HNSJ
Humanities & Natural Sciences Journal
ISSN: (e) 2709-0833
www.hnjournal.net
Peer-Reviewed Journal
Indexed at: NSP - SJIF
Impact Factor 2022 = 4.91

RESEARCH TITLE

Bivariate Dirichlet Distribution: Analysis of Statistical Properties and Parameter Estimation Methods with a Simulation Study

Sara.A. El_Warrad[1], Mohamed Amraja Mohamed[2], and Mohammed A. Asselhab[3]

[1] Statistical Department, Faculty of Arts and Sciences, Benghazi University, Libya.
sara.alwarrad@uob.edu.ly
[2] Statistical Department-faculty of science- Sebha University, Libya.
[3] Mathematical Department-faculty of Arts and sciences- Sebha University, Libya.
HNSJ, 2025, 6(6); https://doi.org/10.53796/hnsj66/12

Received at 07/05/2025 **Accepted at 15/05/2025** **Published at 01/06/2025**

Abstract

In this paper, we study the bivariate Dirichlet distribution and discusses some important statistical properties such as product moments, covariance, and the correlation coefficient. It also introduces a simple method for generating random pairs (X, Y) based on the marginal distribution of X and the conditional distribution of Y|X=x. Furthermore, Estimation of the parameters for the bivariate Dirichlet distribution are derived using method of moments (MME). Maximum likelihood estimator (MLE) are also presented. Finally, it includes a simulation to evaluate the efficiency of the estimators based on bias and mean squared error.

Key Words: Bivariate Dirichlet distribution, Maximum likelihood estimator, moments estimators.

عنوان البحث

التوزيع الثنائي المتغير ديريشليت: تحليل الخصائص الإحصائية وطرق تقدير المعلمات مع دراسة محاكاة

المستخلص

يتناول هذا البحث التوزيع الثنائي المتغير ديريشليت(Bivariate Dirichlet) ، حيث يتم استعراض بعض الخصائص الإحصائية المهمة مثل لحظات الجداء، والتباين المشترك، ومعامل الارتباط. كما يقدم البحث طريقة مبسطة لتوليد أزواج عشوائية (X, Y) بالاعتماد على التوزيع الهامشي لـ X والتوزيع الشرطي لـ Y عند X = x. بالإضافة إلى ذلك، تم اشتقاق تقديرات معلمات التوزيع باستخدام طريقة العزوم(MME) ، كما تم تقديم طريقة التقدير بالاحتمالية العظمى (MLE). ويختتم البحث بمحاكاة لقياس كفاءة التقديرات من حيث الانحياز ومتوسط مربع الخطأ.(MSE)

الكلمات المفتاحية: التوزيع الثنائي المتغير ديريشليت، مقدّر الاحتمالية العظمى، مقدّرات العزوم.

1. Introduction

The bivariate beta distribution is one of the basic distributions in statistics, as it attracts useful applications in several areas; for example, in the modeling of the proportions of substances in a mixture, brand shares, i.e the proportions of brands of some consumer product that are bought by customers (Chatfiled [1975]), proportions of the electorate voting for the candidate in a two candidate election (Hoyer and Mayer [1976]) and the dependence between two soil strength parameters (A_Grivas and Asaoka

[1982]).

Bivariate beta distributions have also been used extensively as a prior in Bayesian statistics (see, for example, Apostolakis and Moieni [1987]). The Dirichlet distribution is multivariate generalization of the beta distribution, hence its alternative name of multivariate beta distribution . Several applications of the Dirichlet distribution are discussed by Wilks [1962], Goodhardt et al [1984], Lange [1995], Bouguila et al [2004], Null [2009] and Wang et al [2011]. Estimation of parameters of bivariate Dirichlet distribution by maximum likelihood is discussed by Nadarajah and Kotz [2007] and estimation parameters of the Dirichlet distribution based on entropy by Sahin et al [2023].

This paper commences in Section 2 with an exposition of the marginal distributions associated with the bivariate Dirichlet distribution. This section further provides a thorough examination of their statistical properties, specifically addressing moments, product moments, covariance, and the correlation coefficient. Following this, Section 3 introduces the conditional distributions and explores their moment characteristics. A practical and accessible approach for generating random varieties from the distribution is outlined in Section 4. Section 5 is dedicated to the derivation of estimators for the distribution's parameters utilizing the method of moments. In addition, the maximum likelihood estimators (MLE) are presented and discussed. To conclude, Section 6 presents a simulation study conducted to assess the efficiency of the proposed estimators. As an alternative, this final section may present a numerical illustration to validate the findings presented in Sections 4 and 5.

The present work focuses on the bivariate Dirichlet distribution, parameterized by positive values a,b,c, and d, which is defined by the subsequent probability density function (pdf):

$$f_{X,Y}(x,y) = K\ x^{a-1} y^{b-1} (1-x)^{c-b-d} (1-x-y)^{d-1} \qquad (1)$$

$for, x \geq 0, y \geq 0, x+y<1, a>0, b>0, c>0 \ and \ d>0$

and

$$K = \frac{1}{\mathrm{B}(a,c)\mathrm{B}(b,d)} \qquad (2)$$

The distribution in (1) is the bivariate form of the Connor and Mosimann’s generalized Dirichlet distribution (see Connor and Mosimann [1969]). It has several applications in many areas, including Bayesian statistics, contingency tables, correspondence analysis, environmental sciences, forensic sciences, geochemistry, image analysis and statistical decision theory (see Gupta and Nadarajah [2004] for illustrations of some of these application areas).

We will define some properties estimator. these properties will help us in deciding whether one estimator is better than anther.

Definition 1: The point estimator $\hat{\theta}$ is an unbiased estimator for the parameter θ if

$E(\hat{\theta}) = \theta$

If the estimator is not unbiased, then the difference

$E(\hat{\theta}) - \theta,$

is called the bias of estimator $\hat{\theta}$.

Definition 2: The mean squared error (MSE) of an estimator $\hat{\theta}$ of the parameter θ is defined as

$$MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2$$

The mean squared error (MSE) can be rewritten as follows :

$$MSE(\hat{\theta}) = E(\hat{\theta} - E(\theta))^2 + (\theta - E(\hat{\theta}))^2$$

$$= Var(\hat{\theta}) + (bias)^2$$

The calculations throughout this paper involve some special function**,** including the beta type I

$$f_X(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1} \qquad 0 < x < 1 \qquad (3)$$

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1}\ (1-t)^{\beta-1}. \qquad (4)$$

The rth moment of X is

$$E(X^r) = \frac{B(\alpha + r, \beta)}{B(\alpha, \beta)}, \qquad (5)$$

The Gauss hypergeometric function

$$_2F_1(\delta, \lambda; \lambda + \mu; x) = \frac{1}{B(\lambda, \mu)} \int_0^1 u^{\lambda-1}(1-u)^{\mu-1}(1-xu)^{-\delta}\, du \qquad (6)$$

where $\text{Re}\ \lambda > 0$, $\text{Re}\ \mu > 0$, which is given in a series form by

$$_2F_1(\delta, \lambda; \mu; x) = \sum_{j=0}^{\infty} \frac{(\delta)_j (\lambda)_j}{(\mu)_j} \frac{x^j}{j!}, \qquad (7)$$

where $|x| < 1$, and $(f)_k = f(f+1)\ldots\ldots(f+k-1)$ denotes the ascending factorial.

The hypergeometric function type I

$$f_X(x) = \frac{\Gamma(\gamma + v - \alpha)\Gamma(\gamma + v - \beta)}{\Gamma(\gamma)\Gamma(v)\Gamma(\gamma + v - \alpha - \beta)} x^{v-1}(1-x)^{\gamma-1}\, {}_2F_1(\alpha, \beta; \gamma; 1-x), \qquad (8)$$

where $0 < x < 1$, $\text{Re}(\gamma + v - \alpha - \beta) > 0, \text{Re}\, v > 0$ *and* $\gamma > 0$. (see Gupta and Nagar [2000], p298).

The rth moment of X is

$$E(X^r) = \frac{\Gamma(\gamma + v - \alpha)\Gamma(\gamma + v - \beta)}{\Gamma(v)\Gamma(\gamma + v - \alpha - \beta)} \frac{\Gamma(v + r)\Gamma(\gamma + v + r - \alpha - \beta)}{\Gamma(\gamma + v + r - \alpha)\Gamma(\gamma + v + r - \beta)}, \qquad (9)$$

where $\text{Re}(\gamma + v + r - \alpha - \beta) > 0$.(for proof see Nagar.D and Alvarez,A.J [2005])

The properties of the above special function can be found in (Gradshteyn and Ryzhik [1980] , p 284,984,286,1040 and 298).

2. Marginal Distributions

This section focuses on the derivation and analysis of the marginal distributions of the components X and Y of the bivariate Dirichlet distribution given by its probability density function (pdf) in Equation (1). We further explore key statistical properties of these marginal distributions, specifically their moments, product moments, covariance, and correlation coefficient.

2.1. The Marginal Distribution of X

Integrating Equation (1), with respect to y, we obtain the marginal pdf of X given by

$$f_X(x) = \frac{1}{\mathrm{B}(a,c)} x^{a-1}(1-x)^{c-1} \qquad 0 < x < 1$$

Note that the marginal pdf of X is $\mathrm{B}_1(a,c)$ see(3).

From Equation (5) , we find that

$$\mathrm{E}(X) = \frac{a}{(a+c)} \tag{10}$$

and

$$Var(X) = \frac{ac}{(a+c+1)(a+c)^2} \tag{11}$$

2.2. The Marginal Distribution of Y

Integrating Equation (1), with respect to x, we obtain the marginal pdf of Y given by

$$f_Y(y) = Ay^{b-1}(1-y)^{(d+a)-1}\,{}_2F_1(a,b+d-c;d+a;1-y) \qquad 0 < y < 1 \tag{12}$$

where $A = \frac{\Gamma(a+c)\Gamma(b+d)}{\Gamma(b)\Gamma(c)\Gamma(d+a)}$ and ${}_2F_1(a,b+d-c;d+a;1-y)$ is defined in (6)

Note that the marginal pdf of Y is $H_1(b,a,b+d-c,a+d)$ see (8).

From Equation (9) , we have

$$\mathrm{E}(Y) = \frac{bc}{(b+d)(a+c)} \tag{13}$$

$$Var(Y) = \frac{bc\left(ba(b+d+1)+cd(a+c+1)+da\right)}{(a+c)^2(b+d+1)^2(a+c+1)} \tag{14}$$

2.3. Product Moments

Theorem1: If X and Y are jointly distributed random variables with the joint pdf in Equation (1), then

$$\mathrm{E}\left(X^nY^m\right) = K\mathrm{B}(a+n,c+m)\mathrm{B}(b+m,d) \tag{15}$$

for $a+n>0, b+m>0, c+m>0$ *and* $K = \frac{1}{\mathrm{B}(a,c)\mathrm{B}(b,d)}$

Proof: Knowing that

$$E\left(X^nY^m\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} X^nY^m f_{X,Y}(x,y)dydx \qquad (16)$$

and substituting with Equation (1) into Equation (16), we get

$$E\left(X^nY^m\right) = K\int_0^1\int_0^{1-x} x^{a+n-1}y^{b+m-1}(1-x)^{c-b-d}\ (1-x-y)^{d-1}dydx$$

$$= K\int_0^1 x^{a+n-1}(1-x)^{c-b-d}\left(\int_0^{1-x} y^{b+m-1}\left[(1-x)\left(1-\frac{y}{1-x}\right)^{d-1}\right]dy\right)dx$$

Using the transformation $u = \frac{y}{1-x}$, we get

$$E\left(X^nY^m\right) = K\int_0^1 x^{a+n-1}(1-x)^{c-b-1}\left(\int_0^1 [u(1-x)]^{b+m-1}(1-u)^{d-1}(1-x)du\right)dx$$

$$= K\int_0^1 x^{a+n-1}(1-x)^{c+m-1}dx\int_0^1 u^{b+m-1}(1-u)^{d-1}du$$

Using Equation (4) in the above equation , we obtain Equation (15).

This completes the proof of the Theorem.

Theorem2: If X and Y are jointly distributed random variables with the joint pdf in Equation (1), then the correlation coefficient of X and Y is given by

$$\rho = -\sqrt{\frac{ab(b+d+1)}{\{cd(a+c+1)+(ba(b+d+1)+da\}}} \qquad (17)$$

Proof: First to compute the covariance of X and Y given by

$$Cov(X,Y) = E(XY) - E(X)E(Y)$$

From Equations (15), (10) and (13), we get

$$Cov(X,Y) = \frac{abc}{(a+c+1)(a+c)(b+d)} - \frac{a}{(a+c)}\frac{bc}{(a+c)(b+d)}$$

$$= -\frac{abc}{(a+c+1)(a+c)^2(b+d)} \qquad (18)$$

Now, we determine the correlation coefficient of X and Y

$$\rho = \frac{\operatorname{cov}(x,y)}{\sqrt{\operatorname{var}(x)\operatorname{var}(y)}}$$

Using Equations (18), (11) and (14), we obtain Equation (17).

3. The Conditional Density Functions:

In this section, we study the conditional distributions of X and Y are jointly distributed with the pdf (1).we also derive the conditional moments.

Theorem 3: If X and Y are the jointly distributed random variables with the joint pdf (1), then the conditional pdf of Y given X=x is given by

$$f_{Y|x}(y\,|\,x) = \frac{(1-x)}{\mathrm{B}(b,d)}\left(\frac{y}{1-x}\right)^{b-1}\left(1-\frac{y}{1-x}\right)^{d-1}, \qquad 0<y<1-x, 0<x<1. \tag{19}$$

Equivalently,

$$f\left(\frac{y}{1-x}\,|_{X=x}\right) = \frac{1}{\mathrm{B}(b,d)}\left(\frac{y}{1-x}\right)^{b-1}\left(1-\frac{y}{1-x}\right)^{d-1}, \qquad 0<\frac{y}{1-x}<1. \tag{20}$$

Proof: The conditional density function $f_{Y|x}(y\,|\,x)$ is

$$f_{Y|x}(y\,|\,x) = \frac{f_{X,Y}(x,y)}{f_X(x)}, \qquad 0<y<1-x,\ 0<,x<1.$$

Using (1) and (3), we have

$$f_{Y|x}(y|\,x) = \frac{1}{\mathrm{B}(a,c)\mathrm{B}(b,d)}x^{a-1}y^{b-1}(1-x)^{c-b-d}(1-x-y)^{d-1}$$

$$\times\left(\frac{1}{\mathrm{B}(a,c)}x^{a-1}(1-x)^{c-1}\right)^{-1}$$

Which can be rewritten as

$$f_{Y|x}(y|\,x) = \frac{1}{\mathrm{B}(b,d)}\frac{y^{b-1}(1-x-y)^{d-1}}{(1-x)^{b+d-1}}, \qquad 0<y<1-x,\ 0<,x<1.$$

leading to the result given in (19).

Obviously, the result in (20) follows, directly.

Theorem 4: If X and Y are the jointly distributed random variables with the joint pdf (1), then the conditional pdf of X given Y =y is given by

$$f_{X|y}(x\,|\,y) = \frac{(1-y)}{\mathrm{B}(a,d)\,{}_2F_1(a,b+d-c;a+d;1-y)}\left(\frac{x}{1-y}\right)^{a-1}\left(1-\frac{x}{1-y}\right)^{d-1}$$

$$\times\left(1-\frac{(1-y)x}{1-y}\right)^{-(b+d-c)} \tag{21}$$

$0<x<1-y, 0<y<1.$

Equivalently,

$$f\left(\frac{x}{1-y}\,|_{Y=y}\right) = \frac{1}{\mathrm{B}(a,d)\,{}_2F_1(a,b+d-c;a+d;1-y)}\left(\frac{x}{1-y}\right)^{a-1}\left(1-\frac{x}{1-y}\right)^{d-1}$$

$$\times\left(1-\frac{(1-y)x}{1-y}\right)^{-(b+d-c)} \tag{22}$$

where $0<\frac{x}{1-y}<1.$

Proof: The conditional density function of $f_{X|y}(x\,|\,y)$ is obtained by diving $f_{X,Y}(x,y)$ in Equation (1) by $f_Y(y)$ in Equation (12). Thus

$$f_{X|y}(x|y)=\frac{1}{B(a,c)B(b,d)}x^{a-1}y^{b-1}(1-x)^{c-b-d}(1-x-y)^{d-1}$$

$$\times\left(\frac{\Gamma(a+c)\Gamma(b+d)}{\Gamma(b)\Gamma(c)\Gamma(a+d)}y^{b-1}(1-y)^{a+d-1}{}_1F_2(a,b+d-c;a+d;1-y)\right)^{-1}$$

$$=\frac{\Gamma(a+d)}{\Gamma(a)\Gamma(d)}\frac{x^{a-1}(1-x)^{c-b-d}(1-x-y)^{d-1}}{(1-y)^{a+d-1}{}_2F_1(a,b+d-c;a+d;1-y)}$$

$0<x<1-y, 0<y<1.$

which leads to the result in Equation (21).

The conditional distribution of X/(1-Y) given Y=y in Equation (22) then follows.

Note that the result (20) belongs to the standard beta family with parameters b and d , and the result (22) belongs to(Libby and Novick's [1982]) generalized beta family with parameters a , d, $1-y$ and $b+d-c$.

Theorem 5: If X and Y are the jointly distributed random variables with the joint pdf (1), then

$$E(Y\mid x)=\frac{(1-x)b}{b+d} \qquad (23)$$

and

$$Var(Y\mid x)=\frac{(1-x)^2bd}{(b+d+1)(b+d)^2} \qquad (24)$$

Proof: Since the distribution of $\frac{Y}{1-x}$ given $X=x$ is $B_1(b,d)$, then from Equation (5) ,we have

$$E\left(\frac{Y}{1-x}|_{X=x}\right)=\frac{b}{b+d} \qquad (25)$$

and

$$Var\left(\frac{Y}{1-x}|_{X=x}\right)=\frac{bd}{(b+d+1)(b+d)^2} \qquad (26)$$

Equations (25) and (26) can be simplified to obtain Equations (23) and (24).

From Equation (23), we see that with non-homoscedastic variance.

Theorem 6: If X and Y are the jointly distributed random variables with the joint pdf (1), then

$$E(X\mid y)=\frac{a(1-y)}{(a+d)}\frac{{}_2F_1(a+1,b+d-c;a+d+1;1-y)}{{}_2F_1(a,b+d-c;a+d;1-y)} \qquad (27)$$

and

$$Var(X\mid y)=\frac{a(1-y)^2}{(a+d)}\frac{1}{{}_2F_1(a,b+d-c;a+d;1-y)}\left\{\left(\frac{(a+1)}{(a+d+1)^2}F_1(a+2,b+d-c;a+d+2;1-y)\right)\right.$$

$$-\frac{a}{(a+d)}\left(\frac{\left({}_2F_1(a+1,b+d-c;a+d+1;1-y)\right)^2}{{}_2F_1(a,b+d-c;a+d;1-y)}\right)\Bigg\} \qquad (28)$$

Proof: The first moment of $\left(\frac{X}{1-y}|_{Y=y}\right)$ is

$$E\left(\frac{X}{1-y}|_{Y=y}\right)=\int_{-\infty}^{\infty}\left(\frac{x}{1-y}\right)\ f\left(\frac{x}{1-y}|_{Y=y}\right)dx$$

$$=\mathrm{M}\int_{0}^{1-y}\left(\frac{x}{1-y}\right)^{(a+1)-1}\left(1-\frac{x}{1-y}\right)^{d-1}\left(1-(1-y)\frac{x}{1-y}\right)^{-(b+d-c)}dx,$$

where $\mathrm{M}=\frac{1}{\mathrm{B}(a,d)\,{}_2F_1(a,b+d-c;a+d;1-y)}$

Solving the above integral by using (6),we get

$$\mathrm{E}\left(\frac{X}{1-y}|_{Y=y}\right)=\frac{\mathrm{B}(a+1,d)}{\mathrm{B}(a,d)}\frac{{}_2F_1(a+1,b+d-c;a+d+1;1-y)}{{}_2F_1(a,b+d-c;a+d;1-y)}$$

$$=\frac{a}{(a+d)}\frac{{}_2F_1(a+1,b+d-c;a+d+1;1-y)}{{}_2F_1(a,b+d-c;a+d;1-y)} \qquad (29)$$

Next, we determine the second moment of $\left(\frac{X}{1-y}|_{Y=y}\right)$

$$\mathrm{E}\left(\left(\frac{X}{1-y}\right)^2|_{Y=y}\right)=\mathrm{M}\int_{0}^{1-y}\left(\frac{x}{1-y}\right)^{(a+2)-1}\left(1-\frac{x}{1-y}\right)^{d-1}\left(1-(1-y)\frac{x}{1-y}\right)^{-(b+d-c)}dx.$$

Using Equation (6),we get

$$\mathrm{E}\left(\left(\frac{X}{1-y}\right)^2|_{Y=y}\right)=\frac{\mathrm{B}(a+2,d)}{\mathrm{B}(a,d)}\frac{{}_2F_1(a+2,b+d-c;a+d+2;1-y)}{{}_2F_1(a,b+d-c;a+d;1-y)}$$

$$=\frac{a(a+1)}{(a+d)(a+d+1)}\frac{{}_2F_1(a+2,b+d-c;a+d+2;1-y)}{{}_2F_1(a,b+d-c;a+d;1-y)}$$

Then the conditional variance is

$$Var\left(\frac{X}{1-y}|_{Y=y}\right)=\frac{a(a+1)}{(a+d)(a+d+1)}\frac{{}_2F_1(a+2,b+d-c;a+d+2;1-y)}{{}_2F_1(a,b+d-c;a+d;1-y)}$$

$$-\frac{a^2}{(a+d)^2}\left(\frac{{}_2F_1(a+1,b+d-c;a+d+1;1-y)}{{}_2F_1(a,b+d-c;a+d;1-y)}\right)^2$$

$$=\frac{a}{(a+d)}\frac{1}{{}_2F_1(a,b+d-c;a+d;1-y)}\left\{\left(\frac{(a+1)}{(a+d+1)}{}_2F_1(a+2,b+d-c;a+d+2;1-y)\right)\right.$$

$$\left.-\frac{a}{(a+d)}\left(\frac{\left({}_2F_1(a+1,b+d-c;a+d+1;1-y)\right)^2}{{}_2F_1(a,b+d-c;a+d;1-y)}\right)\right\} \qquad (30)$$

Equations (29) and (30) can be simplified to obtain Equations (27) and (28).

4. Algorithm for Generation of (X,Y) Observations

This section details a straightforward algorithm for simulating observations (X,Y) from the bivariate Dirichlet distribution defined by its probability density function (pdf) in Equation (1). As demonstrated by the marginal distribution of X being Beta(a,c) (Equation 3) and the conditional distribution of $T = \frac{Y}{1-x}$ being Beta(b,d) (Equation 20), we can exploit this hierarchical structure for data generation. The following algorithm provides a simple method to obtain bivariate observations from the target distribution:

Algorithm 1

Step 1: Generate a value for X from a Beta distribution with parameters a and c.

Step 2: Generate a value for T independently from a Beta distribution with parameters b and d.

Step 3: Calculate the corresponding value for Y using the transformation Y=T(1−X).

Step 4: The resulting pair (X,Y) constitutes a generated observation from the bivariate Dirichlet distribution.

5. Parameter Estimation

This section focuses on the estimation of the parameters of the probability density function (pdf) provided in Equation (1). Specifically, we derive the estimators using the method of moments and subsequently present the maximum likelihood estimators (MLEs).

5.1 Method of moments estimation

Suppose $(X_1, Y_1), \ldots, (X_n, Y_n)$ is a random sample from the distribution (1). Using the first two moments of the marginal distribution of X and the conditional distribution of Y/(1-x) given X=x, we derive the estimators of a,c and b,d respectively, as follows:

If $a \neq b \neq c \neq d$,

For obtaining estimators of a and c, we set

$$\frac{1}{n}\sum_{i=1}^{n} X_i = \mathrm{E}(X) = \overline{X} = \frac{a}{(a+c)}, \qquad (31)$$

and

$$\frac{1}{n}\sum_{i=1}^{n} X_i^{2} = \mathrm{E}(X^2) = Var(X) + (\mathrm{E}(X))^2 = \frac{a(a+1)}{(a+c)(a+c+1)} \qquad (32)$$

For obtains estimators of b and d, we set

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{Y_i}{1-X_i}\right) = \frac{b}{(b+d)} \qquad (33)$$

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{Y_i}{1-X_i}\right)^2 = \frac{b(b+1)}{(b+d)(b+d+1)} \qquad (34)$$

Solving simultaneously Equations ((31)- (34)) for a,b,c and d we get

Respectively the corresponding method of moments estimators (MMEs)

$$\hat{a} = \frac{\overline{X}\left(\overline{X} - \frac{1}{n}\sum_{i=1}^{n} X_i^{\,2}\right)}{\frac{1}{n}\sum_{I=1}^{n} X_i^{\,2} - \overline{X}^2} \tag{35}$$

$$\hat{c} = \frac{\left(1-\overline{X}\right)\left(\overline{X} - \frac{1}{n}\sum_{i=1}^{n} X_i^{\,2}\right)}{\frac{1}{n}\sum_{I=1}^{n} X_i^{\,2} - \overline{X}^2} \tag{36}$$

$$\hat{b} = \frac{V(V-U)}{U-V} \tag{37}$$

and

$$\hat{d} = \frac{(1-V)(V-U)}{U-V} \tag{38}$$

where $V = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{Y_i}{1-X_i}\right)$,and $U = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{Y_i}{1-X_i}\right)^2$.

Clearly $\hat{a}$ and $\hat{c}$ are moments estimators, while $\hat{b}$ and $\hat{d}$ are approximate moments estimators.

5.2. The Maximum Likelihood (MLE)

Let $(X_1, Y_1),...,(X_n, Y_n)$ be a random sample from the distribution (1). Nadarajah and Kotz [2007] studied the maximum likelihood (MLE) estimators of the distribution (1).

Then, the maximum likelihood (MLE) estimators of (a,b,c,d) are obtained by solving the following equations simultaneously for a,b,c and d.

$$\sum_{i=1}^{n} \log x_i = -n\psi(a+c) + n\psi(a) \tag{39}$$

$$\sum_{i=1}^{n} \log y_i - \sum_{i=1}^{n} \log(1-x_i) = -n\psi(b+d) + n\psi(b) \tag{40}$$

$$\sum_{i=1}^{n} \log(1-x_i) = -n\psi(a+c) + n\psi(c) \tag{41}$$

and

$$\sum_{i=1}^{n} \log(1-x_i-y_i) - \sum_{i=1}^{n} \log(1-x_i) = -n\psi(b+d) + n\psi(d) \tag{42}$$

where $\psi(x) = d\log\Gamma(x)/dx$ denotes the digamma function.

There is no exact solution for Equations (39)-(42), but can be solved numerically.

6. Numerical illustration

A numerical comparison of the method of moments and maximum likelihood estimation (MLE) was conducted to evaluate their performance in terms of mean squared error (MSE) and bias. This involved generating 200 independent random datasets from the bivariate

Dirichlet distribution specified in Equation (1), with each dataset having a sample size of $n \in \{5,10,15,20,30\}$. The simulated observations (X,Y) were generated according to Algorithm 1. For each of the 200 simulated datasets, the parameter estimates were obtained for different combinations of a,b,c, and d by applying the formulas derived for the method of moments in Equations (35)-(38) and for the maximum likelihood estimation (MLE) in Equations (39)-(42). Two specific scenarios for the parameter values were examined: Case 1, where (a=2,c=1,b=4,d=3), and Case 2, where (a=4,c=3,b=2,d=1). It should be emphasized that these parameter values were chosen arbitrarily to illustrate the behavior of the estimators. In Case 1, we considered a parameter setting where a,c < b,d while in Case 2, the opposite relationship a,c > b,d was investigated. The resulting bias and MSE for the parameter estimators in both cases are presented in Tables 1 through 4. All numerical computations were performed using a MATLAB program.

Table(1) The Bias of the estimators of (a,b,c,d). when (a=2,c=1,b=4 and d=3).

Parameters / Estimators		a	b	c	d
n=5	MME	3.0309	5.5273	1.3775	4.2905
	MLE	2.9610	5.7576	1.4210	4.4740
n=10	MME	0.8400	1.5475	0.3969	1.1253
	MLE	0.9850	1.6651	0.4724	1.2143
n=15	MME	0.5951	0.8081	0.2322	0.5845
	MLE	0.5618	0.8901	0.2104	0.6429
n=20	MME	0.3646	0.7927	0.1963	0.5669
	MLE	0.3517	0.8500	0.1792	0.6104
n=30	MME	0.2519	0.2395	0.0752	0.2246
	MLE	0.2678	0.2933	0.0773	0.2689

Table(2) MES of the estimators of (a,b,c,d). when (a=2,c=1,b=4 and d=3).

Parameters / Estimators		a	b	c	d
n=5	MME	58.7551	235.3628	17.8314	118.3329
	MLE	51.0677	239.6029	18.2878	121.2371
n=10	MME	5.5171	9.6759	0.8314	5.2540
	MLE	5.5924	10.1622	1.0102	5.5500
n=15	MME	2.3824	5.1362	0.3012	2.7081
	MLE	1.9822	5.3426	0.2420	2.8157
n=20	MME	0.7026	3.7932	0.2040	2.1676
	MLE	0.6564	3.8197	0.1848	2.1977
n=30	MME	0.5246	1.7351	0.0979	0.9372
	MLE	0.4982	1.7256	0.0900	0.9362

Table(3) The Bias of the estimators of (a,b,c,d). when (a=4,c=3,b=2 and d=1).

Parameters / Estimators		a	b	c	d
n=5	MME	4.4189	2.9008	3.0896	1.2041
	MLE	4.5585	2.8990	3.2119	1.2610
n=10	MME	1.3683	0.7995	1.1035	0.3186
	MLE	1.3931	0.8275	1.1264	0.3269
n=15	MME	0.6717	0.4600	0.5452	0.1756
	MLE	0.7373	0.4763	0.5979	0.1806
n=20	MME	0.5211	0.3506	0.4234	0.1510
	MLE	0.5643	0.3623	0.4583	0.1531
n=30	MME	0.3390	0.2164	0.1855	0.1028
	MLE	0.3491	0.2441	0.1870	0.1135

Table(4) MES of the estimators of (a,b,c,d). when (a=4,c=3,b=2 and d=1).

Parameters / Estimators		a	b	c	d
n=5	MME	95.9743	91.3910	45.3577	13.6211
	MLE	94.0243	84.1383	45.0749	13.4095
n=10	MME	10.2379	4.5377	6.3146	0.8359
	MLE	9.7983	4.3752	6.3432	0.8300
n=15	MME	4.0496	1.3517	2.9493	0.2681
	MLE	4.0984	1.0974	3.0680	0.2465
n=20	MME	3.1955	0.9875	1.7836	0.2326
	MLE	2.8719	0.8531	1.6848	0.1965
n=30	MME	1.2952	0. 6279	0.6207	0.1338
	MLE	1.2465	0.5607	0.6959	0.1140

From Table (1), we see that the bias of MMEs for all parameters is less than of the MLEs for all values of n except for the parameter a for n=5,15 and 20, and parameter c for n=15 and 20.

However the differences in the bias are not significant.

From Table (2), we see that the MSE of MMEs of all parameters is less than that of the MLEs for all values of n except for the parameter a for n=5,15,20 and 30, parameter b for n =30, parameter c for n =15,20 and 30 and parameter d for n=30.

Again the differences in the MES between the two estimator are very small.

From Table (3), we see that the bias of MMEs is smallest that of MLEs for all parameters and all values of n except for the parameter b when n=5.

From Table (4), we see that the differences in the MSE of the two estimator are very small, but however the MMEs have smaller MSE in most cases.

We may conclude that both estimators perform approximately the same regarding biasness and MSE.

References

[1] A-Grivas, D. and Asaoka,A. (1982). Slope safety prediction under static and seismic loads, Journal of the Geotechnical Engineer Division Proceedings of the American Society of Civil Engineers, **108**, 713-729

[2] Apostolakis, F.J and Moieni, P. (1987). The foundations of models of dependence in probability safety assessment, Reliability Engineering ,**18**, 177-195.

[3] Bouguila, N et al. (2004). Unsupervised learning of a finite mixture model based on the Dirichlet distribution and its application . IEEE Transactins on Image Processing, **13**(11):1533-1543.

[4] Chatfiled, C. (1975). A marking application of a characterization theorem. In: A Modern course on Distributions in Scientific Work, Vol. 2: "Model Building and Model Selection " (G. P. Patil, S.Kotz and J. K. Ord, eds.),pp. 175-185. Reidel, Dordrecht, The Netherland.

[5] Connor, R. and Mosimann, J. (1969). Concepts of independence for proportions with a generalization of the dirichlet distribution. Journal of the American Statistical Association, 64**,** 194–206.

[6] Godhart, et al. (1984). The Dirichlet: A comprehensive model of bying behavior. Journal of the Royal Statistical Society. Series A(General), pages 621-655.

[7] Gradshteyn, I. and Ryzhik, I.(1980). Table of Integrals, Series, Products, Academic Press.

[8] Gupta, A. K. and Nagar,D.K. (2000). Matrix Variate Distributions. New York: Chapman and Hall/ CR.

[9] Gupta, A. K. and Nadarajah, S. (2004). "Handbook of Beta Distribution and Its Applications". Marcel Dekker, New York.

[10] Hoyer,R. W. and Mayer, L. (1976). The equivalence of various objective functions in a stochastic model of electoral competition. Technical Report, No.114,Series2, Department of Statistic, Princeton University.

[11] Lange, K..(1995). Applications of Dirichlet distributions to forensic match probabilities. Genetica, **96(1-2)** ,107-117.

[12] Libby, D.L. and Novick, M.R.(1982). Multivariate generalized beta-distributions with applications to utility assessment . Advances in Statistical Analysis **91** ,93-106.

[13] Nadarajah, S. and Kotz, S. (2007). Proportions, sums and ratios. Advances in Statistical Analysis, **91**, 93–106.

[14] Nagar, D. and Alvarez, A.J. (2005). Properties of the hypergeomatric function type I distribution. Advances and Applications in Statistical , **5**, No.3, 341–351.

[15] Null, B. (2009). Modeling baseball player ability with a nested Dirichlet distribution. Journal of Quantitative Analysis in Sport , **5**(2).

[16] Sahin, B et al . (2023). Parameter estimation of the Dirichlet distribution based on entropy . Aximos, 12, 947.

[17] Trick, S et al . (2023). Parameter estimation for a bivariate beta distribution with arbitrarybeta marginal and positive correlation .Metron, **81**,163-180

[18] Wang, K et al . (2011). Dirichlet and related distributions .Theory, Methods and Applications, Vol888, chapter 1, John wiley and sons.

[19] Wilks, S . (1962). Mathematical statistics. New York: John wiley and sons.