Abstract: In this article, we define and study the Dirichlet-Lauricella type D distribution. This distribution is a generalization of the Dirichlet type 1 distribution.

AMS Subject Classification: 62H99
Key Words: Appell hypergeometric function, Dirichlet distribution, hypergeometric function, Lauricella function

1. Introduction

The random variables $U_1, \ldots, U_n$ are said to have a Dirichlet type 1 distribution with parameters $a_1, \ldots, a_n; a_{n+1}$, denoted by $(U_1, \ldots, U_n) \sim D1(a_1, \ldots, a_n; a_{n+1})$, if their joint probability density function (p.d.f.) is given by

$$f(u_1, \ldots, u_n) = \frac{\Gamma(\sum_{i=1}^{n+1} a_i)}{\prod_{i=1}^{n+1} \Gamma(a_i)} \prod_{i=1}^{n} u_i^{a_i-1} \left(1 - \sum_{i=1}^{n} u_i\right)^{a_{n+1}-1},$$

$$u_i > 0, \quad i = 1, \ldots, n, \quad \sum_{i=1}^{n} u_i < 1.$$  (1)

The Dirichlet type 1 distribution is a multivariate generalization of the beta distribution. Dirichlet distributions are very often used as prior distributions in Bayesian statistics, and in fact the Dirichlet distribution is the conjugate prior.
of the categorical distribution and multinomial distribution. The Dirichlet type 1 distribution has been studied extensively, for example, see Kotz, Balakrishnan and Johnson [2], and Gupta and Nagar [3].

In this article, we give a generalization of the Dirichlet type 1 distribution. This generalization is based on the Lauricella type D hypergeometric function and thus will be called Dirichlet-Lauricella type D distribution. Recently, Nagar and Gómez [8] have proposed a generalization of Dirichlet type 1 distribution based on the Lauricell’s type B hypergeometric function.

In Section 2, we give definition of Lauricella type D hypergeometric function. We define Dirichlet-Lauricella type D distribution in Section 3. Section 4, deals with several properties such as marginal densities and joint moment.

2. Preliminaries

The Pochhammer symbol \((a)_n\) is 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 Lauricella hypergeometric functions \(F_D^{(n)}\) of several variables is defined as

\[
F_D^{(n)}(a, b_1, \ldots, b_n; c; z_1, \ldots, z_n) = \sum_{j_1, \ldots, j_n=0}^{\infty} \frac{(a)_{j_1+\cdots+j_n} (b_1)_{j_1} \cdots (b_n)_{j_n} z_1^{j_1} \cdots z_n^{j_n}}{(c)_{j_1+\cdots+j_n} j_1! \cdots j_n!}, \quad \max\{|z_1|, \ldots, |z_n|\} < 1. \tag{2}
\]

For \(n = 1\), the Lauricella hypergeometric functions \(F_D^{(n)}\) reduces a Gauss hypergeometric function and for and \(n = 2\) it slides to an Appell hypergeometric function \(F_1\).

The integral representations of \(F_D^{(n)}\) is given by

\[
F_D^{(n)}(a, b_1, \ldots, b_n; c; z_1, \ldots, z_n) = \frac{\Gamma(c)}{\prod_{i=1}^n \Gamma(b_i) \Gamma(c - \sum_{i=1}^n b_i)} \cdot \int_{u_1>0, \ldots, u_n>0} \frac{\prod_{i=1}^n u_i^{b_i-1} (1 - \sum_{i=1}^n u_i)^{c - \sum_{i=1}^n a_i - 1}}{(1 - \sum_{i=1}^n u_i z_i)^a} \prod_{i=1}^n du_i, \tag{3}
\]

where \(\text{Re}(b_i) > 0, i = 1, \ldots, n\) and \(\text{Re}(c - b_1 - \cdots - b_n) > 0\).

Another representation of \(F_D^{(n)}\), in terms of a single integral, is given by

\[
F_D^{(n)}(a, b_1, \ldots, b_n; c; z_1, \ldots, z_n) = \frac{\Gamma(c)}{\Gamma(a) \Gamma(c - a)} \int_0^1 \frac{u^{a-1} (1 - u)^{c-a-1}}{(1 - uz_i)^{b_i}} du,
\]
where \( \text{Re}(c) > \text{Re}(a) > 0 \) and \( |\text{arg}(1 - z_i)| < \pi, i = 1, \ldots, n. \)

For further results and properties of this function the reader is referred to Exton [1], Srivastava and Karlsson [11], and Prudnikov, Brychkov and Marichev [10, Sec. 7.2.4].

Let \( f(\cdot) \) be a continuous function and \( \alpha_i > 0, i = 1, \ldots, r \). The integral

\[
D_r(\alpha_1, \ldots, \alpha_r; f) = \int \cdots \int \prod_{i=1}^{r} x_i^{\alpha_i-1} f\left(\sum_{i=1}^{r} x_i\right) \prod_{i=1}^{r} dx_i
\]  

is known as the Liouville-Dirichlet integral. Substituting \( y_i = x_i/x, i = 1, \ldots, r-1 \) and \( x = \sum_{i=1}^{r} x_i \) with the Jacobian \( J(x_1, \ldots, x_{r-1}, x_r \to y_1, \ldots, y_{r-1}, x) = x^{r-1} \) it is easy to see that

\[
D_n(\alpha_1, \ldots, \alpha_r; f) = \frac{\prod_{i=1}^{r} \Gamma(\alpha_i)}{\Gamma(\sum_{i=1}^{r} \alpha_i)} \int_0^{1} x^{\sum_{i=1}^{r} \alpha_i-1} f(x) \, dx.
\]  

Further, by taking \( f(x) = (1-x)^{\alpha_{r+1}}, \alpha_{r+1} > 0 \) in (4) and (5), the classical Dirichelet integral is evaluated as

\[
\int \cdots \int \prod_{i=1}^{r} u_i^{\alpha_i-1} \left(1 - \sum_{i=1}^{r} x_i\right)^{\alpha_{r+1}-1} \prod_{i=1}^{r} dx_i = \frac{\prod_{i=1}^{r+1} \Gamma(\alpha_i)}{\Gamma(\sum_{i=1}^{r+1} \alpha_i)}.
\]  

Furthermore, by setting \( r = n - 1, \alpha_i = a_i + j_i, i = 1, \ldots, n \) where \( a_1, \ldots, a_n \) are positive real numbers and \( j_1, \ldots, j_n \) are non-negative integers in (6), we get

\[
\int \cdots \int \prod_{i=1}^{n} x_i^{a_i+j_i-1} \left(1 - \sum_{i=1}^{n-1} x_i\right)^{a_n+j_n-1} \prod_{i=1}^{n-1} dx_i = \frac{\prod_{i=1}^{n} \Gamma(a_i) \cdot (a_1)_{j_1} \cdots (a_n)_{j_n}}{\Gamma(\sum_{i=1}^{n} a_i) \cdot (\sum_{i=1}^{n} a_i)_{j_1+\cdots+j_n}}.
\]

3. The Dirichlet-Lauricella Type D Distribution

The Dirichlet-Lauricella type D distribution is defined as follows.
The random variables $U_1, \ldots, U_n$ are said to have a Dirichlet-Lauricella type D distribution with parameters $a_1, \ldots, a_n; c; d; \theta_1, \ldots, \theta_n$, denoted by $(U_1, \ldots, U_n) \sim \text{DLD}(a_1, \ldots, a_n; c; d; \theta_1, \ldots, \theta_n)$, if their joint p.d.f. is given by

$$K_D \prod_{i=1}^{n} u_i^{a_i-1} \frac{(1 - \sum_{i=1}^{n} u_i)^{c - \sum_{i=1}^{n} a_i - 1}}{(1 - \sum_{i=1}^{n} \theta_i u_i)^d}, \quad u_i > 0, \quad i = 1, \ldots, n, \quad \sum_{i=1}^{n} u_i < 1,$$

where $a_1 > 0, \ldots, a_n > 0$, $c - a_1 - \cdots - a_n > 0$, $-1 < \theta_1 < 1, \ldots, -1 < \theta_n < 1$ and $-\infty < d < \infty$. The normalizing constant $K_D$ in (8) is given by

$$K_D^{-1} = \int_{u_1 > 0, \ldots, u_n > 0} \prod_{i=1}^{n} u_i^{a_i-1} \left(1 - \sum_{i=1}^{n} u_i\right)^{c - \sum_{i=1}^{n} a_i - 1} \frac{\prod_{i=1}^{n} \Gamma(a_i) \Gamma(c - \sum_{i=1}^{n} a_i)}{\Gamma(c)} \frac{u_i^{a_i-1}(1 - u_i)^c}{(1 - \theta_i u_i)^d} \prod_{i=1}^{n} du_i$$

where the last line has been obtained by using (3).

From (8), it is clear that for $d = 0$, the Dirichlet-Lauricella type D distribution reduces a Dirichlet type 1 distribution with parameters $a_1, \ldots, a_n$ and $c - \sum_{i=1}^{n} a_i$.

For $n = 1$, the p.d.f. in (8) simplifies to a generalized beta type 1 p.d.f. given by

$$\frac{\Gamma(c)}{\Gamma(a_1) \Gamma(c - a_1)}_2 F_1(a_1, d; c; \theta_1) \frac{u_1^{a_1-1}(1 - u_1)^c}{(1 - \theta_1 u_1)^d}, \quad 0 < u_1 < 1,$$

where $c > a_1 > 0$, $-1 < \theta_1 < 1$ and $_2 F_1$ is the Gauss hypergeometric function.

The generalized beta type 1 distribution defined by the above p.d.f. has been studied by Nagar and Rada-Mora [6], Nagar and Bedoya-Valencia [7], Libby and Novic [4], Pham-Gia and Duong [9].

Further, for $n = 2$, the p.d.f. in (8) slides to a generalized bivariate beta type 1 p.d.f. defined by (Nadarajah and Kotz [5]),

$$\frac{\Gamma(c)}{\Gamma(a_1) \Gamma(a_2) \Gamma(c - a_1 - a_2)}_2 F_1(d, a_1, a_2; c; \theta_1, \theta_2) \frac{u_1^{a_1-1} u_2^{a_2-1}(1 - u_1 - u_2)^c}{(1 - \theta_1 u_1 - \theta_2 u_2)^d}, \quad u_1 > 0, \quad u_2 > 0, \quad u_1 + u_2 < 1,$$

where $a_1 > 0$, $a_2 > 0$, $c > a_1 + a_2$, $-1 < \theta_1 < 1$, $-1 < \theta_2 < 1$ and $_2 F_1$ is the first hypergeometric function of Appell.
Consider the transformation \( Z_i = (1 - \sum_{j=1}^{s} U_j)^{-1} U_i, i = s + 1, \ldots, n. \) Then, \( u_i = (1 - \sum_{j=1}^{s} u_j) z_i, i = s + 1, \ldots, n \) with the Jacobian \( J(u_{s+1}, \ldots, u_n \rightarrow z_{s+1}, \ldots, z_n) = (1 - \sum_{j=1}^{s} u_j)^{n-s}. \) Substituting appropriately in (8), the joint density of \( U_1, \ldots, U_s, Z_{s+1}, \ldots, Z_n \) is given by

\[
K_D \prod_{i=1}^{s} u_i^{a_i-1} \frac{(1 - \sum_{i=1}^{s} u_i)^{c - \sum_{i=1}^{s} a_i - 1} \prod_{i=s+1}^{n} z_i^{a_i-1} (1 - \sum_{i=s+1}^{n} z_i)^{c - \sum_{i=1}^{n} a_i - 1}}{[1 - \sum_{i=1}^{s} \theta_i u_i - (1 - \sum_{j=1}^{s} u_j) \sum_{i=s+1}^{n} \theta_i z_i]^d},
\]

where \( u_i > 0, i = 1, \ldots, s, \sum_{i=1}^{s} u_i < 1, \) \( z_i > 0, i = s + 1, \ldots, n, \) and \( \sum_{i=s+1}^{n} z_i < 1. \)

Now, we find the marginal p.d.f. of \( U_1, \ldots, U_s \) by integrating out \( z_{s+1}, \ldots, z_n \) from the joint density of \( U_1, \ldots, U_s, Z_{s+1}, \ldots, Z_n \) as

\[
K_D \prod_{i=1}^{s} u_i^{a_i-1} \frac{(1 - \sum_{i=1}^{s} u_i)^{c - \sum_{i=1}^{s} a_i - 1}}{(1 - \sum_{i=1}^{s} \theta_i u_i)^d} \times \int \cdots \int z_{s+1} > 0, \ldots, z_n > 0, \sum_{i=s+1}^{n} z_i < 1 \prod_{i=s+1}^{n} dz_i.
\]

Now, using the integral representation of \( F_D^{(n)} \), we derive the marginal p.d.f. of \( U_1, \ldots, U_s \) as

\[
K_{D1} \prod_{i=1}^{s} u_i^{a_i-1} \frac{(1 - \sum_{i=1}^{s} u_i)^{c - \sum_{i=1}^{s} a_i - 1}}{(1 - \sum_{i=1}^{s} \theta_i u_i)^d} \times F_{D}^{(n-s)}(d, a_{s+1}, \ldots, a_n; c - \sum_{i=1}^{s} a_i, \frac{\theta_{s+1} (1 - \sum_{i=1}^{s} u_i)}{1 - \sum_{i=1}^{s} \theta_i u_i}, \ldots, \frac{\theta_n (1 - \sum_{i=1}^{s} u_i)}{1 - \sum_{i=1}^{s} \theta_i u_i}),
\]

where \( u_i > 0, i = 1, \ldots, s, \sum_{i=1}^{s} u_i < 1 \) and

\[
K_{D1}^{-1} = \prod_{i=1}^{s} \frac{\Gamma(a_i) \Gamma(c - \sum_{i=1}^{s} a_i)}{\Gamma(c)} F_D^{(n)}(d, a_1, \ldots, a_n; c; \theta_1, \ldots, \theta_n).
\]

It is interesting to observe that the marginal density of \( U_1, \ldots, U_s \) does not belong to the Dirichlet-Lauricella type D family of distributions and differs by an additional factor containing the Lauricella hypergeometric function \( F_D. \)
From (11), it is straightforward to show that the marginal p.d.f. of $U_s$ is

$$K_D D_2 \frac{u_s^{a_s-1} (1 - u_s)^{c-a_s-1}}{(1 - \theta_s u_s)^{d_s}} F_D^{(n-1)}(d, a_1, \ldots, a_s-1, a_{s+1}, \ldots, a_n; c - a_s;)
\frac{\theta_1 (1 - u_s)}{1 - \theta_s u_s}, \ldots, \frac{\theta_{s-1} (1 - u_s)}{1 - \theta_s u_s}, \frac{\theta_{s+1} (1 - u_s)}{1 - \theta_s u_s}, \ldots, \frac{\theta_n (1 - u_s)}{1 - \theta_s u_s},$$

where

$$K_D^{-1} = \frac{\Gamma(a_s) \Gamma(c - a_s)}{\Gamma(c)} F_D^{(n)}(d, a_1, \ldots, a_n; c; \theta_1, \ldots, \theta_n). \quad (13)$$

The marginal p.d.f. of $Z_{s+1}, \ldots, Z_n$ is given by

$$K_D \prod_{i=s+1}^{n} z_i^{a_i-1} \frac{(1 - \sum_{i=s+1}^{n} z_i)^{c-\sum_{i=1}^{n} a_i-1}}{(1 - \sum_{i=s+1}^{n} \theta_i z_i)^d} \times \int \ldots \int \frac{\prod_{i=1}^{s} u_i^{a_i-1} (1 - \sum_{i=1}^{s} u_i)^{c-\sum_{i=1}^{s} a_i-1}}{[1 - \sum_{j=1}^{s} (\theta_j - \sum_{i=s+1}^{n} \theta_i z_i) u_j / (1 - \sum_{i=s+1}^{n} \theta_i z_i)]} \prod_{i=1}^{s} du_i.$$ 

Now, evaluating the above integral by using (3), we get

$$K_D D_3 \frac{\prod_{i=s+1}^{n} z_i^{a_i-1} (1 - \sum_{i=s+1}^{n} z_i)^{c-\sum_{i=1}^{n} a_i-1}}{(1 - \sum_{i=s+1}^{n} \theta_i z_i)^d} \times F_D^{(s)}(d, a_1, \ldots, a_s; c; \frac{\theta_1 - \sum_{i=s+1}^{n} \theta_i z_i}{1 - \sum_{i=s+1}^{n} \theta_i z_i}, \ldots, \frac{\theta_s - \sum_{i=s+1}^{n} \theta_i z_i}{1 - \sum_{i=s+1}^{n} \theta_i z_i}),$$

where

$$K_D^{-1} = \frac{\prod_{i=s+1}^{n} \Gamma(a_i) \Gamma(c - \sum_{i=1}^{n} a_i)}{\Gamma(c - \sum_{i=1}^{n} a_i)} F_D^{(n)}(d, a_1, \ldots, a_n; c; \theta_1, \ldots, \theta_n). \quad (14)$$

It is well known that if $(U_1, \ldots, U_n) \sim D1(a_1, \ldots, a_n; c - \sum_{i=1}^{n} a_i)$, then

$$\left(\frac{U_1}{\sum_{i=1}^{n} U_i}, \ldots, \frac{U_{n-1}}{\sum_{i=1}^{n} U_i}\right) \sim D1(a_1, \ldots; a_n)$$

and the sum $\sum_{i=1}^{n} U_i$ follows a beta type 1 distribution with parameters $\sum_{i=1}^{n} a_i$ and $c - \sum_{i=1}^{n} a_i$. In the next theorem, we derive similar result for the Dirichlet-Lauricells type D distribution.
Theorem 3.1. Let \( (U_1, \ldots, U_n) \sim \text{DLD}(a_1, \ldots, a_n; c; d; \theta_1, \ldots, \theta_n) \) and define \( U = \sum_{i=1}^{n} U_i \) and \( X_i = U_i / U, \ i = 1, \ldots, n-1 \). Then, the joint p.d.f. of \( X_1, \ldots, X_{n-1} \) is given as

\[
K_D \left[ \frac{\Gamma(\sum_{i=1}^{n} a_i)\Gamma(c - \sum_{i=1}^{n} a_i)}{\Gamma(c)} \prod_{i=1}^{n-1} x_i^{a_i-1} \left( 1 - \sum_{i=1}^{n-1} x_i \right)^{a_n-1} \right]
\times 2F_1 \left( \sum_{i=1}^{n} a_i, d; c; \sum_{i=1}^{n-1} \theta_i x_i + \theta_n \left( 1 - \sum_{i=1}^{n-1} x_i \right) \right),
\]

where \( x_i > 0, i = 1, \ldots, n-1, \sum_{i=1}^{n-1} x_i < 1 \). Further, the p.d.f. of \( U \) is derived as

\[
K_D \left[ \frac{\prod_{i=1}^{n} \Gamma(a_i)}{\Gamma(\sum_{i=1}^{n} a_i)} \right] \frac{\sum_{i=1}^{n} a_i-1(1-u)^{c-\sum_{i=1}^{n} a_i-1}}{[1-\sum_{i=1}^{n-1} \theta_i x_i u - \theta_n u(1-\sum_{i=1}^{n-1} x_i)]^d},
\]

where \( x_i > 0, i = 1, \ldots, n-1, \sum_{i=1}^{n-1} x_i < 1 \) and \( 0 < u < 1 \). Now, integrating \( u \) in the above expression by using the integral representation of \( 2F_1 \), we get the desired result. Further, writing

\[
\left[ 1 - \sum_{i=1}^{n-1} \theta_i x_i u - \theta_n u \left( 1 - \sum_{i=1}^{n-1} x_i \right) \right]^{-d}
\]

\[
= \sum_{j_1, \ldots, j_n=0}^{\infty} (d)_{j_1, \ldots, j_n} (\theta_1 x_1 u)^{j_1} \cdots (\theta_{n-1} x_{n-1} u)^{j_{n-1}} [\theta_n (1 - \sum_{i=1}^{n-1} x_i) u]^{j_n} \frac{1}{j_1! \cdots j_n!}
\]

the joint p.d.f. of \( (X_1, \ldots, X_{n-1}) \) and \( U \) is rewritten as

\[
K_D u^{\sum_{i=1}^{n} a_i-1}(1-u)^{c-\sum_{i=1}^{n} a_i-1} \sum_{j_1, \ldots, j_n=0}^{\infty} (d)_{j_1, \ldots, j_n} \theta_1^{j_1} \cdots \theta_{n-1}^{j_{n-1}} \theta_n^{j_n} \frac{1}{j_1! \cdots j_n!}
\]
\[
\times \frac{(\theta_1 u)^{j_1} \cdots (\theta_{n-1} u)^{j_{n-1}} (\theta_n u)^{j_n}}{j_1! \cdots j_n!} \prod_{i=1}^{n-1} x_i^{a_i+j_i-1} \left(1 - \sum_{i=1}^{n-1} x_i\right)^{a_n+j_n-1}
\]

Now, integrating the above expression with respect to \(x_1, \ldots, x_{n-1}\) by using (7) and summing the resulting series by applying (2), we get the desired result. \(\square\)

By definition, the product moments are obtained as

\[
E\left[\prod_{i=1}^{n} u_i^{r_i}\right] = \int \cdots \int \prod_{i=1}^{n} u_i^{a_i+r_i-1} \left(1 - \sum_{i=1}^{n} u_i\right)^{c - \sum_{i=1}^{n} a_i - 1} \prod_{i=1}^{n} du_i
\]

\[
= \frac{\Gamma(c) \prod_{i=1}^{n} \Gamma(a_i + r_i)}{\Gamma(c + r) \prod_{i=1}^{n} \Gamma(a_i)}
\times F_D^{(n)}(d, a_1 + r_1, \ldots, a_n + r_n; c + r; \theta_1, \ldots, \theta_n)
\times F_D^{(n)}(d, a_1, \ldots, a_n; c; \theta_1, \ldots, \theta_n)
\]

where \(r = \sum_{i=1}^{n} r_i\), \(\text{Re}(a_i + r_i) > 0, i = 1, \ldots, n\) and \(\text{Re}(c + r) > 0\). Further

\[
E\left[(1 - \sum_{i=1}^{n} u_i)^{h}\right] = \int \cdots \int \prod_{i=1}^{n} u_i^{a_i-1} \left(1 - \sum_{i=1}^{n} u_i\right)^{c + h - \sum_{i=1}^{n} a_i - 1} \prod_{i=1}^{n} du_i
\]

\[
= \frac{\Gamma(c) \Gamma(c + h - \sum_{i=1}^{n} a_i)}{\Gamma(c + h) \Gamma(c - \sum_{i=1}^{n} a_i)}
\times F_D^{(n)}(d, a_1, \ldots, a_n; c + h; \theta_1, \ldots, \theta_n)
\times F_D^{(n)}(d, a_1, \ldots, a_n; c; \theta_1, \ldots, \theta_n)
\]

**Acknowledgments**

The research work of DKN and AG was supported by the Sistema Universitario de Investigación, Universidad de Antioquia under the project No. IN10182CE.

**References**


