayesian probability theory, if the posterior distribution belongs to the same family as the prior distribution, then the prior and posterior are called conjugate distributions, and the prior is called a conjugate prior. In case of multinomial distribution, the usual conjugate prior is the Dirichlet distribution. In the present case, if

$$p(s_1, \dots, s_n, f | x_1, \dots, x_n) = \binom{s_1 + \cdots + s_n + f}{s_1, \dots, s_n, f} x_1^{s_1} \cdots x_n^{s_n} (1 - x_1 - \cdots - x_n)^f$$

and

$$p(x_1, \dots, x_n) = C(\alpha_1, \dots, \alpha_n, \beta, \gamma, \xi) \frac{x_1^{\alpha_1 - 1} \cdots x_n^{\alpha_n - 1} (1 - x_1 - \cdots - x_n)^{\beta - 1}}{[1 + \xi(x_1 + \cdots + x_n)]^{\gamma}},$$

where $x_1 > 0, \dots, x_n > 0$, and $x_1 + \cdots + x_n < 1$, then

$$p(x_1, \dots, x_n | s_1, \dots, s_n, f) = C(\alpha_1 + s_1, \dots, \alpha_n + s_n, \beta + f, \gamma, \xi) \times \frac{x_1^{\alpha_1+s_1-1} \cdots x_n^{\alpha_n+s_n-1}(1 - x_1 - \cdots - x_n)^{\beta+f-1}}{[1 + \xi(x_1 + \cdots + x_n)]^\gamma}.$$

Thus, the multivariate family of distributions considered in this article is conjugate prior for the multinomial distribution.

Figure 1 gives some graphs of the p.d.f. define by (3.1) for different values of the parameters. A wide range of shapes arise out of the multivariate Gauss hypergeometric p.d.f.

In the next theorem, by applying a linear transformation to the multivariate Gauss hypergeometric variables, we define a generalization of the Dirichlet type 2 distribution.

3.2. Theorem. *Let $(X_1, \dots, X_n) \sim \mathrm{GH}(\alpha_1, \dots, \alpha_n, \beta, \gamma, \xi)$. Define $Y_1, \dots, Y_n$ as $Y_i = X_i/(1 - \sum_{i=1}^n X_i)$, $i = 1, \dots, n$. Then, the p.d.f. of $(Y_1, \dots, Y_n)$ is*

$$C(\alpha_1, \dots, \alpha_n, \beta, \gamma, \xi) \frac{\prod_{i=1}^n y_i^{\alpha_i-1} \left(1 + \sum_{i=1}^n y_i\right)^{\gamma - \sum_{i=1}^n \alpha_i - \beta}}{\left[1 + (1 + \xi) \sum_{i=1}^n y_i\right]^\gamma},$$

where $y_i > 0$, $i = 1, \dots, n$ and $C(\alpha_1, \dots, \alpha_n, \beta, \gamma, \xi)$ is the normalizing constant given in (3.2).

Proof. Transforming $X_i = Y_i/(1+\sum_{i=1}^n Y_i)$ with the Jacobian $J(x_1, \dots, x_n \to y_1, \dots, y_n) = (1 + \sum_{i=1}^n y_i)^{-(n+1)}$ in the p.d.f. (3.1), we obtain the desired result. □

Note that, if $\xi = 0$ or $\gamma = 0$, then the p.d.f. given in the above theorem slides to a Dirichlet type 2 p.d.f. given by

$$\frac{\Gamma\left(\sum_{i=1}^n \alpha_i + \beta\right)}{\prod_{i=1}^n \Gamma(\alpha_i)\Gamma(\beta)} \frac{\prod_{i=1}^n y_i^{\alpha_i-1}}{\left(1 + \sum_{i=1}^n y_i\right)^{\sum_{i=1}^n \alpha_i + \beta}}, \quad y_i > 0, \quad i = 1, \dots, n,$$

and in this case we write $(Y_1, \dots, Y_n) \sim \mathrm{D2}(\alpha_1, \dots, \alpha_n; \beta)$.

4. Marginal Distribution

It is well known that if $(X_1, \dots, X_n) \sim \mathrm{D1}(\alpha_1, \dots, \alpha_n; \beta)$, then for $1 \le s \le n$, $(X_1, \dots, X_s) \sim \mathrm{D1}(\alpha_1, \dots, \alpha_s; \beta + \sum_{i=s+1}^n \alpha_i)$. In this section, we derive similar result for multivariate generalization of the Gauss hypergeometric distribution defined by the p.d.f. (3.1).

4.1. Theorem. *Let $(X_1, \dots, X_n) \sim \mathrm{GH}(\alpha_1, \dots, \alpha_n, \beta, \gamma, \xi)$. Then, for $1 \le s \le n$, the joint p.d.f. of $X_1, \dots, X_s$ is*

$$K_1(\alpha_1, \dots, \alpha_n, \beta, \gamma, \xi) \frac{\prod_{i=1}^s x_i^{\alpha_i-1} \left(1 - \sum_{i=1}^s x_i\right)^{\beta + \sum_{i=s+1}^n \alpha_i - 1}}{\left(1 + \xi \sum_{i=1}^s x_i\right)^\gamma} \times {}_2F_1\left(\sum_{i=s+1}^n \alpha_i, \gamma; \sum_{i=s+1}^n \alpha_i + \beta; -\frac{\xi\left(1 - \sum_{i=1}^s x_i\right)}{1 + \xi \sum_{i=1}^s x_i}\right), \tag{4.1}$$

for $x_1 > 0, \dots, x_s > 0$ and $\sum_{i=1}^s x_i < 1$, where

$$[K_1(\alpha_1, \dots, \alpha_n, \beta, \gamma, \xi)]^{-1} = \frac{\prod_{i=1}^s \Gamma(\alpha_i)\Gamma\left(\sum_{i=s+1}^n \alpha_i + \beta\right)}{\Gamma\left(\sum_{i=1}^n \alpha_i + \beta\right)} {}_2F_1\left(\sum_{i=1}^n \alpha_i, \gamma; \sum_{i=1}^n \alpha_i + \beta; -\xi\right).$$

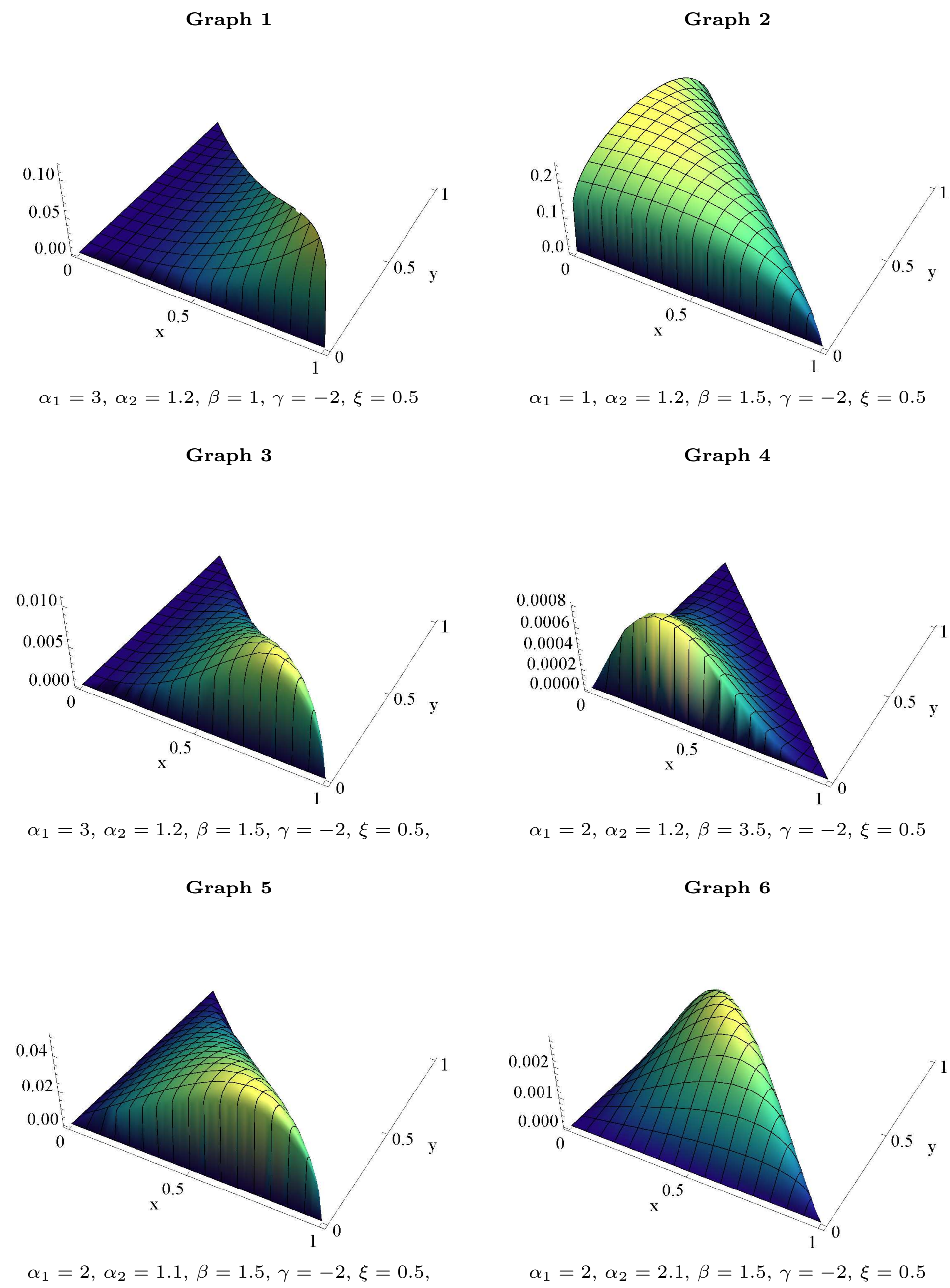

Figure 1. p.d.f. of the multivariate Gauss hypergeometric distribution.

Proof. To calculate the marginal p.d.f. of $X_1, \ldots, X_s$, we integrate (3.1) with respect to $x_{s+1}, \ldots, x_n$, to obtain

$$C(\alpha_1, \ldots, \alpha_n, \beta, \gamma, \xi) \prod_{i=1}^{s} x_i^{\alpha_i - 1}$$
$$\times \underset{\substack{x_{s+1}>0,\ldots,x_n>0 \\ \sum_{i=s+1}^n x_i < 1-\sum_{i=1}^s x_i}}{\int \cdots \int} \frac{\prod_{i=s+1}^n x_i^{\alpha_i-1}(1-\sum_{i=1}^s x_i - \sum_{i=s+1}^n x_i)^{\beta-1}}{(1+\xi\sum_{i=1}^s x_i + \xi \sum_{i=s+1}^n x_i)^{\gamma}} \mathrm{d}x_{s+1}\cdots \mathrm{d}x_n,$$

where $0 < x_i$, $i = 1, \ldots, s$. Now, substituting $z_j = x_j/(1-\sum_{i=1}^s x_i)$ for $j = s+1, \ldots, n$ with the Jacobian $J(x_{s+1}, \ldots, x_n \to z_{s+1}, \ldots, z_n) = \left(1-\sum_{i=1}^s x_i\right)^{n-s}$, the marginal p.d.f. of $X_1, \ldots, X_s$ is obtained as

$$C(\alpha_1, \ldots, \alpha_n, \beta, \gamma, \xi) \frac{\prod_{i=1}^s x_i^{\alpha_i-1}\left(1-\sum_{i=1}^s x_i\right)^{\beta+\sum_{i=s+1}^n \alpha_i - 1}}{\left(1+\xi\sum_{i=1}^s x_i\right)^{\gamma}}$$
$$\times \underset{\substack{z_{s+1}>0,\ldots,z_n>0 \\ \sum_{i=s+1}^n z_i<1}}{\int \cdots \int} \frac{\prod_{i=s+1}^n z_i^{\alpha_i-1}(1-\sum_{i=s+1}^n z_i)^{\beta-1}}{\left[1+\xi(1-\sum_{i=1}^s x_i)\sum_{i=s+1}^n z_i/(1+\xi\sum_{i=1}^s x_i)\right]^{\gamma}} \mathrm{d}z_{s+1}\cdots\mathrm{d}z_n. \tag{4.2}$$

Further, using the Liouville-Dirichlet integral (2.7), we can evaluate the above integral as

$$\frac{\prod_{i=s+1}^n \Gamma(\alpha_i)}{\Gamma\left(\sum_{i=s+1}^n \alpha_i\right)} \int_0^1 \frac{z^{\sum_{i=s+1}^n \alpha_i - 1}(1-z)^{\beta-1}}{\left[1+\xi\left(1-\sum_{i=1}^s x_i\right)z/(1+\xi\sum_{i=1}^s x_i)\right]^{\gamma}}\, \mathrm{d}z$$
$$= \frac{\prod_{i=s+1}^n \Gamma(\alpha_i)\Gamma(\beta)}{\Gamma\left(\sum_{i=s+1}^n \alpha_i + \beta\right)}\, {}_2F_1\left(\sum_{i=s+1}^n \alpha_i, \gamma; \sum_{i=s+1}^n \alpha_i + \beta; -\frac{\xi\left(1-\sum_{i=1}^s x_i\right)}{1+\xi\sum_{i=1}^s x_i}\right).$$

Finally, substituting this last expression, as well as the value of $C(\alpha_1, \ldots, \alpha_n, \beta, \gamma, \xi)$, in (4.2) and simplifying, we obtain the desired result. □

Note that the marginal p.d.f. of $X_1, \ldots, X_s$, obtained in the previous theorem, differs from the multivariate Gauss hypergeometric p.d.f. by a factor that involves the ${}_2F_1$ function.

4.2. Corollary. *Let* $(X_1, \ldots, X_n) \sim \mathrm{GH}(\alpha_1, \ldots, \alpha_n, \beta, \gamma, \xi)$*. Then, for* k=1, …, n*, the p.d.f. of* X_k *is*

$$K_2(\alpha_1, \ldots, \alpha_n, \beta, \gamma, \xi)\ \frac{x_k^{\alpha_k-1}(1-x_k)^{\beta+\sum_{i(\neq k)=1}^n \alpha_i - 1}}{(1+\xi x_k)^{\gamma}}$$
$$\times\, {}_2F_1\left(\sum_{i(\neq k)=1}^n \alpha_i, \gamma; \sum_{i(\neq k)=1}^n \alpha_i + \beta; -\frac{\xi(1-x_k)}{1+\xi x_k}\right),$$

for $0 < x_k < 1$*, where*

$$[K_2(\alpha_1, \ldots, \alpha_n, \beta, \gamma, \xi)]^{-1}$$
$$= \frac{\Gamma(\alpha_k)\Gamma(\sum_{i(\neq k)=1}^n \alpha_i + \beta)}{\Gamma\left(\sum_{i=1}^n \alpha_i + \beta\right)}\, {}_2F_1\left(\sum_{i=1}^n \alpha_i, \gamma; \sum_{i=1}^n \alpha_i + \beta; -\xi\right).$$

4.3. Theorem. *Let* $(X_1, \ldots, X_n) \sim \mathrm{GH}(\alpha_1, \ldots, \alpha_n, \beta, \gamma, \xi)$ *and for* $s = 1, \ldots, n-1$*, define the random variables* $Y_i = X_i/(1-\sum_{i=1}^s X_i)$*,* $i = s+1, \ldots, n$*. Then, the p.d.f. of*

$(Y_{s+1},\dots,Y_n)$ *is*

$$K_3(\alpha_1,\dots,\alpha_n,\beta,\gamma,\xi)\frac{\prod_{i=s+1}^n y_i^{\alpha_i-1}(1-\sum_{i=s+1}^n y_i)^{\beta-1}}{\left(1+\xi\sum_{i=s+1}^n y_i\right)^{\gamma}}$$
$$\times\,{}_2F_1\left(\sum_{i=1}^s \alpha_i,\gamma;\sum_{i=1}^n \alpha_i+\beta;-\frac{\xi(1-\sum_{i=s+1}^n y_i)}{1+\xi\sum_{i=s+1}^n y_i}\right),\quad y_i>0,\quad i=1,\dots,n,$$

where

$$[K_3(\alpha_1,\dots,\alpha_n,\beta,\gamma,\xi)]^{-1}=\frac{\prod_{i=s+1}^n\Gamma(\alpha_i)\Gamma(\beta)}{\Gamma(\sum_{i=s+1}^n\alpha_i+\beta)}\,{}_2F_1\left(\sum_{i=1}^n\alpha_i,\gamma;\sum_{i=1}^n\alpha_i+\beta;-\xi\right).$$

Proof. Applying the transformation $Y_i=X_i/(1-\sum_{i=1}^s X_i)$, $i=s+1,\dots,n$ with the Jacobian $J(x_{s+1},\dots,x_n\to y_{s+1},\dots,y_n)=(1-\sum_{i=1}^s x_i)^{n-s}$ in (3.1) and integrating with respect to $x_1,\dots,x_s$, we obtain the p.d.f. of $(Y_{s+1},\dots,Y_n)$ as

$$C(\alpha_1,\dots,\alpha_n,\beta,\gamma,\xi)\ \frac{\prod_{i=s+1}^n y_i^{\alpha_i-1}(1-\sum_{i=s+1}^n y_i)^{\beta-1}}{(1+\xi\sum_{i=s+1}^n y_i)^{\gamma}}$$
$$\times\underset{\substack{x_1>0,\dots,x_s>0,\\ \sum_{i=1}^s x_i<1}}{\int\cdots\int}\frac{\prod_{i=1}^s x_i^{\alpha_i-1}(1-\sum_{i=1}^s x_i)^{\sum_{i=s+1}^n\alpha_i+\beta-1}\,\mathrm{d}x_1\cdots\mathrm{d}x_s}{\left[1+\xi(1-\sum_{i=s+1}^n y_i)\sum_{i=1}^s x_i/(1+\xi\sum_{i=s+1}^n y_i)\right]^{\gamma}},$$

where $0<y_i, i=s+1,\dots,n$ and $\sum_{i=s+1}^n y_i<1$. Now, evaluating the above integral using the Liouville-Dirichlet integral (2.7), we get

$$C(\alpha_1,\dots,\alpha_n,\beta,\gamma,\xi)\ \frac{\prod_{i=s+1}^n y_i^{\alpha_i-1}(1-\sum_{i=s+1}^n y_i)^{\beta-1}}{(1+\xi\sum_{i=s+1}^n y_i)^{\gamma}}$$
$$\times\frac{\prod_{i=1}^s\Gamma(\alpha_i)}{\Gamma(\sum_{i=1}^s\alpha_i)}\frac{\Gamma(\sum_{i=1}^s\alpha_i)\Gamma(\sum_{i=s+1}^n\alpha_i+\beta)}{\Gamma(\sum_{i=1}^n\alpha_i+\beta)}$$
$$\times\,{}_2F_1\left(\sum_{i=1}^s\alpha_i,\gamma;\sum_{i=1}^n\alpha_i+\beta;-\frac{\xi(1-\sum_{i=s+1}^n y_i)}{1+\xi\sum_{i=s+1}^n y_i}\right).$$

Finally, substituting for $C(\alpha_1,\dots,\alpha_n,\beta,\gamma,\xi)$ in the above expression and simplifying, we obtain the desired result. □

The following theorem gives the distribution of partial sums of random variables whose joint distribution is multivariate Gauss hypergeometric.

4.4. Theorem. *Let* $(X_1,\dots,X_n)\sim\mathrm{GH}(\alpha_1,\dots,\alpha_n,\beta,\gamma,\xi)$ *and* $n_1,\dots,n_\ell$ *be non-negative integers such as* $\sum_{i=1}^\ell n_i=n$. *Define,* $\alpha_{(i)}=\sum_{j=n_{i-1}^*+1}^{n_i^*}\alpha_i$, $n_0^*=0$, $n_i^*=\sum_{j=1}^i n_j$, $i=1,\dots,\ell$, $Z_j=X_j/X_{(i)}$, $j=n_{i-1}^*+1,\dots,n_i^*-1$ *and* $X_{(i)}=\sum_{j=n_{i-1}^*+1}^{n_i^*}X_j$, $i=1,\dots,\ell$. *Then*

(i) $(X_{(1)},\dots,X_{(\ell)})$ *and* $(Z_{n_{i-1}^*+1},\dots,Z_{n_i^*-1})$, $i=1,\dots,\ell$, *are independently distributed,*

(ii) $(X_{(1)},\dots,X_{(\ell)})\sim\mathrm{GH}(\alpha_{(1)},\dots,\alpha_{(\ell)},\beta,\gamma,\xi)$ *and*

(iii) $(Z_{n_{i-1}^*+1},\dots,Z_{n_i^*-1})\sim\mathrm{D1}(\alpha_{n_{i-1}^*+1},\dots,\alpha_{n_i^*-1};\alpha_{n_i^*})$, $i=1,\dots,\ell$.

Proof. Transforming $Z_j=X_j/X_{(i)}$ and $X_{(i)}=\sum_{j=n_{i-1}^*+1}^{n_i^*}X_j$, $j=n_{i-1}^*+1,\dots,n_i^*-1$, $i=1,\dots,\ell$, with the Jacobian

$$J(x_1,\dots,x_n\to z_1,\dots,z_{n_1-1},x_{(1)},\dots,z_{n_{\ell-1}^*+1},\dots,z_{n_\ell^*-1},x_{(\ell)})=\prod_{i=1}^{\ell}x_{(i)}^{n_i-1},$$

in the p.d.f. (3.1), we obtain the joint p.d.f. of $Z_{n^*_{i-1}+1}, \ldots, Z_{n^*_i-1}, X_{(i)}, i=1,\ldots,\ell$, as being proportional to

$$\text{(4.3)} \qquad \frac{\prod_{i=1}^{\ell} x_{(i)}^{\alpha_{(i)}-1}(1-\sum_{i=1}^{\ell} x_{(i)})^{\beta-1}}{(1+\xi\sum_{i=1}^{\ell} x_{(i)})^{\gamma}} \prod_{i=1}^{\ell}\left[\prod_{j=n^*_{i-1}+1}^{n^*_i-1} z_j^{\alpha_j-1}\left(1-\sum_{j=n^*_{i-1}+1}^{n^*_i-1} z_j\right)^{\alpha_{n^*_i}-1}\right],$$

where $x_{(i)}>0$, $i=1,\ldots,\ell$, $\sum_{i=1}^{\ell} x_{(i)}<1$, $z_j>0$, $j=n^*_{i-1}+1,\ldots,n^*_i-1$, $\sum_{j=n^*_{i-1}+1}^{n^*_i-1} z_j<1, i=1,\ldots,\ell$. From the factorization (4.3), it is clear that $(X_{(1)},\ldots,X_{(\ell)})$ and $(Z_{n^*_{i-1}+1},\ldots,Z_{n^*_i-1})$, $i=1,\ldots,\ell$, are independently distributed. $(X_{(1)},\ldots,X_{(\ell)}) \sim \mathrm{GH}(\alpha_{(1)},\ldots,\alpha_{(\ell)},\beta,\gamma,\xi)$ and $(Z_{n^*_{i-1}+1},\ldots,Z_{n^*_i-1}) \sim \mathrm{D1}(\alpha_{n^*_{i-1}+1},\ldots,\alpha_{n^*_i-1};\alpha_{n^*_i})$, $i=1,\ldots,\ell$. □

4.5. Corollary. *Let* $(X_1,\ldots,X_n) \sim \mathrm{GH}(\alpha_1,\ldots,\alpha_n,\beta,\gamma,\xi)$ *and define* $Z_i = X_i/Z$ *for* $i=1,\ldots,n-1$ *and* $Z=\sum_{j=1}^n X_j$. *Then,* $(Z_1, \ldots, Z_{n-1})$ *and* Z *are independent,* $(Z_1,\ldots,Z_{n-1}) \sim \mathrm{D1}\,(\alpha_1,\ldots,\alpha_{n-1};\alpha_n)$ *and* $Z \sim \mathrm{GH}(\sum_{i=1}^n \alpha_i,\beta,\gamma,\xi)$.

4.6. Corollary. *If* $(X_1,\ldots,X_n) \sim \mathrm{GH}(\alpha_1,\ldots,\alpha_n,\beta,\gamma,\xi)$, *then*

$$\frac{\sum_{i=1}^s X_i}{\sum_{j=1}^n X_j} \sim \mathrm{B1}\left(\sum_{i=1}^s \alpha_i, \sum_{i=s+1}^n \alpha_i\right), \quad s<n.$$

4.7. Corollary. *If* $(X_1,\ldots,X_n) \sim \mathrm{GH}(\alpha_1,\ldots,\alpha_n,\beta,\gamma,\xi)$, *then* $(X_1,\ldots,X_i+X_j,\ldots,X_n) \sim \mathrm{GH}(\alpha_1,\ldots,\alpha_i+\alpha_j,\ldots,\alpha_n,\beta,\gamma,\xi)$.

4.8. Theorem. *Let* $(X_1,\ldots,X_n) \sim \mathrm{GH}(\alpha_1,\ldots,\alpha_n,\beta,\gamma,\xi)$ *and* $n_1,\ldots,n_\ell$ *be non-negative integers such as* $\sum_{i=1}^{\ell} n_i = n$. *Define,* $\alpha_{(i)} = \sum_{j=n^*_{i-1}+1}^{n^*_i} \alpha_i$, $n^*_0=0$, $n^*_i=\sum_{j=1}^i n_j$, $i=1,\ldots,\ell$, $W_j = X_j/X_{n^*_i}$, $j=n^*_{i-1}+1,\ldots,n^*_i-1$ *and* $X_{(i)}=\sum_{j=n^*_{i-1}+1}^{n^*_i} X_j$, $i=1,\ldots,\ell$. *Then*

- *(i)* $(X_{(1)},\ldots,X_{(\ell)})$ *and* $(W_{n^*_{i-1}+1},\ldots,W_{n^*_i-1})$, $i=1,\ldots,\ell$, *are independently distributed,*
- *(ii)* $(X_{(1)},\ldots,X_{(\ell)}) \sim \mathrm{GH}(\alpha_{(1)},\ldots,\alpha_{(\ell)},\beta,\gamma,\xi)$ *and*
- *(iii)* $(W_{n^*_{i-1}+1},\ldots,W_{n^*_i-1}) \sim \mathrm{D2}(\alpha_{n^*_{i-1}+1},\ldots,\alpha_{n^*_i-1};\alpha_{n^*_i})$, $i=1,\ldots,\ell$.

4.9. Corollary. *Let* $(X_1,\ldots,X_n) \sim \mathrm{GH}(\alpha_1,\ldots,\alpha_n,\beta,\gamma,\xi)$ *and define* $W_i = X_i/X_n$ *for* $i=1,\ldots,n-1$ *and* $Z=\sum_{j=1}^n X_j$. *Then,* $(W_1, \ldots, W_{n-1})$ *and* Z *are independent,* $(Z_1,\ldots,Z_{n-1}) \sim \mathrm{D2}\,(\alpha_1,\ldots,\alpha_{n-1};\alpha_n)$ *and* $Z \sim \mathrm{GH}(\sum_{i=1}^n \alpha_i,\beta,\gamma,\xi)$.

4.10. Corollary. *If* $(X_1,\ldots,X_n) \sim \mathrm{GH}(\alpha_1,\ldots,\alpha_n,\beta,\gamma,\xi)$, *then*

$$\frac{\sum_{i=1}^s X_i}{\sum_{j=s+1}^n X_j} \sim \mathrm{B2}\left(\sum_{i=1}^s \alpha_i, \sum_{i=s+1}^n \alpha_i\right), \quad s<n.$$

5. Joint Moments

We derive the joint moments of random variables jointly distributed as multivariate Gauss hypergeometric. These moments will facilitate us to compute several expected values such as mean and variance.

Using (3.1) and (3.2), the the joint moments of $X_1, \ldots, X_n$ are obtained as

$$\begin{aligned} E(X_1^{r_1}\cdots X_n^{r_n}) &= \frac{C(\alpha_1,\ldots,\alpha_n,\beta,\gamma,\xi)}{C(\alpha_1+r_1,\ldots,\alpha_n+r_n,\beta,\gamma,\xi)} \\ &= \frac{\Gamma\left(\sum_{i=1}^n \alpha_i+\beta\right)\prod_{i=1}^n \Gamma(\alpha_i+r_i)}{\Gamma\left[\sum_{i=1}^n(\alpha_i+r_i)+\beta\right]\prod_{i=1}^n\Gamma(\alpha_i)} \\ &\quad \times \frac{{}_2F_1(\sum_{i=1}^n(\alpha_i+r_i),\gamma;\sum_{i=1}^n(\alpha_i+r_i)+\beta;-\xi)}{{}_2F_1(\sum_{i=1}^n\alpha_i,\gamma;\sum_{i=1}^n\alpha_i+\beta;-\xi)}. \end{aligned}$$

By substituting appropriately in the above expression, the following expected values can easily be obtained:

$$E(X_i) = \frac{\alpha_i}{\sum_{i=1}^n \alpha_i+\beta}\,\frac{{}_2F_1(\sum_{i=1}^n\alpha_i+1,\gamma;\sum_{i=1}^n\alpha_i+\beta+1;-\xi)}{{}_2F_1(\sum_{i=1}^n\alpha_i,\gamma;\sum_{i=1}^n\alpha_i+\beta;-\xi)},$$

$$\begin{aligned} E(X_i^2) &= \frac{\alpha_i(\alpha_i+1)}{(\sum_{i=1}^n\alpha_i+\beta)(\sum_{i=1}^n\alpha_i+\beta+1)} \\ &\quad\times \frac{{}_2F_1(\sum_{i=1}^n\alpha_i+2,\gamma;\sum_{i=1}^n\alpha_i+\beta+2;-\xi)}{{}_2F_1(\sum_{i=1}^n\alpha_i,\gamma;\sum_{i=1}^n\alpha_i+\beta;-\xi)}, \end{aligned}$$

$$\begin{aligned} \mathrm{Var}(X_i) &= \frac{\alpha_i}{\sum_{i=1}^n\alpha_i+\beta}\left[\frac{(\alpha_i+1)\ {}_2F_1(\sum_{i=1}^n\alpha_i+2,\gamma;\sum_{i=1}^n\alpha_i+\beta+2;-\xi)}{(\sum_{i=1}^n\alpha_i+\beta+1)\ {}_2F_1(\sum_{i=1}^n\alpha_i,\gamma;\sum_{i=1}^n\alpha_i+\beta;-\xi)}\right. \\ &\quad\left. -\frac{\alpha_i}{\sum_{i=1}^n\alpha_i+\beta}\left\{\frac{{}_2F_1(\sum_{i=1}^n\alpha_i+1,\gamma;\sum_{i=1}^n\alpha_i+\beta+1;-\xi)}{{}_2F_1(\sum_{i=1}^n\alpha_i,\gamma;\sum_{i=1}^n\alpha_i+\beta;-\xi)}\right\}^2\right], \end{aligned}$$

$$\begin{aligned} E(X_iX_j) &= \frac{\alpha_i\alpha_j}{(\sum_{i=1}^n\alpha_i+\beta)(\sum_{i=1}^n\alpha_i+\beta+1)} \\ &\quad\times\frac{{}_2F_1(\sum_{i=1}^n\alpha_i+2,\gamma;\sum_{i=1}^n\alpha_i+\beta+2;-\xi)}{{}_2F_1(\sum_{i=1}^n\alpha_i,\gamma;\sum_{i=1}^n\alpha_i+\beta;-\xi)},\quad i\neq j, \end{aligned}$$

and finally for $i \neq j$,

$$\begin{aligned} \mathrm{Cov}(X_i,X_j) &= \frac{\alpha_i\alpha_j}{\sum_{i=1}^n\alpha_i+\beta}\left[\frac{{}_2F_1(\sum_{i=1}^n\alpha_i+2,\gamma;\sum_{i=1}^n\alpha_i+\beta+2;-\xi)}{(\sum_{i=1}^n\alpha_i+\beta+1){}_2F_1(\sum_{i=1}^n\alpha_i,\gamma;\sum_{i=1}^n\alpha_i+\beta;-\xi)}\right. \\ &\quad\left. -\frac{1}{\sum_{i=1}^n\alpha_i+\beta}\left\{\frac{{}_2F_1(\sum_{i=1}^n\alpha_i+1,\gamma;\sum_{i=1}^n\alpha_i+\beta+1;-\xi)}{{}_2F_1(\sum_{i=1}^n\alpha_i,\gamma;\sum_{i=1}^n\alpha_i+\beta;-\xi)}\right\}^2\right]. \end{aligned}$$

Using the definition and the above expressions, one can calculate the correlation between X_i and X_j.

6. Factorizations

In this section we give several factorizations of the multivariate Gauss hypergeometric p.d.f..

6.1. Theorem. *Let* $(X_1,\ldots,X_n)\sim \mathrm{GH}(\alpha_1,\ldots,\alpha_n,\beta,\gamma,\xi)$. *For* $i=1,\ldots,n-1$ *define* $Y_i=\sum_{j=1}^i X_j/\sum_{j=1}^{i+1}X_j$ *and* $Y_n=\sum_{j=1}^n X_j$. *Then, the random variables* $Y_1,\ldots,Y_n$ *are independent,* $Y_i\sim \mathrm{B1}(\sum_{j=1}^i\alpha_j,\alpha_{i+1})$, $i=1,\ldots,n-1$ *and* $Y_n\sim\mathrm{GH}(\sum_{i=1}^n\alpha_i,\beta,\gamma,\xi)$.

Proof. From the transformation given in the theorem, we obtain $x_1 = y_n\prod_{i=1}^{n-1}y_i$, $x_2=y_n(1-y_1)\prod_{i=2}^{n-1}y_i,\ldots,x_{n-1}=y_n(1-y_{n-2})y_{n-1}$, and $x_n=y_n(1-y_{n-1})$. with the Jacobian $J(x_1,\dots,x_n\to y_1,\dots,y_n)=\prod_{i=2}^n y_i^{i-1}$. Now, making appropriate substitutions in the joint p.d.f. of $X_1,\dots,X_n$, we obtain

$$C(\alpha_1,\dots,\alpha_n,\beta,\gamma,\xi)\ \left(y_n\prod_{i=1}^{n-1}y_i\right)^{\alpha_1-1}\prod_{j=2}^{n}\left[y_n(1-y_{j-1})\prod_{i=j}^{n-1}y_i\right]^{\alpha_j-1}$$
$$\times\frac{(1-y_n)^{\beta-1}}{(1+\xi y_n)^{\gamma}}\prod_{i=2}^{n}y_i^{i-1}.$$

Further, writing

$$C(\alpha_1,\dots,\alpha_n,\beta,\gamma,\xi)=\prod_{j=1}^{n-1}\left[B\left(\sum_{i=1}^{j}\alpha_i,\alpha_{j+1}\right)\right]^{-1}$$
$$\times\left[B\left(\sum_{i=1}^{n}\alpha_i,\beta\right){}_2F_1\left(\sum_{i=1}^{n}\alpha_i,\gamma;\sum_{i=1}^{n}\alpha_i+\beta;-\xi\right)\right]^{-1},$$

the above expression is simplified as

$$\left[\prod_{j=1}^{n-1}\frac{y_j^{\sum_{i=1}^{j}\alpha_i-1}(1-y_j)^{\alpha_{j+1}-1}}{B(\sum_{i=1}^{j}\alpha_i,\alpha_{j+1})}\right]C\left(\sum_{i=1}^{n}\alpha_i,\beta,\gamma,\xi\right)\frac{y_n^{\sum_{i=1}^{n}\alpha_i-1}(1-y_n)^{\beta-1}}{(1+\xi y_n)^{\gamma}},$$

where $0<y_1,\dots,y_n<1$. Now, from the above factorization, we get the result. □

6.2. Theorem. *Let* $(X_1,\dots,X_n)\sim\mathrm{GH}(\alpha_1,\dots,\alpha_n,\beta,\gamma,\xi)$. *Define* $Z_n=\sum_{j=1}^n X_j$ *and* $Z_i=X_{i+1}/\sum_{j=1}^i X_j$, *for* $i=1,\dots,n-1$. *Then,* $Z_1,\ \dots,\ Z_n$ *are independent,* $Z_i\sim\mathrm{B2}(\alpha_{i+1},\sum_{j=1}^i\alpha_j)$, $i=1,\dots,n-1$ *and* $Z_n\sim\mathrm{GH}(\sum_{i=1}^n\alpha_i,\beta,\gamma,\xi)$.

Proof. This result is obtain from Theorem 6.1, by observing that $Z_i=(1-Y_i)/Y_i$, for $i=1,\dots,n-1$, $Z_n=Y_n$ and $(1-Y_i)/Y_i\sim\mathrm{B2}(\alpha_{i+1},\sum_{j=1}^i\alpha_j)$, where $Y_i\sim\mathrm{B1}(\sum_{j=1}^i\alpha_j,\alpha_{i+1})$. □

6.3. Theorem. *Let* $(X_1,\dots,X_n)\sim\mathrm{GH}(\alpha_1,\dots,\alpha_n,\beta,\gamma,\xi)$. *Define* $W_n=\sum_{j=1}^n X_j$ *and* $W_i=\sum_{j=1}^i X_j/X_{i+1}$, *for* $i=1,\dots,n-1$. *Then,* $W_1,\ \dots,\ W_n$, *are independent,* $W_i\sim\mathrm{B2}(\sum_{j=1}^i\alpha_j,\alpha_{i+1})$ *for* $i=1,\dots,n-1$ *and* $W_n\sim\mathrm{GH}(\sum_{i=1}^n\alpha_i,\beta,\gamma,\xi)$.

Proof. The result is obtained from Theorem 6.2 by taking into account that $W_i=1/Z_i$ for $i=1,\dots,n-1$, $W_n=Z_n$ and $1/Z_i\sim\mathrm{B2}(\sum_{j=1}^i\alpha_j,\alpha_{i+1})$, where $Z_i\sim\mathrm{B2}(\alpha_{i+1},\sum_{j=1}^i\alpha_j)$. □

6.4. Theorem. *Let* $(X_1,\dots,X_n)\sim\mathrm{GH}(\alpha_1,\dots,\alpha_n,\beta,\gamma,\xi)$. *Define* $Y_n=\sum_{j=1}^n X_j$ *and for* $i=1,\dots,n-1$ $Y_i=X_i/\sum_{j=i}^n X_j$. *Then,* $Y_1,\dots,Y_n$ *are independent,* $Y_i\sim\mathrm{B1}(\alpha_i,\sum_{j=i+1}^n\alpha_j)$, *for* $i=1,\dots,n-1$ *and* $Y_n\sim\mathrm{GH}(\sum_{i=1}^n\alpha_i,\beta,\gamma,\xi)$.

Proof. Making the substitution $x_1=y_ny_1$, $x_2=y_ny_2(1-y_1)$, $\dots$, $x_{n-1}=y_ny_{n-1}(1-y_1)\cdots(1-y_{n-2})$ and $x_n=y_n(1-y_1)\cdots(1-y_{n-1})$ with the Jacobian $J(x_1,\dots,x_n\to y_1,\dots,y_n)=y_n^{n-1}\prod_{i=1}^{n-2}(1-y_i)^{n-i-1}$ in (3.1), we obtain the joint p.d.f. of $Y_1,\dots,Y_n$ as

$$C(\alpha_1,\dots,\alpha_n,\beta,\gamma,\xi)\prod_{i=1}^{n-1}\left[y_ny_i\prod_{j=1}^{i-1}(1-y_j)\right]^{\alpha_i-1}\left[y_n\prod_{j=1}^{n-1}(1-y_j)\right]^{\alpha_n-1}$$
$$\times\frac{(1-y_n)^{\beta-1}}{(1+\xi y_n)^{\gamma}}y_n^{n-1}\prod_{i=1}^{n-2}(1-y_i)^{n-i-1},$$

which can be re-written as

$$\left[\prod_{i=1}^{n-1} \frac{y_i^{\alpha_i-1}(1-y_i)^{\sum_{j=i+1}^{n}\alpha_j-1}}{B(\alpha_i, \sum_{j=i+1}^{n}\alpha_j)}\right] C\left(\sum_{i=1}^{n}\alpha_i, \beta, \gamma, \xi\right) \frac{y_n^{\sum_{i=1}^{n}\alpha_i-1}(1-y_n)^{\beta-1}}{(1+\xi y_n)^{\gamma}}.$$

Now, the desired result follows from the above factorization. □

6.5. Theorem. *Let* $(X_1, \ldots, X_n) \sim \mathrm{GH}(\alpha_1, \ldots, \alpha_n, \beta, \gamma, \xi)$. *Define* $Z_n = \sum_{j=1}^{n} X_j$ *and* $Z_i = X_i / \sum_{j=i+1}^{n} X_j$, *for* $i = 1, \ldots, n-1$. *Then,* $Z_1, \ldots, Z_n$, *are independent,* $Z_i \sim \mathrm{B2}(\alpha_i, \sum_{j=i+1}^{n}\alpha_j)$, *for* $i = 1, \ldots, n-1$ *and* $Z_n \sim \mathrm{GH}(\sum_{i=1}^{n}\alpha_i, \beta, \gamma, \xi)$.

Proof. The result is obtained from Theorem 6.4, by noting that $Z_i = Y_i/(1-Y_i)$ for $i = 1, \ldots, n-1$, $Z_n = Y_n$ and $Y_i/(1-Y_i) \sim \mathrm{B2}(\alpha_i, \sum_{j=i+1}^{n}\alpha_j)$ for $Y_i \sim \mathrm{B1}(\alpha_i, \sum_{j=i+1}^{n}\alpha_j)$. □

6.6. Theorem. *Let* $(X_1, \ldots, X_n) \sim \mathrm{GH}(\alpha_1, \ldots, \alpha_n, \beta, \gamma, \xi)$. *Define* $W_n = \sum_{j=1}^{n} X_j$ *and* $W_i = \sum_{j=i+1}^{n} X_j / X_i$, *for* $i = 1, \ldots, n-1$. *Then,* $W_1, \ldots, W_n$ *are independent,* $W_i \sim \mathrm{B2}(\sum_{j=i+1}^{n}\alpha_j, \alpha_i)$ *for* $i = 1, \ldots, n-1$ *and* $W_n \sim \mathrm{GH}(\sum_{i=1}^{n}\alpha_i, \beta, \gamma, \xi)$.

Proof. This result follows from Theorem 6.5, by observing that $W_i = 1/Z_i$ for $i = 1, \ldots, n-1$, $W_n = Z_n$ and $1/Z_i \sim \mathrm{B2}(\sum_{j=i+1}^{n}\alpha_j, \alpha_i)$, where $Z_i \sim \mathrm{B2}(\alpha_i, \sum_{j=i+1}^{n}\alpha_j)$. □

7. Data application

Here, we illustrate the use of the multivariate Gauss hypergeometric p.d.f. We use the following data taken from Aitchison [2]:

	sand	silt	clay
1	0.775	0.195	0.030
2	0.719	0.249	0.032
3	0.507	0.361	0.132
4	0.522	0.409	0.066
5	0.700	0.265	0.035
6	0.665	0.322	0.013
7	0.431	0.553	0.016
8	0.534	0.368	0.098
9	0.155	0.544	0.301
10	0.317	0.415	0.268
11	0.657	0.278	0.065
12	0.704	0.290	0.006
13	0.174	0.536	0.290
14	0.106	0.698	0.196
15	0.382	0.431	0.187
16	0.108	0.527	0.365
17	0.184	0.507	0.309
18	0.046	0.474	0.480
19	0.156	0.504	0.340
20	0.319	0.451	0.230
21	0.095	0.535	0.370
22	0.171	0.480	0.349
23	0.105	0.554	0.341
24	0.048	0.547	0.410
25	0.026	0.452	0.522
26	0.114	0.527	0.359

27	0.067	0.469	0.464
28	0.069	0.497	0.434
29	0.040	0.449	0.511
30	0.074	0.516	0.409
31	0.048	0.495	0.457
32	0.045	0.485	0.470
33	0.066	0.521	0.413
34	0.067	0.473	0.459
35	0.074	0.456	0.469
36	0.060	0.489	0.451
37	0.063	0.538	0.399
38	0.025	0.480	0.495
39	0.020	0.478	0.502

The data are on the sediment composition in an Arctic lake. The second column gives relative frequencies of sand. The third column gives relative frequencies of silt. The fourth column gives relative frequencies of clay.

We fitted the multivariate Gauss hypergeometric p.d.f. in (3.1) to the pairwise data on (sand, clay) and (silt, clay). We also fitted the Dirichlet p.d.f., the particular case of (3.1) for $\gamma = 0$ and $\xi = 0$. The method of maximum likelihood was used for the fitting.

For the first pair, we obtained the estimates:

- $\widehat{\alpha}_1 = 6.574 \times 10^{-1}(1.330 \times 10^{-1})$, $\widehat{\alpha}_2 = 7.957 \times 10^{-1}(1.686 \times 10^{-1})$, $\widehat{\beta} = 10.452(2.339)$, $\widehat{\gamma} = -11.033(2.818)$, $\widehat{\xi} = 24540.9(1288584)$ with $\log L = 60.7$ for the bivariate Gauss hypergeometric p.d.f.
- $\widehat{\alpha}_1 = 1.021(1.698 \times 10^{-1})$, $\widehat{\alpha}_2 = 1.299(2.186 \times 10^{-1})$, $\widehat{\beta} = 2.319(4.004 \times 10^{-1})$ with $\log L = 39.5$ for the Dirichlet p.d.f.

For the second pair, we obtained the estimates:

- $\widehat{\alpha}_1 = 4.182(9.407 \times 10^{-1})$, $\widehat{\alpha}_2 = 2.168(4.617 \times 10^{-1})$, $\widehat{\beta} = 7.802 \times 10^{-1}(1.638 \times 10^{-1})$, $\widehat{\gamma} = 4.072(1.635)$, $\widehat{\xi} = 302.0(5031.1)$ with $\log L = 44.6$ for the bivariate Gauss hypergeometric p.d.f.
- $\widehat{\alpha_1} = 2.316(3.999 \times 10^{-1})$, $\widehat{\alpha_2} = 1.297(2.183 \times 10^{-1})$, $\widehat{\beta} = 1.019 \times 10^{-1}(1.695 \times 10^{-1})$ with $\log L = 39.5$ for the Dirichlet p.d.f.

Here, $\log L$ denotes the maximized log-likelihood value and the numbers within brackets are the standard errors computed by inverting the observation information matrices.

It follows by the standard likelihood ratio test that the bivariate Gauss hypergeometric distribution provides a significantly better fit for both data sets. Contours of the fitted bivariate Gauss hypergeometric p.d.f.s are shown in Figures 2 and 3. Also shown in the figures are the actual observed data. The fitted p.d.f.s do appear to capture the pattern in the data.

Acknowledgement

The research work of DKN and DB-V was supported by the Comité para el Desarrollo de la Investigación, Universidad de Antioquia research grant No. IN560CE.

References

[1] C. Armero and M. Bayarri, Prior assessments for predictions in queues, *The Statistician*, **43** (1994), no. 1, 139–153.

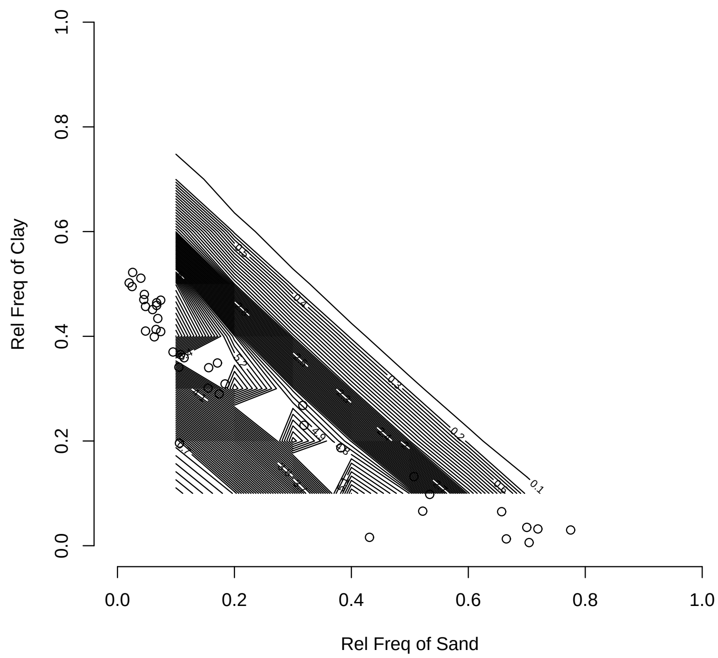

Figure 2. Relative frequencies of sand and clay and contours of the fitted bivariate Gauss hypergeometric p.d.f.

[2] J. Aitchison, *The Statistical Analysis of Compositional Data*, The Blackburn Press, Caldwell, NJ, 2003.

[3] Liliam Cardeño, Daya K. Nagar and Luz Estela Sánchez, Beta type 3 distribution and its multivariate generalization, *Tamsui Oxford Journal of Mathematical Sciences*, **21**(2005), no. 2, 225–241.

[4] J.-Y. Dauxois, Bayesian inference for linear growth birth and death processes, *Journal of Statistical Planning and Inference*, **121** (2004), no. 1, 1–19.

[5] Peter S. Fader and Bruce G. S. Hardie, A note on modelling underreported Poisson counts, *Journal of Applied Statistics*, **27** (2000), no. 8, 953–964 .

[6] A. K. Gupta and D. K. Nagar, *Matrix Variate Distributions*, Chapman & Hall/CRC, Boca Raton, 2000.

[7] A. K. Gupta and D. K. Nagar, Properties of matrix variate beta type 3 distribution, *International Journal of Mathematics and Mathematical Sciences*, **2009**(2009), Art. ID 308518, 18 pp.

[8] Rameshwar D. Gupta and Donald St. P. Richards, The history of Dirichlet and Liouville distributions, *Internationals Statistical Review*, **69**(2001), no. 3, 433-446.

[9] A. K. Gupta and D. Song, Generalized Lioville distribution, *Computters & Mathematics with Applicattions*, **32**(1996), no. 2, 103-109.

[10] N. L. Johnson, S. Kotz and N. Balakrishnan, *Continuous Univariate Distributions-2*, Second Edition, John Wiley & Sons, New York, 1994.

[11] S. Kotz, N. Balakrishnana and N. L. Johnson, *Continuous Multivariate Distributions-1*, Second Edition, John Wiley & Sons, New York (2000).

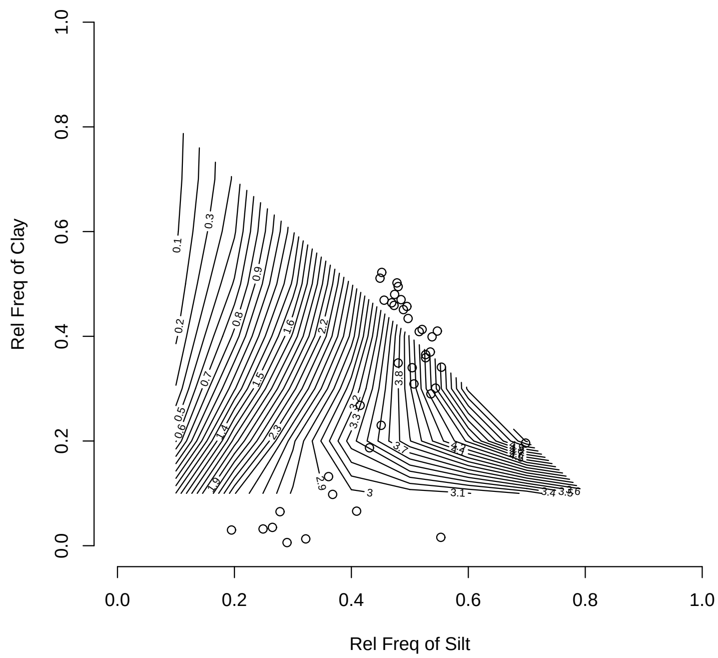

Figure 3. Relative frequencies of silt and clay and contours of the fitted bivariate Gauss hypergeometric p.d.f.

[12] D. L. Libby and M. R. Novic, Multivariate generalized beta distributions with applications to utility assessment, *J. Educ. Statist*, **7** (1982), no. 4, 271–294.

[13] Y. L. Luke, *The Special Functions and Their Approximations, Vol. 1*, Academic Press, New York, 1969.

[14] 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 - London, 1979.

[15] S. Nadarajah, Sums, products and ratios of generalized beta variables, *Statistical Papers*, **47**(2006), no. 1, 69–90.

[16] Daya K. Nagar, Paula Andrea Bran-Cardona and A. K. Gupta, Multivariate generalization of the hypergeometric function type 1 distribution, *Acta Applicandae Mathematicae*, **105**(2009), no. 1, 111-122.

[17] Daya K. Nagar and Erika Alejandra Rada-Mora, Properties of multivariate beta distributions, *Far East Journal of Theoretical Statistics*, **24**(2008), no. 1, 73–94.

[18] Daya K. Nagar and Fabio Humberto Sepúlveda-Murillo, Multivariate generalization of the confluent hypergeometric function kind 1 distribution, *International Journal of Mathematics and Mathematical Sciences*, **2008**, Art. ID 152808, 13 pp.

[19] T. Pham-Gia and Q. P. Duong, The generalized beta- and F-distributions in statistical modelling, *Mathematics and Computer Modelling*, **12** (1989), no.12, 1613-1625.

[20] Luz Estela Sánchez and Daya K. Nagar, Distribution of the product and the quotient of independent beta type 3 variables, *Far East Journal of Theoretical Statistics*, **17**(2005), no. 2, 239–251.

[21] José María Sarabia and Enrique Castillo, Bivariate distributions based on the generalized three-parameter beta distribution, *Advances in Distribution Theory, Order Statistics, and Inference*, Stat. Ind. Technol., Birkhäuser Boston, Boston, MA, 2006.

[22] B. D. Sivazlian, On multivariate extensión of the gamma and beta distributions, *Siam Journal of Applied Mathematics*, **41**(1981), no. 2, 205-209.

[23] D. Song and A. K. Gupta, Properties of generalized Liouville distribution, *Random Operators and Stochastic Equations*, **5**(1997), no. 5, 337-348.