Revista Colombiana de Estadística - Theoretical Statistics
January 2026, volume 49, issue 1, pp. 131–159
http://doi.org/10.15446/rce.v49n1.118380

Bivariate Simplex Distribution

Distribución Simplex Bivariada

Emerson Amaral[1,a], Lucas S. Vieira[1,b], Lizandra C. Fabio[1,c], Vanessa Barros[2,d], Jalmar M. F. Carrasco[1,e]

[1]Department of Statistics, Institute of Mathematics and Statistics, Federal University of Bahia, Salvador, Brazil

[2]Department of Mathematics, Institute of Mathematics and Statistics, Federal University of Bahia, Salvador, Brazil

Abstract

This article proposes a bivariate Simplex distribution for modeling continuous outcomes constrained to the interval $(0, 1)$, which can represent proportions, rates, or indices. We derive analytical expressions to calculate the dependence between the variables and examine its relationship with the association parameter. Parameters are estimated using the maximum likelihood method, and their performance is assessed through Monte Carlo simulations. The simulations explore various aspects of the bivariate distribution, including different surfaces and contour graphs. To demonstrate the methodology and properties of the proposed model, we present two empirical applications in the areas of Psychometry and Jurimetry. Supplementary material can be accessed via the following link: `https://github.com/carrascojalmar/BSD.git` includes user-friendly code and simulation results.

***Keywords*:** Copula; Jurimetric; Monte Carlo; Psychometry; Simplex distribution.

Resumen

Este artículo propone una distribución Simplex bivariada para modelar resultados continuos restringidos al intervalo (0,1), que puede representar proporciones, tasas o índices. Se derivan expresiones analíticas para calcular la dependencia entre las variables y examinar su relación con el parámetro de asociación. Los parámetros se estiman mediante el método de máxima verosimilitud y su rendimiento se evalúa mediante simulaciones de Monte

[a]B.Sc. E-mail: emerson.mrlv@gmail.com
[b]M.Sc. E-mail: lucasestatistica@hotmail.com
[c]Ph.D. E-mail: lizandra.fabio@ufba.com
[d]Ph.D. E-mail: vbarrosoliveira@gmail.com
[e]Ph.D. E-mail: carrasco.jalmar@ufba.com

Carlo. Las simulaciones exploran diversos aspectos de la distribución bivariada, incluyendo diferentes superficies y gráficos de contorno. Para demostrar la metodología y las propiedades del modelo propuesto, se presentan dos aplicaciones empíricas en las áreas de psicometría y jurimetría. Material complementario disponible en: `https://github.com/carrascojalmar/BSD.git` incluye código intuitivo y resultados de simulación.

***Palabras clave*:** Cópula; Distribución simplex; Jurimétrica; Monte Carlo; Psicometría.

1. Introduction

Relevant scientific studies have provided data describing intrinsic phenomena regarding rates, fractions, proportions, or indices. For instance, the distribution supported on the interval (0,1) plays a crucial role in research and application in finance (Gómez-Déniz et al., 2014) and (Biswas et al., 2021), public health (Mazucheli et al., 2019) and (Biswas & Chakraborty, 2019) and demographics (Andreopoulos et al., 2019). In this context, the Beta and Simplex distributions are particularly prominent, with their density functions capable of assuming different shapes depending on parameter values.

Bivariate distributions are essential in practice because they allow simultaneous analysis and decision-making regarding two related or dependent variables. Methods for constructing joint distributions for random variables can be found in Lai & Balakrishnan (2009) and Kotz et al. (2019). Specifically, for the Beta distribution, Barros (2015) proposed parameter estimation methods for the bivariate Beta distribution introduced by Nadarajah & Kotz (2005). Further studies Arnold & Ng (2011), who explored the bivariate Beta distributions for correlated data, and Gupta & Wong (1985), who examined two bivariate Beta distributions derived from the Morgenstern curves and the bivariate Dirichlet distribution, respectively. Other notable contributions include Sarabia & Castillo (2006), who studied various bivariate extensions under a Bayesian framework, and Olkin & Liu (2003) who demonstrated a positively dependent bivariate Beta distribution via the likelihood ratio. The extension of the Beta distribution to the multivariate case $(0,1)^s$ was investigated by Jones (2002), while Machado Moschen & Carvalho (2023) analyzed the Beta distribution proposed by Olkin & Trikalinos (2015) using both classic and Bayesian approach.

Despite these advances, the bivariate Simplex distribution remains relatively unexplored for modeling the distribution of two proportions, such as the proportion of budget allocated to different sectors. Bivariate distributions are often constructed using copula functions, which allow the analysis of dependence structures between two random variables independently of their marginal distributions. This approach offers flexibility in combining different types of marginal distributions. Therefore, this article proposes deriving the bivariate Simplex distribution via copulas[1] as an alternative method for analyzing bivariate data constrained to the

[1] A copula describes a joint distribution function in terms of its marginals and is widely used in empirical analysis across various fields, including survival analysis, actuarial sciences, marketing,

standard unit interval. This contribution is significant for the Simplex distribution framework.

The article is structured as follows: Section 2 reviews the properties and inferential processes associated with the Simplex distribution. Section 3 introduces the bivariate Simplex distribution via copulas, develops analytical expressions for calculating dependence between variables, and provides estimators using the maximum likelihood method. A Monte Carlo simulation study is conducted to investigate the asymptotic behavior of these estimators. Section 5 applies the proposed methodology to two real datasets from the fields of Psychometry and Jurimetry to assess its empirical performance. Finally, Section 6 summarizes the conclusions.

2. Preliminaries

The Simplex distribution was proposed by Barndorff-Nielsen & Jørgensen (1991); later introduced into a class of dispersion models by Jorgensen (1997), which extended the generalized linear models (GLMs) (Nelder & Wedderburn, 1972). The Simplex distribution is very convenient and flexible regarding data restricted to the continuous unit interval (0,1), which can be interpreted as proportions, rates, or indices. Let y be a random variable that follows a Simplex distribution, with parameters $\mu \in (0,1)$ and $\sigma^2 > 0$. The probability density function (pdf) of this distribution is given by

$$f(y;\mu,\sigma^2) = \{2\pi\sigma^2[y(1-y)]^3\}^{-1/2} \exp\Big\{-\frac{1}{2\sigma^2}d(y;\mu)\Big\}, \tag{1}$$

where $0 < y < 1$ and $d(y;\mu) = (y-\mu)^2/y(1-y)\mu^2(1-\mu)^2$ is the unit deviation. The expected value and variance of Y are given by $\mathrm{E}(Y) = \mu$ and

$$\mathrm{Var}(Y) = \mu(1-\mu) - \sqrt{\frac{1}{2\sigma^2}} \exp\Big\{\frac{1}{2\sigma^2\mu^2(1-\mu)^2}\Big\}\Gamma\Big\{\frac{1}{2}, \frac{1}{2\sigma^2\mu^2(1-\mu)^2}\Big\},$$

where $\Gamma(a,b) = \int_b^\infty x^{a-1}e^{-x}dx$ is the incomplete gamma function. In addition, the variance function is given by $V(\mu) = \mu^3(1-\mu)^3$.

The Simplex distribution can take several shapes depending on the parameter values (μ, σ^2), as shown in Figure 1. If y is a random variable that follows the Simplex distribution with mean μ and dispersion parameter σ^2, it frequently is denoted by $y \sim S(\mu,\sigma^2)$. It was shown that (i) $\mathrm{E}[d'(y;\mu)] = 0$, (ii) $\mathrm{Var}[d(y;\mu)] = 2(\sigma^2)^2$, (iii) $\mathrm{E}[d(y;\mu)] = \sigma^2$, (iv) $\mathrm{E}[(y-\mu)d(y;\mu)] = 0$, (v) $\mathrm{E}[(y-\mu)d^2(y;\mu)] = 0$, (vi) $\mathrm{E}[(y-\mu)d''(y;\mu)] = -2\sigma^2$ and (vii) $\frac{1}{2}\mathrm{E}[(d''(y;\mu)] = \frac{3\sigma^2}{\mu(1-\mu)} + \frac{1}{\mu^3(1-\mu)^3}$, where $d'(y;\mu) = \partial d(y;\mu)/\partial\mu$ e $d''(y;\mu) = \partial^2 d(y;\mu)/\partial\mu^2$; for more details see Song & Tan (2000).

medical statistics, and econometrics.

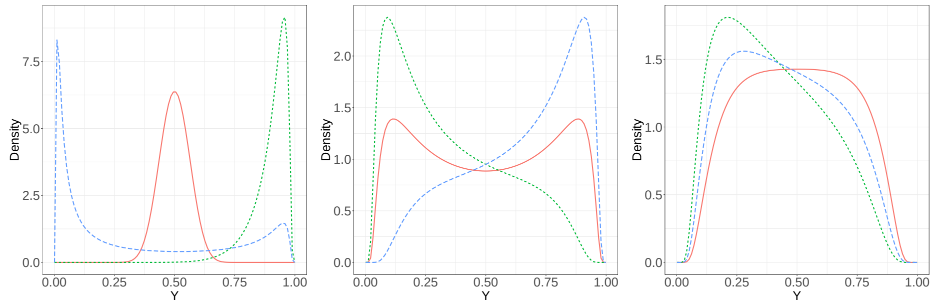

Figure 1: Density graph of the Simplex Distribution for different parameter values.

Let $y_1, y_2, \ldots, y_n$ be a random sample, such that y_i follows the Simplex distribution, given in (1) for all $i = 1, \ldots, n$. The likelihood function for the independent observations is defined as $\mathrm{L}(\boldsymbol{\theta}; \boldsymbol{y}) = \prod_{i=1}^{n} f(y_i; \boldsymbol{\theta})$, where $\boldsymbol{\theta} = (\mu, \sigma^2)^\top$. The logarithm of the likelihood function is expressed in the form $\ell(\boldsymbol{\theta}; \boldsymbol{y}) = \sum_{i=1}^{n} \ell_i(\boldsymbol{\theta}; y_i)$, where $\ell_i(\boldsymbol{\theta}, y_i) = -\log(2\pi)/2 - \log(\sigma^2)/2 - 3\log[y_i(1-y_i)]/2 - d(y_i; \mu)/2\sigma^2$. The maximum likelihood estimators for μ and σ^2 are found by simultaneously solving the estimation equations, i.e., $\partial\ell(\boldsymbol{\theta}; \boldsymbol{y})/\partial\mu = 0$ e $\partial\ell(\boldsymbol{\theta}; \boldsymbol{y})/\partial\sigma^2 = 0$, where

$$\frac{\partial\ell(\boldsymbol{\theta}; \boldsymbol{y})}{\partial\mu} = -\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(-\frac{2(y_i-\mu)}{\mu(1-\mu)}\left[d(y_i,\mu) + \frac{1}{\mu^2(1-\mu)^2}\right]\right),$$
$$\frac{\partial\ell(\boldsymbol{\theta}; \boldsymbol{y})}{\partial\sigma^2} = -\frac{n}{2\sigma^2} + \sum_{i=1}^{n}\frac{d(y_i,\mu)}{2(\sigma^2)^2}.$$

It is easy to find that $\widehat{\sigma}^2 = \sum_{i=1}^{n} d(y_i, \mu)/n$. The second derivatives of $\ell(\boldsymbol{\theta}; \boldsymbol{y})$ related to the parameter vector are given by $\partial^2\ell(\boldsymbol{\theta}; \boldsymbol{y})/\partial\mu^2 = \sum_{i=1}^{n} -d''(y_i,\mu)/2\sigma^2$, $\partial^2\ell(\boldsymbol{\theta}; \boldsymbol{y})/\partial(\sigma^2)^2 = n/2\sigma^4 - \sum_{i=1}^{n} d(y_i,\mu)/(\sigma^2)^3$ and $\partial^2\ell(\boldsymbol{\theta}; \boldsymbol{y})/\partial\mu\partial\sigma^2 = -\sum_{i=1}^{n} d'(y_i,\mu)/(\sigma^2)^2$ where $d''(y_i,\mu) = \partial^2 d(y_i,\mu)/\partial\mu^2$. Using properties (i), (iii), and (vii) above, the Fisher's information matrix is given by

$$K(\boldsymbol{\theta}) = K(\mu, \sigma^2) = \begin{pmatrix} K_{\mu\mu} & K_{\mu\sigma^2} \\ K_{\sigma^2\mu} & K_{\sigma^2\sigma^2} \end{pmatrix},$$

where $K_{\mu\mu} = -\mathrm{E}[\partial^2\ell(\boldsymbol{\theta}; \boldsymbol{y})/\partial\mu^2] = 3n/\mu(1-\mu) + n/\sigma^2\mu^3(1-\mu)^3$, $K_{\mu\sigma^2} = K_{\sigma^2\mu} = -\mathrm{E}[\partial^2\ell(\boldsymbol{\theta}; \boldsymbol{y})/\partial\mu\partial\sigma^2] = 0$ and $K_{\sigma^2\sigma^2} = -\mathrm{E}[\partial^2\ell(\boldsymbol{\theta}; \boldsymbol{y})/\partial(\sigma^2)^2] = -n/2\sigma^4$ respectively. Under general regularity conditions, the maximum likelihood estimators are consistent, and the asymptotic distribution of $\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ is normal with mean zero and covariance matrix $K^{-1}(\boldsymbol{\theta})$.

3. Bivariate Simplex Distribution

In this section, we derive the bivariate Simplex distribution via copulas because bivariate distributions based on copulas are powerful tools for modeling complex dependence structures between two variables. By separating the marginal distributions from the dependency, copulas offer more flexibility and accuracy, especially in fields where non-linear relationships and tail dependencies are essential. Analytical expressions are derived for the expected dependence between variables.

The copula was introduced by Sklar (1959), although similar ideas and results can be found in Hoeffding (1940). The copula function is one of several ways of generating multivariate distributions. A copula is defined as a joint distribution function of the form

$$C(u_1, u_2, \ldots, u_k) = P(U_1 \leq u_1, U_2 \leq u_2, \ldots, U_k \leq u_k), \tag{2}$$

where $0 \leq u_j \leq 1$, $U_j \sim U(0,1)$, for every $j = 1, 2, \ldots, k$. Suppose $H(\cdot)$ is a k-dimensional cumulative distribution function with marginals $F_1(\cdot), \ldots, F_k(\cdot)$. Then, according to Sklar (1959), there is a k-dimensional copula C such that, for every $(y_1, \ldots, y_k) \in (-\infty, \infty)^k$, $H(y_1, \ldots, y_k) = C(F_1(y_1), \ldots, F_k(y_k))$. In this sense, C is unique if $F_1, \ldots, F_k$ are all continuous. The above result guarantees that we can find the joint distribution of k random variables $y_1, \ldots, y_k$, i.e., given a set of continuous random variables with marginal distribution functions $F_1(y_1), \ldots, F_k(y_k)$ and a distribution function $H(y_1, \ldots, y_k)$; the joint density function is given by

$$\begin{aligned} h(y_1, ..., y_k) \quad &= \quad \frac{\partial^k H(y_1, ..., y_k)}{\partial y_1 ... \partial y_k} = \frac{\partial^k C(F_1(y_1), ..., F_k(y_k))}{\partial F_1(y_1) ... \partial F_k(y_k)} \frac{\partial F_1(y_1)}{\partial y_1} \times ... \times \\ & \frac{\partial F_k(y_k)}{\partial y_k} = c(F_1(y_1) ... F_k(y_k)) \prod_{i=1}^{k} f_i(y_i), \end{aligned}$$

where $c(F_1(y_1), ..., F_k(y_k)) = \dfrac{\partial^k C(F_1(y_1), ..., F_k(y_k))}{\partial F_1(y_1) \ldots \partial F_k(y_k)}$ and $f_i(y_i) = \dfrac{\partial F_i(y_i)}{\partial y_i}$. In particular, if y_1 and y_2 are two random variables with a joint distribution function $F(y_1, y_2)$ and continuous marginal distribution functions $F_1(y_1)$ and $F_2(y_2)$, respectively, then there exists a single copula $C : [0,1]^2 \rightarrow [0,1]$ such that for all $(y_1, y_2) \in R^2$ and $F(y_1, y_2) = C(F_1(y_1), F_2(y_2))$, expressing a bivariate density function with marginals F_1 and F_2.

The FGM copula (Farlie-Gumbel-Morgenstern copula) is a relatively simple and classical copula used to model the dependence between two random variables. While it is not as flexible as other copulas, it provides a basic structure to capture weak positive or negative dependence. The FGM copula is defined as

$$C(F_1(y_1), F_2(y_2); \lambda) = F_1(y_1)F_2(y_2) + \lambda F_1(y_1)F_2(y_2)(1 - F_1(y_1))(1 - F_2(y_2)),$$

where $\lambda \in [-1, 1]$ is the dependence parameter. If $\lambda = 0$, the copula reduces to the independence copula, meaning $C(F_1(y_1), F_2(y_2)) = F_1(y_1) \times F_2(y_2)$, which indicates that the two random variables are independent; for $\lambda > 0$ (Positive dependence), meaning the two variables tend to increase together (though weakly); for $\lambda < 0$ (Negative dependence), meaning that when one variable increases, the other tends to decrease (weak negative dependence).

The Farlie-Gumbel-Morgenstern (FGM) copula is a valuable tool in the realm of dependence modeling due to its simplicity and interpretability. Its structure offers an intuitive way to capture weak dependence between variables, making it particularly useful in scenarios where more complex copulas may be unnecessary, or overfitting may be a concern. The FGM copula is symmetric, which allows it to treat both variables equally, and its dependence parameter λ is easy to interpret, making it accessible for analysts and practitioners alike. While it does not capture extreme tail dependence, this feature can be advantageous in applications where only moderate associations are needed, such as in educational settings, initial exploratory analyses, or simple bivariate models. The FGM copula's tractability and low computational demands make it a practical choice when the goal is to understand basic dependence structures without the overhead of more complex copulas, allowing it to strike a balance between simplicity and effectiveness.

Thus, in this particular paper, we assume that the variables y_1 and y_2 follow univariate Simplex distributions, and the density function $f(y_1, y_2)$, is as follows:

$$\begin{aligned} f(y_1, y_2; \boldsymbol{\theta}) = \{2\pi\sigma_1^2[y_1(1-y_1)]^3\}^{-1/2} \exp\left\{-\frac{1}{2\sigma_1^2} d(y_1; \mu_1)\right\} \times \\ \{2\pi\sigma_2^2[y_2(1-y_2)]^3\}^{-1/2} \exp\left\{-\frac{1}{2\sigma_2^2} d(y_2; \mu_2)\right\} \times \\ \{1 + \lambda[1 - 2F_1(y_1)][1 - 2F_2(y_2)]\}, \end{aligned} \quad (3)$$

where $F_1(y_1)$ and $F_2(y_2)$ are the distribution functions of y_1 and y_2, respectively, and $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \lambda)^\top$ is the vector of unknown parameters, were μ_1, μ_2 and σ_1, σ_2 represent the location and dispersion parameters, respectively and λ is the dependency parameter. We denote $\boldsymbol{y} \sim S^2(\boldsymbol{\mu}, \boldsymbol{\sigma}^2, \lambda)$ the random variable vector $\boldsymbol{y} = (y_1, y_2)^\top$ which follows the bivariate Simplex distribution with $\boldsymbol{\mu} = (\mu_1, \mu_2)^\top$ and $\boldsymbol{\sigma} = (\sigma_1, \sigma_2)^\top$ parameters.

3.1. Maximum Likelihood Estimation

Let $\boldsymbol{y} = (\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n)^\top$ represent a vector, where $\boldsymbol{y}_i = (y_{i1}, y_{i2})^\top$ follow the bivariate Simplex distribution given in (3) , for $i = 1, \ldots, n$. The likelihood function of all pairs of observations is defined as $\mathrm{L}(\boldsymbol{\theta}, \boldsymbol{y}) = \prod_{i=1}^n f(\boldsymbol{y}_i; \boldsymbol{\theta})$, where $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \lambda)^\top$ and the logarithm of the likelihood function is given by $\ell(\boldsymbol{\theta}; \boldsymbol{y}) = \sum_{i=1}^n \ell_i(\boldsymbol{\theta}; \boldsymbol{y}_i)$, where

$$\begin{aligned} \ell_i(\boldsymbol{\theta}; \boldsymbol{y}_i) = -\log(2\pi) - \frac{1}{2}\log(\sigma_1^2) - \frac{3}{2}\log[y_{i1}(1-y_{i1})] - \frac{1}{2\sigma_1^2} d(y_{i1}; \mu_1) \\ -\frac{1}{2}\log(\sigma_2^2) - \frac{3}{2}\log[y_{i2}(1-y_{i2})] - \frac{1}{2\sigma_2^2} d(y_{i2}; \mu_2) \\ + \log\{1 + \lambda[1 - 2F_1(y_{i1})][1 - 2F_2(y_{i2})]\}. \end{aligned} \quad (4)$$

By partially deriving the logarithm of the likelihood function with respect to the vector of parameters, the elements of the score vector $\boldsymbol{U}(\boldsymbol{\theta}) = (U_{\mu_1}, U_{\mu_2}, U_{\sigma_1^2}, U_{\sigma_2^2},$

$U_\lambda)^\top$, are obtained from the expressions:

$$
\begin{aligned}
U_{\mu_1} = \frac{\partial \ell(\boldsymbol{\theta}; \boldsymbol{y})}{\partial \mu_1} &= \sum_{i=1}^{n} \frac{d(y_{i1}; \mu_1)}{\sigma_1^2 \mu_1 (1-\mu_1)} \left[(y_{i1} - \mu_1) + \frac{y_{i1}(1-y_{i1})}{y_{i1} - \mu_1} \right] + \\
&\quad \frac{2\lambda\{2F_2(y_{i2}) - 1\}\dot{F}_{1\mu_1}}{1 + \lambda\{2F_1(y_{i1}) - 1\}\{2F_2(y_{i2}) - 1\}}, \\
U_{\mu_2} = \frac{\partial \ell(\boldsymbol{\theta}; \boldsymbol{y})}{\partial \mu_2} &= \sum_{i=1}^{n} \frac{d(y_{i2}; \mu_2)}{\sigma_2^2 \mu_2 (1-\mu_2)} \left[(y_{i2} - \mu_2) + \frac{y_{i2}(1-y_{i2})}{y_{i2} - \mu_2} \right] + \\
&\quad \frac{2\lambda\{2F_1(y_{i1}) - 1\}\dot{F}_{2\mu_2}}{1 + \lambda\{2F_2(y_{i2}) - 1\}\ 2F_1(y_{i1}) - 1\}}, \\
U_{\sigma_1^2} = \frac{\partial \ell(\boldsymbol{\theta}; \boldsymbol{y})}{\partial \sigma_1^2} &= \sum_{i=1}^{n} \frac{1}{2\sigma_1^2} \left[\frac{d(y_{i1}; \mu_1)}{\sigma_1^2} - 1 \right] + \frac{2\lambda\{2F_2(y_{i2}) - 1\}\dot{F}_{1\sigma_1^2}}{1 + \lambda\{2F_1(y_{i1}) - 1\}\{2F_2(y_{i2}) - 1\}}, \\
U_{\sigma_2^2} = \frac{\partial \ell(\boldsymbol{\theta}; \boldsymbol{y})}{\partial \sigma_2^2} &= \sum_{i=1}^{n} \frac{1}{2\sigma_2^2} \left[\frac{d(y_{i2}; \mu_2)}{\sigma_2^2} - 1 \right] + \frac{2\lambda\{2F_1(y_{i1}) - 1\}\dot{F}_{2\sigma_2^2}}{1 + \lambda\{2F_2(y_{i2}) - 1\}\{2F_1(y_{i1}) - 1\}}, \\
U_{\lambda} = \frac{\partial \ell(\boldsymbol{\theta}; \boldsymbol{y})}{\partial \lambda} &= \sum_{i=1}^{n} \frac{\{2F_1(y_{i1}) - 1\}\{2F_2(y_{i2}) - 1\}}{1 + \lambda\{2F_1(y_{i1}) - 1\}\{2F_2(y_{i2}) - 1\}},
\end{aligned}
$$

where $\dot{F}_{1\mu_1} = \partial F_1(y_1)/\partial \mu_1$, $\dot{F}_{2\mu_2} = \partial F_2(y_2)/\partial \mu_2$, $\dot{F}_{1\sigma_1^2} = \partial F_1(y_1)/\partial \sigma_1^2$, e $\dot{F}_{2\sigma_2^2} = \partial F_2(y_2)/\partial \sigma_2^2$. Similarly to the univariate case, it is possible to find the observed information matrix, $J(\boldsymbol{\theta}) = \partial \ell(\boldsymbol{\theta}; \boldsymbol{y})/\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top$, whose partial derivatives are given by

$$
\begin{aligned}
J_{\mu_1\mu_1}(\boldsymbol{\theta}) =& \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \mu_1^2} = \sum_{i=1}^{n} - \frac{1}{\sigma_1^2 \mu_1^2 (1-\mu_1)^2} \Big[\frac{3y_{i1}(1-y_{i1})}{\mu_1^2(1-\mu_1)^2} - \frac{2}{\mu_1(1-\mu_1)} + \\
& 3d(y_{i1}; \mu_1) \big[(y_{i1} - \mu_1)^2 + 2y_{i1}(1-y_{i1}) \big] \Big] + \sum_{i=1}^{n} \frac{2\lambda\{2F_2(y_{i2}) - 1\}}{G_{y_{i1}y_{i2}}^2} \times \\
& \Big[\ddot{F}_{1\mu_1} G_{y_{i1}y_{i2}} - 2\lambda\{F_2(y_{i2}) - 1\}\dot{F}_{1\mu_1} \Big], \\
J_{\mu_1\sigma_1^2}(\boldsymbol{\theta}) =& \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \mu_1 \partial \sigma_1^2} = \sum_{i=1}^{n} - \frac{d(y_{i1}; \mu_1)}{\sigma_1^4 \mu_1 (1-\mu_1)} \Big[y_{i1} - \mu_1 + \frac{y_{i1}(1-y_{i1})}{y_{i1} - \mu_1} \Big] + \\
& \frac{2\lambda\{2F_2(y_{i2}) - 1\}}{G_{y_{i1}y_{i2}}^2} \Big[\ddot{F}_{1\mu_1\sigma_1^2} G_{y_{i1}y_{i2}} - 2\lambda \dot{F}_{1\mu_1} \dot{F}_{1\sigma_1^2} \{2F_2(y_{i2}) - 1\} \Big], \\
J_{\mu_1\mu_2}(\boldsymbol{\theta}) =& \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \mu_1 \partial \mu_2} = \sum_{i=1}^{n} \frac{4\lambda \dot{F}_{1\mu_1} \dot{F}_{2\mu_2}}{G_{y_{i1}y_{i2}}^2}, \\
J_{\mu_1\sigma_2^2}(\boldsymbol{\theta}) =& \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \mu_1 \partial \sigma_2^2} = \sum_{i=1}^{n} \frac{4\lambda \dot{F}_{1\mu_1} \dot{F}_{2\sigma_2^2}}{G_{y_{i1}y_{i2}}^2},
\end{aligned}
$$

$$J_{\mu_1\lambda}(\boldsymbol{\theta}) = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\mu_1\partial\lambda} = \sum_{i=1}^{n} \frac{2\{2F_2(y_{i2})-1\}\dot{F}_{1\mu_1}}{G^2_{y_{i1}y_{i2}}},$$

$$J_{\mu_2\mu_2}(\boldsymbol{\theta}) = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\mu_2^2} = \sum_{i=1}^{n} -\frac{1}{\sigma_2^2\mu_2^2(1-\mu_2)^2}\Big[\frac{3y_{i2}(1-y_{i2})}{\mu_2^2(1-\mu_2)^2} - \frac{2}{\mu_2(1-\mu_2)} +$$
$$3d(y_{i2};\mu_2)\big[(y_{i2}-\mu_1)^2 + 2y_{i2}(1-y_{i2})\big]\Big] + \sum_{i=1}^{n} \frac{2\lambda\{2F_1(y_{i1})-1\}}{G^2_{y_{i1}y_{i2}}} \times$$
$$\Big[\ddot{F}_{2\mu_2}G_{y_{i1}y_{i2}} - 2\lambda\{F_1(y_{i1})-1\}\dot{F}_{2\mu_2}\Big],$$

$$J_{\mu_2\sigma_1^2}(\boldsymbol{\theta}) = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\mu_2\partial\sigma_1^2} = \sum_{i=1}^{n} \frac{4\lambda\dot{F}_{2\mu_2}\dot{F}_{1\sigma_1^2}}{G^2_{y_{i1}y_{i2}}},$$

$$J_{\mu_2\sigma_2^2}(\boldsymbol{\theta}) = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\mu_2\partial\sigma_2^2} = \sum_{i=1}^{n} -\frac{d(y_{i2};\mu_2)}{\sigma_2^4\mu_2(1-\mu_2)}\Big[y_{i2}-\mu_2+\frac{y_{i2}(1-y_{i2})}{y_{i2}-\mu_2}\Big] +$$
$$\frac{2\lambda\{2F_1(y_{i1})-1\}}{G^2_{y_{i1}y_{i2}}}\Big[\ddot{F}_{2\mu_2\sigma_2^2}G_{y_{i1}y_{i2}} - 2\lambda\dot{F}_{2\mu_2}\dot{F}_{2\sigma_2^2}\{2F_1(y_{i1})-1\}\Big],$$

$$J_{\mu_2\lambda}(\boldsymbol{\theta}) = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\mu_2\partial\lambda} = \sum_{i=1}^{n} \frac{2\{2F_1(y_{i1})-1\}\dot{F}_{2\mu_2}}{G^2_{y_{i1}y_{i2}}},$$

$$J_{\sigma_1^2\sigma_1^2}(\boldsymbol{\theta}) = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial(\sigma_1^2)^2} = \sum_{i=1}^{n} \frac{1}{2\sigma_1^4} - \frac{d(y_{i1};\mu_1)}{\sigma_1^6} + \frac{2\lambda\{2F_2(y_{i2})-1\}}{G_{y_{i1}y_{i2}}} \times$$
$$\left[\ddot{F}_{1\sigma_1^2} - \frac{2\lambda\{2F_2(y_{i2})-1\}\dot{F}_{1\sigma_1^2}}{G^2_{y_{i1}y_{i2}}}\right],$$

$$J_{\sigma_1^2\sigma_2^2}(\boldsymbol{\theta}) = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\sigma_1^2\partial\sigma_2^2} = \sum_{i=1}^{n} \frac{4\lambda\dot{F}_{1\sigma_1^2}\dot{F}_{2\sigma_2^2}}{G^2_{y_{i1}y_{i2}}},$$

$$J_{\sigma_1^2\lambda}(\boldsymbol{\theta}) = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\sigma_1^2\partial\lambda} = \sum_{i=1}^{n} \frac{2\{2F_2(y_{i2})-1\}\dot{F}_{1\sigma_1^2}}{G^2_{y_{i1}y_{i2}}},$$

$$J_{\sigma_2^2\sigma_2^2}(\boldsymbol{\theta}) = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial(\sigma_2^2)^2} = \sum_{i=1}^{n} \frac{1}{2\sigma_2^4} - \frac{d(y_{i2};\mu_2)}{\sigma_2^6} + \frac{2\lambda\{2F_1(y_{i1})-1\}}{G_{y_{i1}y_{i2}}} \times$$
$$\left[\ddot{F}_{2\sigma_2^2} - \frac{2\lambda\{2F_1(y_{i1})-1\}\dot{F}_{2\sigma_2^2}}{G^2_{y_{i1}y_{i2}}}\right],$$

$$J_{\sigma_2^2\lambda}(\boldsymbol{\theta}) = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\sigma_2^2\partial\lambda} = \sum_{i=1}^{n} \frac{2\{2F_2(y_{i2})-1\}\dot{F}_{2\sigma_2^2}}{G^2_{y_{i1}y_{i2}}},$$

$$J_{\lambda\lambda}(\boldsymbol{\theta}) = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\lambda^2} = \sum_{i=1}^{n} -\left[\frac{\{2F_1(y_{i1})-1\}\{2F_2(y_{i2})-1\}}{G_{y_{i1}y_{i2}}}\right]^2,$$

where $\ddot{F}_{1_{\mu_1}} = \partial \dot{F}_{1_{\mu_1}}/\partial\mu_1$, $\ddot{F}_{2_{\mu_2}} = \partial \dot{F}_{2_{\mu_2}}/\partial\mu_2$, $\ddot{F}_{1_{\sigma_1^2}} = \partial \dot{F}_{1_{\sigma_1^2}}/\partial\sigma_1^2$, $\ddot{F}_{2_{\sigma_2^2}} = \partial \dot{F}_{2_{\sigma_2^2}}$ $/\partial\sigma_2^2$, $\ddot{F}_{1_{\mu_1\sigma_1^2}} = \partial F_1(y_{i1})/\partial\mu_1\partial\sigma_1^2$, $\ddot{F}_{2_{\mu_2\sigma_2^2}} = \partial F_2(y_{i2})/\partial\mu_2\partial\sigma_2^2$ e $G_{y_{i1}y_{i2}} = 1 + \lambda\{2F_1(y_{i1})-1\}\{2F_2(y_{i2})-1\}$. Under certain regularity conditions, the maximum likelihood estimator $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ approximates a Normal distribution with zero mean and variance and covariance matrix $J^{-1}(\boldsymbol{\theta})$; allowing confidence intervals to be found, hypotheses to be tested and predictions to be made. Ensuring parameter identifiability is essential for reliable inference and meaningful interpretation of model components, particularly in copula-based constructions where dependence and marginal parameters are estimated simultaneously. Regarding identifiability, the proposed bivariate Simplex model satisfies the usual conditions for copula-based models with continuous marginals. The marginal parameters $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$ uniquely determine the corresponding Simplex distributions, while the copula parameter λ controls the dependence structure through a monotonic relationship with concordance measures such as Kendall's τ and Spearman's ρ. Hence, distinct values of λ yield distinct joint distributions. Although the FGM copula captures only weak dependence, our simulation study indicates that the maximum likelihood estimator of λ remains stable and asymptotically unbiased as the sample size increases, supporting the practical identifiability and robustness of this parameter.

3.2. Moments

In this sub-section, we derive the joint expectation $\mathrm{E}(y_1y_2)$ stated in the theorem below. For algebraic convenience, we will denote $y_1 = x$ and $y_2 = y$. The proof involves several transformations and the use of special functions (Bessel and Struve), which we state explicitly for clarity.

Theorem 1. *Let $(x,y)^\top$ be a random vector following the bivariate Simplex distribution, where $(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \lambda)^\top$ is the vector of parameters. The joint moment of x and y is given by*

$$\mathrm{E}[xy] = \mu_x\mu_y + \lambda\left[\frac{r_x^2\pi}{2}\left(\frac{1}{a_x\xi_x} + \frac{1}{a_x} + A_x\right) - \mu_x\right]\left[\frac{r_y^2\pi}{2}\left(\frac{1}{a_y\xi_y} + \frac{1}{a_y} + A_y\right) - \mu_y\right],$$

where, for $m \in \{x,y\}$, $A_m = 1 - 2\Big(K_0(2a_m)L_{-1}(2a_m) + K_1(2a_m)L_0(2a_m)\Big)$,

$$a_m = \frac{(\xi_m+1)^2}{\sigma_m^2\xi_m}, \quad r_m = \frac{1}{\sigma_m\sqrt{2\pi}}, \quad \xi_m = \frac{1}{\mu_m} - 1,$$

and where $K_\nu(\cdot)$ denotes the modified Bessel function of the second kind and $L_\nu(\cdot)$ denotes the modified Struve function.

Proof. We want to calculate $\mathrm{E}[xy] = \int_0^1 \int_0^1 xy\, f(x, y)\, dx\, dy$, where $f(x, y)$ is given in Equation (3). After substituting (3) and rearranging terms, we obtain

$$\mathrm{E}[xy] = \mu_x\mu_y + \lambda \left[2\int_0^1 xg(x;\mu_x,\sigma_x^2)G(x;\mu_x,\sigma_x^2)dx - \mu_x\right] \times \left[2\int_0^1 yh(y;\mu_y,\sigma_y^2)H(y;\mu_y,\sigma_y^2)dy - \mu_y\right], \quad (5)$$

where g and h denote the univariate Simplex densities and G, H their corresponding cumulative distribution functions. Therefore, it is sufficient to compute the univariate integral

$$I = \int_0^1 x\, g(x;\mu,\sigma^2)\, G(x;\mu,\sigma^2)\, dx.$$

From expression (1) we obtain:

$$I = \int_0^1 x\{2\pi\sigma^2[x(1-x)]^3\}^{-1/2} \exp\left(-\frac{d(x,\mu)}{2\sigma^2}\right) \times \left(\int_0^x \{2\pi\sigma^2[t(1-t)]^3\}^{-1/2} \exp\left(-\frac{d(t,\mu)}{2\sigma^2}\right)dt\right)dx,$$

where $d(t;\mu) = (t-\mu)^2/t(1-t)\mu^2(1-\mu)^2$. Letting $r = (2\pi\sigma^2)^{-1/2}$, the previous expression can be simplified by

$$I = r^2 \int_0^1 \frac{1}{x^{1/2}(1-x)^{3/2}} \exp(-\frac{d(x,\mu)}{2\sigma^2}) \int_0^x \frac{1}{t^{3/2}(1-t)^{3/2}} \exp(-\frac{d(t,\mu)}{2\sigma^2})dtdx.$$

Let $\mu = \frac{1}{\xi+1}$. By applying the change of variable $t = \frac{1}{z+1}$ to the inner integral, we obtain:

$$\int_0^x \frac{1}{t^{3/2}(1-t)^{3/2}} \exp\Big(-\tfrac{d(t,\mu)}{2\sigma^2}\Big)\, dt = \int_\infty^{\frac{1}{x}-1} \frac{(z+1)^3}{z^{3/2}} \exp\Big(-\tfrac{1}{2\sigma^2}\frac{(\xi-z)^2(\xi+1)^2}{z\xi^2}\Big)\Big(-\frac{dz}{(z+1)^2}\Big).$$

Simplifying and substituting this result into the expression for I, we obtain

$$I = r^2 \int_0^1 \frac{1}{x^{1/2}(1-x)^{3/2}} \exp\Big(-\tfrac{d(x,\mu)}{2\sigma^2}\Big) \times \int_{\frac{1}{x}-1}^\infty \frac{z+1}{z^{3/2}} \exp\Big(-\tfrac{1}{2\sigma^2}\frac{(\xi-z)^2(\xi+1)^2}{z\xi^2}\Big)dz\, dx.$$

Applying again the same change of variable $x = \frac{1}{\omega+1}$ to the outer integral yields

$$I = r^2 \int_0^\infty \frac{1}{\omega^{3/2}} \exp\Big(-\tfrac{1}{2\sigma^2}\frac{(\xi-\omega)^2(\xi+1)^2}{\omega\xi^2}\Big) \times \int_\omega^\infty \frac{z+1}{z^{3/2}} \exp\Big(-\tfrac{1}{2\sigma^2}\frac{(\xi-z)^2(\xi+1)^2}{z\xi^2}\Big)dz\, d\omega.$$

We observe that, setting $\beta = \frac{(\xi+1)^2}{\sigma^2}$ and $\gamma = \frac{(\xi+1)^2}{\sigma^2\xi^2}$, the exponential term in z can be written as $\exp\{-(\beta z^{-1} + \gamma z)/2\}\exp\{a\}$, where $a = \frac{(\xi+1)^2}{\sigma^2\xi}$. The same form holds for the term in ω. We thus obtain the following simplified expression:

$$\begin{aligned} I &= r^2 \exp(2a) \int_0^\infty \frac{1}{\omega^{3/2}} \exp\Big(-\frac{1}{2}(\beta\omega^{-1} + \gamma\omega)\Big) \times \\ &\int_\omega^\infty \frac{z+1}{z^{3/2}} \exp\Big(-\frac{1}{2}(\beta z^{-1} + \gamma z)\Big) dz d\omega. \end{aligned}$$

Once again, we perform a change of variables in the inner integral. Let $p = z/\omega$. Note that here ω is treated as a constant, so $dz = \omega\, dp$. We then obtain

$$\begin{aligned} I &= r^2 \exp(2a) \int_0^\infty \int_1^\infty \frac{1}{p^{3/2}} \left(\frac{1}{\omega^2} + \frac{p}{\omega}\right) \times \\ &\exp\Big(-\tfrac{1}{2}\left(\beta\omega^{-1} + \gamma\omega + \beta(\omega p)^{-1} + \gamma\omega p\right)\Big)\, dp\, d\omega. \end{aligned}$$

Now we regroup terms and exchange the order of integration, obtaining

$$\begin{aligned} I &= r^2 \exp(2a) \int_1^\infty \int_0^\infty \frac{1}{p^{3/2}} \left(\frac{1}{\omega^2} + \frac{p}{\omega}\right) \times \\ &\exp\Big(-\tfrac{1}{2}\big((\beta + \tfrac{\beta}{p})\,\omega^{-1} + (\gamma + \gamma p)\,\omega\big)\Big)\, d\omega\, dp. \end{aligned} \tag{6}$$

We intend to apply Proposition 1 to Equation (6). For this purpose, we note that the integral above can be written as the sum of two integrals, $I = I_1 + I_2$, where

$$I_1 = r^2 \exp(2a) \int_1^\infty \frac{1}{p^{3/2}} \int_0^\infty \omega^{\nu_1 - 1} \exp\Big(-\tfrac{1}{2}((\beta + \beta/p)\omega^{-1} + (\gamma + \gamma p)\omega)\Big) d\omega dp,$$
$$\text{where } \nu_1 = -1,$$

and

$$I_2 = r^2 \exp(2a) \int_1^\infty \frac{1}{p^{1/2}} \int_0^\infty \omega^{\nu_2 - 1} \exp\Big(-\frac{1}{2}((\beta + \beta/p)\omega^{-1} + (\gamma + \gamma p)\omega)\Big) d\omega dp,$$
$$\text{where } \nu_2 = 0.$$

Now we apply Proposition 1 to I_1 and I_2, substitute these results into expression (6), and obtain the following simplified integral involving Bessel functions:

$$I = 2r^2 \exp(2a) \left\{ \frac{1}{\xi} \int_1^\infty \frac{1}{p} K_1\Big(a\frac{p+1}{p^{1/2}}\Big) dp + \int_1^\infty \frac{1}{p^{3/2}} K_0\Big(a\frac{p+1}{p^{1/2}}\Big) \right\}.$$

Applying the change of variables $q = \sqrt{p}$, we obtain:

$$I = r^2 \exp(2a)\{\frac{1}{\xi}\int_1^\infty \frac{1}{q} K_1\Big(a\frac{q^2+1}{q}\Big) dq + \int_1^\infty K_0\Big(a\frac{q^2+1}{q}\Big) dq\}. \tag{7}$$

For simplicity, we denote J_0 and J_1 by the following integrals

$$J_0 = \int_1^\infty K_0\Big(a\frac{q^2+1}{q}\Big) dq, \quad \text{and} \quad J_1 = \int_1^\infty \frac{1}{q} K_1\Big(a\frac{q^2+1}{q}\Big) dq.$$

We recall that expression (7) represents the integral $\int_0^1 xg(x;\mu,\sigma^2)G(x;\mu,\sigma^2)\,dx$ introduced at the beginning of this proof. Therefore, substituting J_0 and J_1 into the identity (7), we obtain:

$$I = \int_0^1 xg(x;\mu,\sigma^2)G(x;\mu,\sigma^2)dx = r^2 \exp(2a)\{J_1/\xi + J_0\}. \tag{8}$$

Note that the integrals J_0 and J_1 were calculated in Lemma 2 and Corollary 1 respectively:

$$J_1 = \frac{\pi}{4a} \exp -2a, \tag{9}$$

and

$$J_0 = \frac{\pi}{4a}\exp(-2a) + \Big[\frac{\pi}{4} - \frac{\pi}{2}(K_0(2a)L_{-1}(2a) + K_1(2a)L_0(2a))\Big]. \tag{10}$$

Substituting (8) into (5), we obtain the following expression for $\mathrm{E}[xy]$:

$$\begin{aligned}\mathrm{E}[xy] &= \mu_x\mu_y + \lambda\Big[2r_x^2 \exp 2a_x(J_1^x/\xi_x + J_0^x) - \mu_x\Big] \times \\ &\quad \Big[2r_y^2 \exp 2a_y(J_1^y/\xi_y + J_0^y) - \mu_y\Big],\end{aligned} \tag{11}$$

where, for $m \in \{x, y\}$, we have

$$r_m = \frac{1}{\sigma_m\sqrt{2\pi}}, \qquad a_m = \frac{(\xi_m+1)^2}{\sigma_m^2\xi_m}, \qquad \xi_m = \frac{1}{\mu_m} - 1,$$

$$J_0^m = \int_1^\infty K_0\left(a_m\frac{m^2+1}{m}\right) dm, \qquad J_1^m = \int_1^\infty \frac{1}{m} K_1\left(a_m\frac{m^2+1}{m}\right) dm.$$

Finally, substituting the identities (9) and (10) into (11) yields the result.

□

Although the derivation of the proposed model involves special functions (modified Bessel and Struve functions) and some numerical integration, the computational cost is quite modest. These functions are efficiently implemented in standard scientific libraries, and the required integrals are evaluated once per likelihood computation. In practice, the estimation procedure scales well with the sample size: in the Monte Carlo experiments (Section 4) involving 1,000 replications with $n = 1000$, the complete simulation study was executed within a few minutes on a standard computer. This demonstrates that the proposed approach is computationally tractable and suitable for large-sample or resampling analyses such as bootstrapping.

4. Simulation Study

In this section, we perform a Monte Carlo simulation study to assess the asymptotic behavior of the maximum likelihood estimators for the bivariate Simplex distribution. The numerical results are derived on $R = 1,000$ Monte Carlo replications, with sample sizes $n = 50$, 100, 150, 200, and $1,000$. The random response vector $\boldsymbol{y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)^\top$, where $\boldsymbol{y}_i = (y_{1i}, y_{2i})^\top$ is generated by using the algorithm described in Johnson (1987). The algorithm involves the following steps: (i) Generate two independent random variables u_1 and v with uniform distributions, $U(0,1)$; (ii) Compute: $A = \lambda(2u_1 - 1) - 1$, $B = [1 - \lambda(2u_1 - 1)]^2 + 4v\lambda(2u_1 - 1)$ and $u_2 = 2v/(\sqrt{B} - A)$; (iii) Apply the inverse transformation method to obtain $y_1 = F_1^{-1}(u_1)$ and $y_2 = F_2^{-1}(u_2)$, where $F_1(\cdot)$ and $F_2(\cdot)$ are the cumulative distribution function of y_1 and y_2, respectively. The mean, bias, root mean square error (RMSE), and the 95% confidence interval coverage probability are computed based on the following expressions:

$$\overline{\theta}_j = \frac{1}{R}\sum_{i=1}^{R} \widehat{\theta}_j^{(i)}, \qquad \text{Bias}(\theta_j) = \overline{\theta}_j - \theta_j \qquad \text{and} \qquad \text{RMSE}(\theta_j) = \sqrt{\frac{1}{R}\sum_{i=1}^{N} (\widehat{\theta}_j^{(i)} - \theta_j)^2},$$

where $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)^\top = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \lambda)^\top$. We considered four scenarios, in which the parameter λ takes on values of -1, 0 and 1 to perform the behavior of $\widehat{\boldsymbol{\theta}}$.

4.1. Scenario 1

In this scenario, the following vectors are taken as the true values of the parameters $\boldsymbol{\theta}_1 = (0.5, 0.5, 2, 2, 1)^\top$, $\boldsymbol{\theta}_2 = (0.5, 0.5, 5, 5, 1)^\top$ and $\boldsymbol{\theta}_3 = (0.9, 0.9, \sqrt{11}, \sqrt{11}, 1)^\top$. Figure 1 (see supplementary material[2]) illustrates the surface and contour plots of the generated samples. The joint moments $\text{E}(y_1 y_2)$ for $\boldsymbol{\theta_1}$, $\boldsymbol{\theta_2}$ and $\boldsymbol{\theta_3}$ are 0.36, 0.40 and 1.24, respectively. For $\boldsymbol{\theta}_1$, the generated samples are concentrated in the interval $(0.25; 0.75)$, indicating a bimodal behavior. Similarly, for $\boldsymbol{\theta}_2$ and $\boldsymbol{\theta}_3$, the samples are concentrated near zero and one simultaneously and near one, respectively. Figure 2 displays the simulation results for this scenario. The parameter vectors $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2$, and $\boldsymbol{\theta}_3$ are represented by the colors red, green, and blue, respectively. Solid and dashed lines in the figure correspond to the parameters associated with y_1 and y_2, respectively. As expected, the bias and root mean square error

[2] `https://github.com/carrascojalmar/BSD.git`

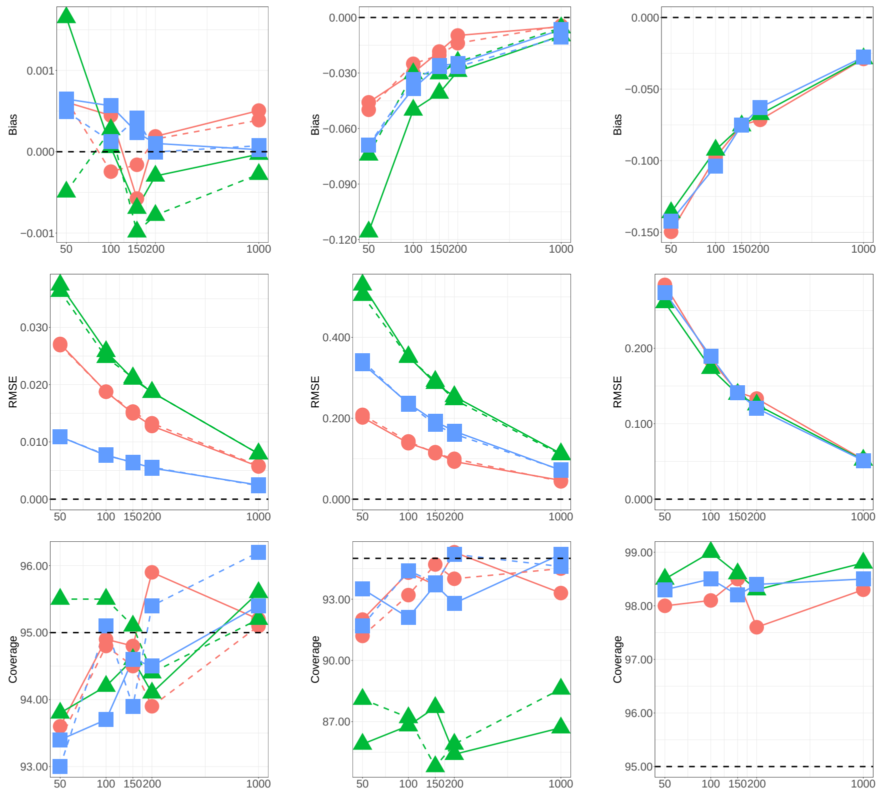

FIGURE 2: Bias (row 1), RMSE (row 2) and Coverage (row 3) of the parameters μ_1 and μ_2 (column 1), σ_1^2 and σ_2^2 (column 2) and λ (column 3) for Scenario 1.

(RMSE) approach zero as the sample size increases, indicating that the maximum likelihood estimators are asymptotically unbiased. The probability of coverage for the parameters μ_1 and μ_2 are close to the nominal 95% level across different parameter vectors $\boldsymbol{\theta}$ and sample sizes. However, for parameters σ_1^2 and σ_2^2, the coverage probability is underestimated when samples are generated using the $\boldsymbol{\theta}_2$ parameter. The coverage of the λ parameters is overestimated for different sample sizes. Tables 1-3 (in the supplementary material) show the results of this scenario.

4.2. Scenario 2

In this scenario, the following vectors are taken as the true values of the parameters $\boldsymbol{\theta}_1 = (0.5, 0.5, 2, 2,\text{-}1)^\top$, $\boldsymbol{\theta}_2 = (0.5, 0.5, 5, 5,\text{-}1)^\top$ and $\boldsymbol{\theta}_3 = (0.9, 0.9, \sqrt{11}, \sqrt{11},\text{-}1)^\top$. Figure 2 (see supplementary material) shows the surface and contour plots of the generated samples. The joint moments $\mathrm{E}(y_1y_2)$ for $\boldsymbol{\theta_1}$, $\boldsymbol{\theta_2}$ and $\boldsymbol{\theta_3}$ in this scenario are 0.14, 0.10 and 0.38, respectively. For $\boldsymbol{\theta}_1$, the samples are concentrated within the interval $(0.25; 0.75)$, indicating a bimodal behavior that is the inverse of the behavior observed in the first scenario. Similarly, for $\boldsymbol{\theta}_2$ and $\boldsymbol{\theta}_3$, the samples are concentrated near zero and one. Figure 3 presents the simulation results for this scenario. Again, the colors red, green, and blue for the parameter vectors $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2$, and $\boldsymbol{\theta}_3$, respectively. The solid and dashed lines correspond to the parameters associated with the variables y_1 and y_2, respectively.

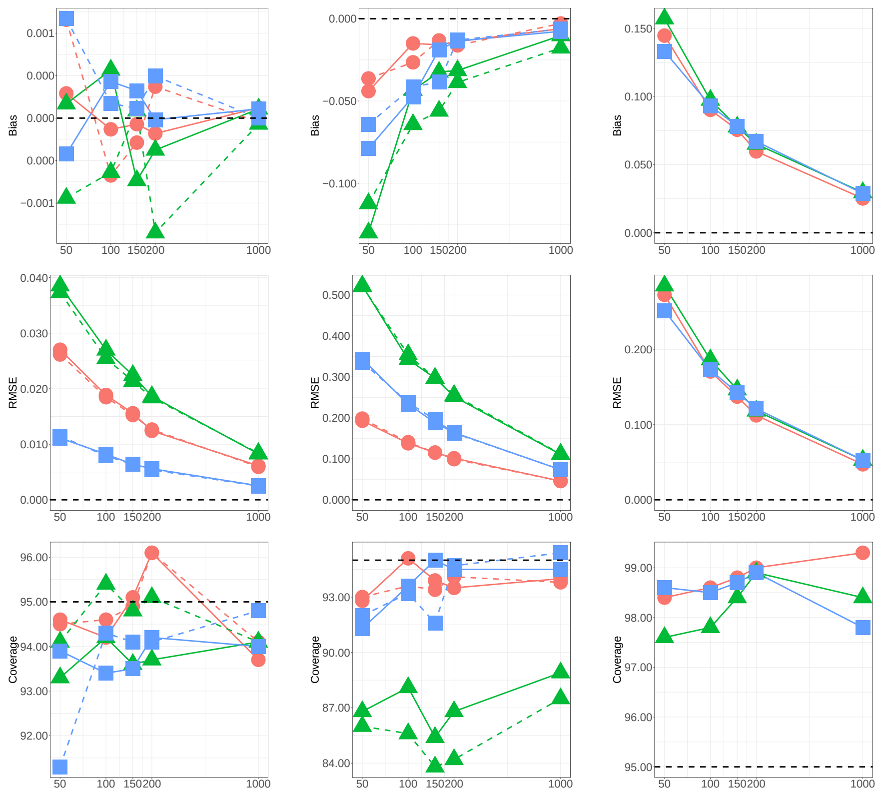

FIGURE 3: Bias (row 1), RMSE (row 2) and Coverage (row 3) of parameters μ_1 and μ_2 (column 1), σ_1^2 and σ_2^2 (column 2) and λ (column 3) for Scenario 2.

In this scenario, as expected, the bias and the root mean square error (RMSE) approach zero as the sample size increases, indicating that the maximum likelihood estimators are asymptotically unbiased. The coverage probability for the parameters μ_1 and μ_2 are close to the nominal 95% level across different parameter vectors $\boldsymbol{\theta}$ and sample sizes. However, for the parameters σ_1^2 and σ_2^2, the coverage is underestimated when $\boldsymbol{\theta}_2$ is assumed. Conversely, the coverage for the parameter λ is overestimated regardless of the parameter vector $\boldsymbol{\theta}$ and the various sample sizes. The results can also be found in the supplementary material in Tables 4-6.

4.3. Scenario 3

In this scenario, the following vectors are taken as the true values of the parameters $\boldsymbol{\theta}_1 = (0.5, 0.5, 2, 2, 0)^\top$, $\boldsymbol{\theta}_2 = (0.5, 0.5, 5, 5, 0)^\top$ and $\boldsymbol{\theta}_3 = (0.9, 0.9, \sqrt{11}, \sqrt{11}, 0)^\top$. Figure 3 (see supplementary material) illustrates the surface and contour plots for the generated samples. The joint moments $\mathrm{E}(y_1 y_2)$ for $\boldsymbol{\theta_1}$, $\boldsymbol{\theta_2}$ and $\boldsymbol{\theta_3}$ are 0.26, 0.25 and 0.81, accordingly. For $\boldsymbol{\theta}_1$, the generated samples are concentrated around 0, 5, indicating an unimodal behavior. Similarly, for $\boldsymbol{\theta}_2$ and $\boldsymbol{\theta}_3$, the samples are concentrated near zero and one. Figure 4 presents the simulation results in this scenario. The parameter vectors $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2$ and $\boldsymbol{\theta}_3$ are represented by the red, green, and blue, respectively. Solid and dashed lines correspond to the parameters associated with y_1 and y_2, respectively.

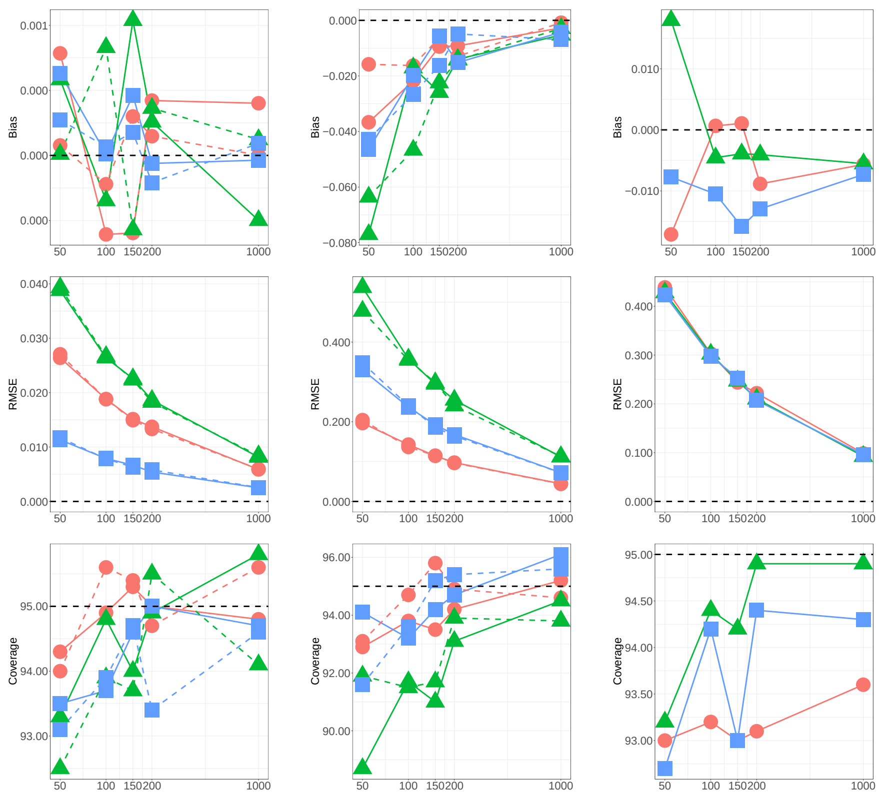

FIGURE 4: Bias (row 1), RMSE (row 2) and Coverage (row 3) of the parameters μ_1 and μ_2 (column 1), σ_1^2 and σ_2^2 (column 2) and λ (column 3) for the Scenario 3.

In this scenario, the bias and root mean square error (RMSE) decreases to zero when the sample size increases. The coverage probability for the parameters μ_1, μ_2, σ_1^2 and σ_2^2 are close to the nominal 95% level across different parameters vector $\boldsymbol{\theta}$ and sample sizes. However, for the parameter λ, only $\boldsymbol{\theta}_2$ achieves coverage close to the nominal 95% level, while for $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_3$, the coverage is underestimated. Again, the results can also be found in Tables 7-9 in the supplementary material.

4.4. Scenario 4

In this scenario, we performed an additional Monte Carlo sensitivity analysis to assess the robustness of the proposed model under alternative copula structures and marginal specifications. The data were generated from the bivariate Simplex distribution via the FGM copula, representing the true data-generating process. Using the same simulated sample, we fitted: Bivariate Simplex models combined with the Clayton, Gaussian, Frank, FGM, and Gumbel copulas Durante & Sempi (2016), and Bivariate Beta models combined with the same set of copulas. Appendix B presents the algebraic form of each copula considered. We considered three sub-scenarios corresponding to different levels of dependence, defined by the parameter vectors $\boldsymbol{\theta}_1 = (0.5, 0.5, \sqrt{5.5}, \sqrt{5.5}, 0.1)^\top$, $\boldsymbol{\theta}_2 = (0.5, 0.5, \sqrt{50}, \sqrt{25}, 0.50)^\top$, $\boldsymbol{\theta}_3 = (0.5, 0.5, \sqrt{5.5}, \sqrt{25}, 0.25)^\top$. These settings allow us to examine model performance under weak, moderate, and intermediate dependence, respectively. For each fitted model, we computed the log-likelihood and two standard information criteria used for model selection:

$$\text{AIC} = -2\ell(\widehat{\boldsymbol{\theta}}) + 2p \quad \text{and} \quad \text{BIC} = -2\ell(\widehat{\boldsymbol{\theta}}) + p\log(n),$$

where $\ell(\widehat{\boldsymbol{\theta}})$ is the maximized log-likelihood and p denotes the number of estimated parameters. The results for the sub-scenario 2 are presented in Table 1, while the outcomes for sub-scenarios 1 and 3 are provided in the supplementary material attached to this manuscript. Figure 5 displays the surface and contour plots corresponding to a random sample generated under sub-scenario 2. The figure illustrates the joint distribution of (y_1, y_2) and reveals that the simulated observations are concentrated near the corners $(0,0)$, $(0,1)$, $(1,0)$, and $(1,1)$, that is, around the extreme regions of the unit square. This pattern reflects the combined effect of moderate dependence ($\lambda = 0.5$) and the asymmetric dispersion parameters $(\sigma_1^2, \sigma_2^2) = (\sqrt{50}, \sqrt{25})$, which together produce heavier concentration at the boundaries.

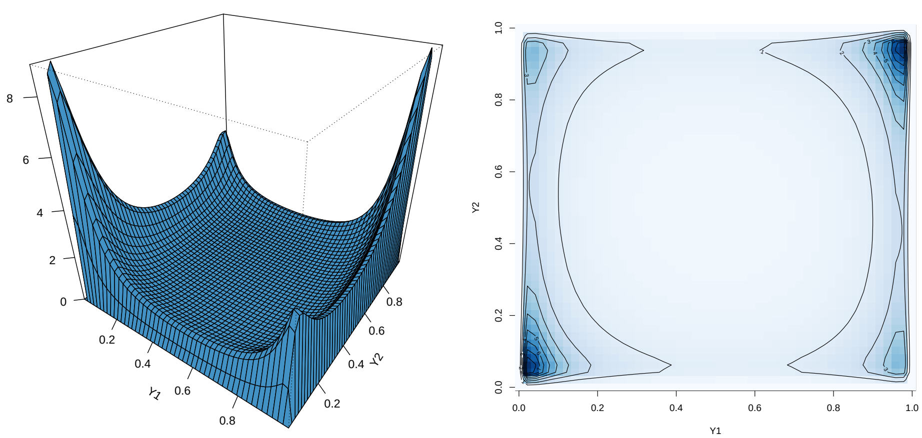

FIGURE 5: Surface and contour graphs for $\boldsymbol{\theta}_2 = (0.5, 0.5, \sqrt{50}, \sqrt{25}, 0.5)^\top$, sub-scenario 2.

Table 1: Simulation results for model comparison in sub-scenario 2, $\boldsymbol{\theta}_2 = (0.5, 0.5, \sqrt{50}, \sqrt{25}, 0.5)^\top$, with $R = 1000$ replicates.

n	Model	Loglik	AIC	BIC	Correlation
50	Simplex-FGM	21.376	-32.751	-23.191	0.187
	Simplex-Frank	21.347	-32.695	-23.134	
	Simplex-Gaussian	21.347	-32.695	-23.134	
	Simplex-Clayton	21.165	-32.331	-22.771	
	Simplex-Gumbel	20.192	-30.383	-20.823	
	Beta-Clayton	7.183	-4.366	5.194	
	Beta-Gaussian	7.070	-4.140	5.421	
	Beta-Frank	7.054	-4.108	5.452	
	Beta-FGM	7.036	-4.072	5.488	
	Beta-Gumbel	5.907	-1.815	7.745	
100	Simplex-FGM	40.016	-70.033	-57.007	0.184
	Simplex-Frank	39.958	-69.915	-56.889	
	Simplex-Gaussian	39.826	-69.652	-56.626	
	Simplex-Clayton	39.311	-68.622	-55.596	
	Simplex-Gumbel	37.829	-65.657	-52.631	
	Beta-Clayton	12.673	-15.347	-2.321	
	Beta-Gaussian	12.607	-15.215	-2.189	
	Beta-Frank	12.602	-15.203	-2.177	
	Beta-FGM	12.572	-15.144	-2.118	
	Beta-Gumbel	10.449	-10.897	2.128	
150	Simplex-FGM	58.633	-107.266	-92.213	0.177
	Simplex-Frank	58.585	-107.169	-92.116	
	Simplex-Gaussian	58.440	-106.880	-91.827	
	Simplex-Clayton	57.721	-105.441	-90.388	
	Simplex-Gumbel	55.682	-101.365	-86.311	
	Beta-Clayton	17.913	-25.826	-10.773	
	Beta-Gaussian	17.817	-25.634	-10.581	
	Beta-Frank	17.783	-25.566	-10.512	
	Beta-FGM	17.732	-25.463	-10.410	
	Beta-Gumbel	14.894	-19.788	-4.735	
200	Simplex-FGM	77.706	-145.412	-128.921	0.176
	Simplex-Frank	77.662	-145.324	-128.833	
	Simplex-Gaussian	77.419	-144.838	-128.346	
	Simplex-Clayton	76.456	-142.911	-126.420	
	Simplex-Gumbel	73.946	-137.892	-121.400	
	Beta-Clayton	23.591	-37.183	-20.691	
	Beta-Gaussian	23.483	-36.966	-20.475	
	Beta-Frank	23.471	-36.943	-20.451	
	Beta-FGM	23.385	-36.771	-20.279	
	Beta-Gumbel	19.737	-29.474	-12.983	
500	Simplex-FGM	187.789	-365.578	-344.505	0.170
	Simplex-Frank	187.710	-365.420	-344.347	
	Simplex-Gaussian	187.044	-364.088	-343.015	
	Simplex-Clayton	184.746	-359.493	-338.420	
	Simplex-Gumbel	179.678	-349.357	-328.284	
	Beta-Frank	55.116	-100.231	-79.158	
	Beta-Gaussian	55.109	-100.218	-79.145	
	Beta-Clayton	55.019	-100.039	-78.966	
	Beta-FGM	54.940	-99.880	-78.807	
	Beta-Gumbel	47.012	-84.025	-62.952	

The results in Table 1 indicate that the bivariate Simplex model with FGM copula consistently achieved the lowest AIC and BIC values in sub-scenario 2, confirming its superior fit when both the marginal and dependence structures are correctly specified. Similar patterns were observed for sub-scenarios 1 and 3, in which the Simplex-FGM model also outperformed the competing models. In contrast, when the same data were fitted using bivariate Beta models, all copula combinations yielded substantially higher AIC and BIC values, revealing a poorer fit under marginal misspecification.

5. Applications

5.1. Global Trends in Mental Health Disorder Data

In this section, we analyze a real-world dataset that illustrates the practical relevance and flexibility of the proposed bivariate Simplex distribution. The dataset, publicly available on the Kaggle platform under the title “Global Trends in Mental Health Disorder", compiles information from countries around the world on the prevalence of several mental health disorders, including schizophrenia, bipolar disorder, eating disorders, anxiety disorders, drug use disorders, depression, and alcohol use disorders. By making such data accessible, the source aims to foster understanding of how these conditions affect populations globally and to support policy decisions, prevention strategies, and resource allocation.

For our analysis, we focus on two outcomes related to the 2017 data: y_1 corresponds to the prevalence of alcohol use disorders (in percentage) in the same year, while y_2 represents the prevalence of depression (in percentage) in each country or region. Both variables are continuous and naturally restricted to the unit interval $(0, 1)$ when expressed as proportions, making them suitable for modeling under the proposed bivariate Simplex framework. This dataset provides an excellent opportunity to evaluate the empirical performance of the proposed model relative to the traditional bivariate Beta distribution. Figure 8 presents the boxplots and histograms for the variables y_1 and y_2. The graph reveals the presence of outliers in the Alcohol use disorders, which correspond to the Belarus (#19); Eastern Europe (#62); Russia (#168); Estonia (#70); Ukraine (#215); United States (#218); Latvia (#113); Moldova (#135); Central Europe, Eastern Europe, and Central Asia (#41); Mongolia (#13); Kazakhstan (#106) and Scotland (#175), in the depression, which correspond to the Greenland (#81), Lesotho (#115) and Morocco (#138).

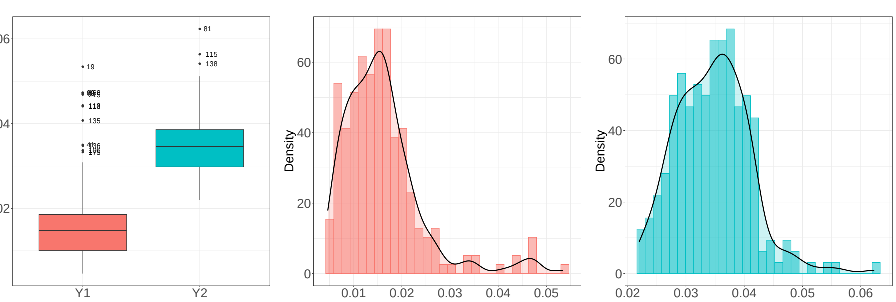

Figure 6: Boxplot and histograms for the y_1(Percentage of people with alcohol use disorders) and y_2 (Percentage of people with depression) variables.

Table 2 presents descriptive measures of position, dispersion, and shape for the variables y_1 (percentage of people with alcohol use disorders) and y_2 (percentage of people with depression). The average prevalence of depression across countries is approximately 3.4%, with values ranging from 2.2% to 6.2%. For alcohol use disorders, the mean prevalence is about 1.6%, ranging from 0.4% to 5.3%. Both variables exhibit moderate dispersion relative to their means, as indicated by the standard errors, and display positive skewness (0.656 for y_1 and 1.860 for y_2), suggesting that most countries present low to medium prevalence rates, while a few exhibit substantially higher values. The high kurtosis values (4.495 for y_1 and 7.746 for y_2) indicate leptokurtic behavior, meaning that the distributions are more peaked than the normal distribution, with heavy tails. Overall, these summary statistics confirm that both variables are asymmetric and right-skewed, with variability concentrated in the lower range of the unit interval. Such characteristics reinforce the suitability of the proposed bivariate Simplex model, which is particularly appropriate for modeling bounded, positively skewed data on $(0, 1)$.

Table 2: Descriptive measures of position, dispersion, asymmetry, kurtosis, and relative position of variables y_1 and y_2.

	Min.	Max.	Q_1	Q_3	Mean	Median	Standard Errors	Asymmetry	Kurtosis
y_1	0.004	0.053	0.010	0.018	0.016	0.015	0.004	1.860	7.746
y_2	0.022	0.062	0.029	0.038	0.034	0.035	0.006	0.656	4.495

Table 3 presents the parameter estimates, standard errors, and 95% confidence intervals obtained from the fitted bivariate Simplex and bivariate Beta models using the FGM copula. For both models, the estimated means $\widehat{\mu}_1 = 0.0159$ and $\widehat{\mu}_2 = 0.0345$ indicate that, on average, about 3.5% of the population suffers from depression and 1.6% from alcohol use disorders, respectively. The dispersion parameters $\widehat{\sigma}_1^2 = 4.05$ and $\widehat{\sigma}_2^2 = 1.03$ in the Simplex model suggest higher variability in alcohol use disorders compared to depression rates across countries. The estimated dependence parameter $\widehat{\lambda} = -0.458$ reveals a moderate negative association between alcohol use disorders and depression, implying that countries with higher prevalence of one condition tend to have lower prevalence of the other. When comparing the competing models, the bivariate Simplex–FGM model outperforms the bivariate Beta–FGM model, achieving higher log-likelihood (1679.97 versus 1672.01) and lower AIC (−3349.94 versus −3334.02) and BIC (−3332.73 versus −3316.81) values. These results provide strong empirical evidence of the superior fit and flexibility of the proposed model for describing bounded, asymmetric, and correlated data.

Figure 7 displays the estimated joint density surface and the corresponding contour plot of the fitted bivariate Simplex-FGM model, obtained using the maximum likelihood estimates reported in Table 3. The plots provide a visual representation of the dependence structure between alcohol use disorders (y_1) and depression (y_2). The shape of the fitted surface is consistent with the empirical distribution of the data, concentrating most of the probability mass in regions corresponding to low-to-moderate prevalence levels and reflecting the negative association captured by the estimate of λ. These graphical results, together with the information criteria, support the adequacy of the proposed bivariate Simplex model for describing the joint behavior of the two mental health indicators.

Table 3: Simplex and Beta Models:Estimates, standard errors, and confidence intervals with a confidence coefficient of 95%.

Parameter	Metric	Simplex	Beta
	Estimate	0.0159	0.0159
μ_1	Standard Error	0.0005	0.0005
	CI95%	(0.0149 ; 0.0169)	(0.0149 ; 0.0168)
	Estimate	0.0345	0.0344
μ_2	Standard Error	0.0004	0.0004
	CI95%	(0.0337 ; 0.0353)	(0.0336 ; 0.0352)
	Estimate	4.0489	0.0593
σ_1^2	Standard Error	0.1879	0.0029
	CI95%	(3.6807 ; 4.4172)	(0.0539 ; 0.0648)
	Estimate	1.0322	0.0340
σ_2^2	Standard Error	0.0480	0.0015
	CI95%	(0.9382 ; 1.1263)	(0.0309 ; 0.0371)
	Estimate	−0.4577	−0.4573
λ	Standard Error	0.1930	0.1965
	CI95%	(−0.8360 ; −0.0794)	(−0.8426 ; −0.0720)
logLik		1679.969	1672.009
AIC		−3349.938	−3334.017
BIC		−3332.726	−3316.805

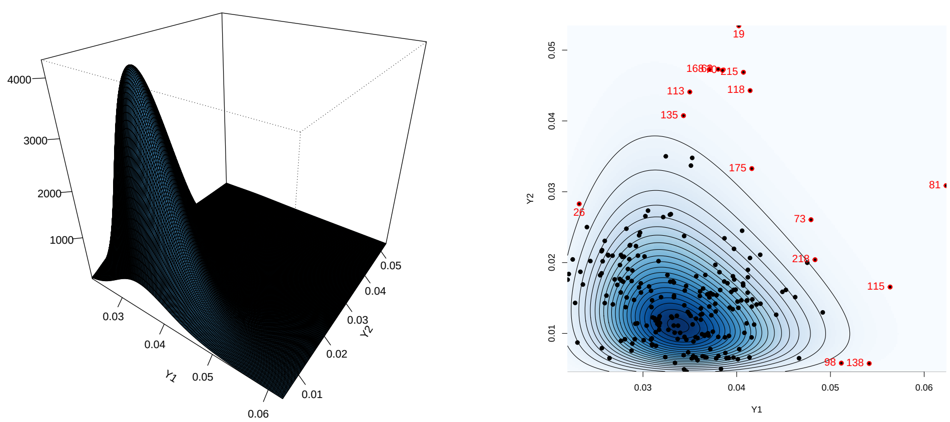

Figure 7: Estimated joint density surface (left) and contour plot (right) of the fitted bivariate Simplex-FGM model for depression and alcohol use disorders.

5.2. Jurimetric Data

In this section, we illustrate the practical applicability of the proposed bivariate Simplex distribution via the FGM copula using a real dataset from the Jurimetry field. The aim of this analysis is not to compare alternative models but rather to demonstrate that the proposed model provides an adequate and interpretable representation for bounded bivariate data arising in applied contexts. The dataset was obtained from the 2014 Annual Report of the Regional Labor Court of the 5th Region (TRT5) in Bahia, Brazil. The mission of TRT5 is to promote justice in labor relations with efficiency, transparency, and swiftness, contributing to social harmony and strengthening citizenship within Bahia. For this study, the outcomes are defined as follows: y_1, the Congestion Rate, represents the proportion of unresolved cases relative to the total cases processed within a year; and y_2, the Conciliation Index, corresponds to the percentage of sentences and decisions resolved through agreements relative to the total number of final decisions issued by the 88 Labor Courts in Bahia. Figure 8 presents the boxplots and histograms for the variables y_1 (Congestion Rate) and y_2 (Conciliation Index). The graph reveals the presence of outliers in the Conciliation Index, which correspond to the courts located in the cities of Itamaraju (#33), Simões Filho (#84), Santo Amaro (#80), and Candeias (#10 and #11).

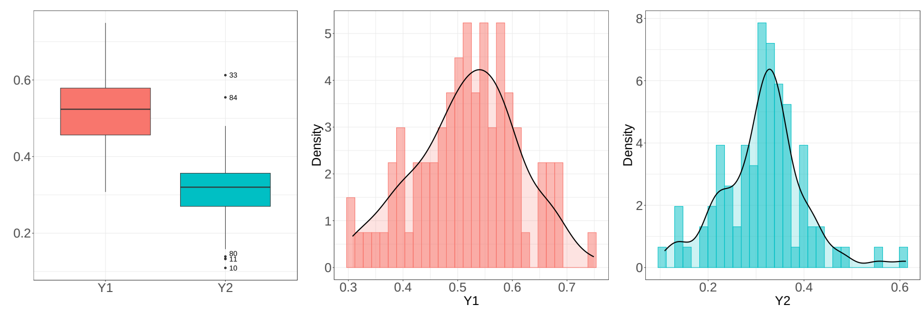

FIGURE 8: Boxplot and histograms for the y_1(Congestion Rate) and y_2 (Conciliation Index) variables.

Table 4 highlights the central tendency and dispersion measures for the variables.

TABLE 4: Descriptive measures of position, dispersion, asymmetry, kurtosis, and relative position of variables y_1 and y_2.

	Min.	Max.	Q_1	Q_3	Mean	Median	Standard Errors	Asymmetry	Kurtosis
y_1	0.308	0.749	0.456	0.579	0.524	0.524	0.095	-0.160	-0.382
y_2	0.109	0.613	0.270	0.357	0.315	0.320	0.086	0.288	1.273

As shown in Table 4, the average Conciliation Index is notably lower than the Congestion Rate, despite conciliation being the fastest route to a resolution. Given that both y_1 and y_2 are constrained to the (0,1) interval, we propose using a model that accounts for this characteristic. Accordingly, we applied the bivariate Simplex distribution defined in Section 3 to obtain the estimates, standard errors, and confidence intervals, which are reported in Table 5. The estimate for the joint expectation, $\widehat{E}(y_1y_2)$, is 0.162, and the maximum likelihood estimators are obtained when the dependence parameter $\lambda = 0.077$ (positive dependence). This suggests that the Congestion Rate and Conciliation Index tend to increase together.

Table 5: Estimates, standard errors, and confidence intervals with a confidence coefficient of 95%.

Parameter	Estimate	Standard error	Confidence Interval (95%)
μ_1	0.5178	0.0100	(0.4982 ; 0.5375)
μ_2	0.3132	0.0096	(0.2943 ; 0.3321)
σ_1^2	0.7987	0.0602	(0.6807 ; 0.9167)
σ_2^2	0.9498	0.0716	(0.8094 ; 1.0802)
AIC	-333.4668		
BIC	-321.0802		

Figure 9 displays the surface and contour plots based on the estimates obtained. The contour graph, in particular, demonstrates a good fit of the model to the data.

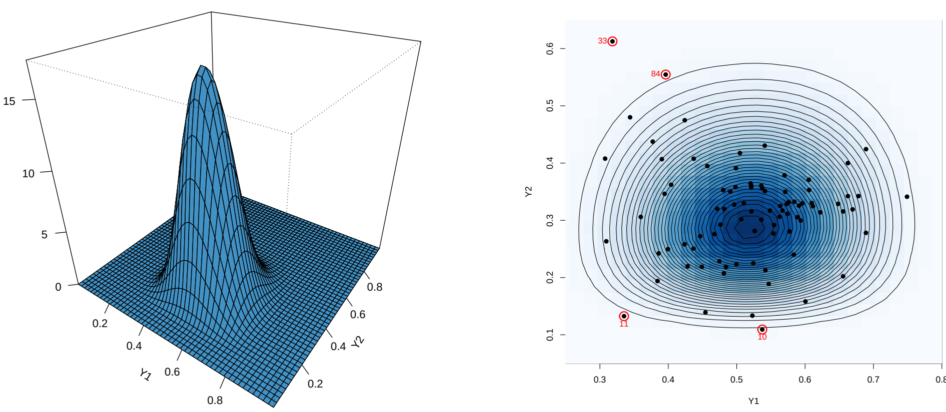

Figure 9: Congestion Rate surface graph (left) and contour graph (right) (y_1) and the Conciliation Index (y_2).

6. Conclusions

This article introduced a bivariate Simplex distribution constructed via the FGM copula as a flexible alternative for modeling continuous data constrained to the unit interval $(0, 1)$. The proposed formulation allows joint inference on two proportion-type variables and retains analytical tractability, leading to closed-form expressions for the joint density, likelihood function, and the joint expectation $\mathrm{E}(Y_1 Y_2)$. Through an extensive Monte Carlo study, we examined the asymptotic properties of the maximum likelihood estimators under different dependence structures. The results demonstrated that the estimators are consistent and asymptotically unbiased across scenarios, with the dependence parameter λ taking values of 1, -1, and 0. Although the FGM copula captures only weak dependence ($[-1/3,\ 1/3]$), its adoption in this work is justified by its analytical simplicity, symmetry, and interpretability. These features facilitate explicit derivations and provide a foundation for further extensions to copulas capable of modeling stronger or asymmetric dependence. To assess robustness, we also compared the proposed model with alternatives based on Frank, Gaussian, Clayton, and Gumbel copulas. The results confirmed the stability of the estimation procedure and the potential for extending the framework to richer dependence structures. Finally, the proposed methodology was applied to two empirical datasets drawn from the fields of Psychometry and Jurimetry: (i) global trends in mental health disorders worldwide, and (ii) the Annual Report of the Regional Labor Court of the 5^{th} Region (TRT5) in Bahia, Brazil. The empirical analyses corroborated the practical suitability of the proposed model, underscoring its effectiveness in modeling bounded bivariate outcomes and its potential applicability to other research domains involving proportions and rate-based data.

Acknowledgments

The first and second authors gratefully acknowledge the undergraduate and graduate scholarship provided by Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq and Fundação de Amparo à Pesquisa do Estado da Bahia – Fapesb, respectively. Also, the last author gratefully acknowledges the funding provided by Programa Institucional de Internacionalização – CAPES – PrInt – (Grant No. 88887.935790/2024-00). Vanessa Barros is partially supported by FAPEMIG APQ-01341-21 project.

[Received: February 2025 — Accepted: December 2025]

References

Andreopoulos, P., Bersimis, G. F., Tragaki, A. & Rovolis, A. (2019), 'Mortality modeling using probability distributions. application in greek mortality data', *Communications in Statistics-Theory and Methods* **48**, 127–140.

Arnold, B. C. & Ng, H. K. T. (2011), 'Flexible bivariate beta distributions', *Journal of Multivariate Analysis* **102**, 1194–1202.

Barndorff-Nielsen, O. E. & Jørgensen, B. (1991), 'Some parametric models on the simplex', *Journal of multivariate analysis* **39**, 106–116.

Barros, O. A. d. (2015), Estimação dos parâmetros da distribuição beta bivariada: aplicações em severidade de doenças em plantas, PhD thesis, University of São Paulo, Brazil.

Biswas, A. & Chakraborty, S. (2019), '$r = p(y \leq x)$ for unit-Lindley distribution: inference with an application in public health', *https://arxiv.org/abs/1904.06181* .

Biswas, A., Chakraborty, S. & Mukherjee, M. (2021), 'On estimation of stress–strength reliability with log-Lindley distribution', *Journal of Statistical Computation and Simulation* **91**, 128–150.

Durante, F. & Sempi, C. (2016), *Principles of Copula Theory*, CRC Press, Taylor & Francis Group, Boca Raton, FL.

Gómez-Déniz, E., Sordo, M. A. & Calderín-Ojeda, E. (2014), 'The log–Lindley distribution as an alternative to the beta regression model with applications in insurance', *Insurance: Mathematics and Economics* **54**, 49–57.

Gupta, A. K. & Wong, C. (1985), 'On three and five parameter bivariate beta distributions', *Metrika* **32**, 85–91.

Hoeffding, W. (1940), 'Masstabinvariante korrelationstheorie', *Schriften des Mathematischen Instituts und Instituts fur Angewandte Mathematik der Universitat Berlin* **5**, 181–233.

Johnson, M. E. (1987), *Multivariate Statistical Simulation*, John Wiley & Sons, New York.

Jones, M. C. (2002), 'Multivariate t and beta distributions associated with the multivariate F distribution', *Metrika* **54**, 215–231.

Jorgensen, B. (1997), *The theory of dispersion models*, CRC Press.

Kotz, S., Balakrishnan, N. & Johnson, N. L. (2019), *Continuous multivariate distributions: Models and applications*, John Wiley and Sons.

Kropáč, O. (1982), 'Some properties and applications of probability distributions based on McDonald function', *Aplikace matematiky* **27**, 285–302.

Lai, C. D. & Balakrishnan, N. (2009), *Continuous bivariate distributions*, Springer.

Machado Moschen, L. & Carvalho, L. M. (2023), 'Bivariate beta distribution: parameter inference and diagnostics', *https://arxiv.org/abs/2303.01271*.

Mazucheli, J., Menezes, A. F. B. & Chakraborty, S. (2019), 'On the one parameter unit-Lindley distribution and its associated regression model for proportion data', *Journal of Applied Statistics* **46**, 700–714.

Nadarajah, S. & Kotz, S. (2005), 'Some bivariate beta distributions', *Statistics* **39**, 457–466.

Nelder, J. A. & Wedderburn, R. W. (1972), 'Generalized linear models', *Journal of the Royal Statistical Society: Series A (General)* **135**, 370–384.

Olkin, I. & Liu, R. (2003), 'A bivariate beta distribution', *Statistics and Probability Letters* **62**, 407–412.

Olkin, I. & Trikalinos, T. A. (2015), 'Constructions for a bivariate beta distribution', *Statistics and Probability Letters* **96**, 54–60.

Olver, F. W. J., Lozier, D. W., Boisvert, R. F. & Clark, C. W. (2010), *NIST handbook of mathematical functions*, Cambridge University Press.

Sarabia, J. M. & Castillo, E. (2006), 'Bivariate distributions based on the generalized three-parameter beta distribution', *Advances in distribution theory, order statistics, and inference* **31**, 85–110.

Sklar, M. (1959), 'Fonctions de répartition a n dimensions et leurs marges', *I'nstitut Statistique de l'Université de Paris* **8**, 229–231.

Song, P. X.-K. & Tan, M. (2000), 'Marginal models for longitudinal continuous proportional data', *Biometrics* **56**, 496–502.

Appendix A. Mathematical Results used in the Proof of Theorem 1

Let K_ν be the modified Bessel function of the second kind. From the definition, we know that for ν real and z positive, $K_\nu(z)$ is real. Additionally, the following symmetry holds $K_{-\nu}(z) = K_\nu(z)$. Below we state some properties of this function.

Proposition 1. *Let β and γ be positive real numbers, and let $\nu \in \mathbb{R}$. Then we have the following integral representation for the modified Bessel function of the second kind:*

$$\int_0^\infty x^{\nu-1} e^{-\frac{1}{2}(\beta x^{-1}+\gamma x)} dx = 2(\frac{\beta}{\gamma})^{\nu/2} K_\nu(\sqrt{\beta\gamma}).$$

Proof. See Kropáč (1982). □

Proposition 2. *Let $z \in \mathbb{C}$ such that the real part of z, $\Re(z)$, is positive. Then,*

1. $\int_0^\infty K_1(2z\cosh(t))dt = \frac{1}{2}K_{1/2}^2(z)$
2. $\int_0^\infty K_0(2z\cosh(t))\cosh(t)dt = \frac{1}{2}K_{1/2}^2(z)$.

Proof. See expression 10.32.17 (`https://dlmf.nist.gov/10.32`). □

Corollary 1. *Let $a \in \mathbb{C}$ be such that $\Re(a) > 0$. Then*

$$\int_1^\infty \frac{1}{x} K_1\Big(a\frac{x^2+1}{x}\Big)dx = \frac{\pi}{4a}e^{-2a}.$$

Proof. Considering $x = e^t$ we have

$$\int_1^\infty \frac{1}{x} K_1\Big(a\frac{x^2+1}{x}\Big)dx = \int_0^\infty K_1(2a\cosh(t))dt.$$

Using Proposition 2 with $z = a$ and the fact that $K_{1/2}(a) = \sqrt{\frac{\pi}{2a}}e^{-a}$ we obtain the result. □

Additionally, we introduce the notion of asymptotic equivalence and little-o notation:

Definition 1. Two functions f and g are asymptotic equivalents as x approaches a if

$$\lim_{x \to a} \frac{f(x)}{g(x)} = 1.$$

This relationship is denoted by $f \sim_a g$.

We observe that the asymptotic equivalence relation is transitive, that is, if $f \sim_a g$ and $g \sim_a h$, then $f \sim_a h$.

Definition 2. A function $f(x)$ is $o(g(x))$ for $x \to a$ if $f(x)$ grows slower than $g(x)$ as x approaches a. In simpler terms,

$$\lim_{x \to a} \frac{f(x)}{g(x)} = 0.$$

From now on, we denote L_ν as the modified Struve function, and $\text{ph}(z)$ denotes the phase of the complex number z.

Lemma 1. *Let $z \in \mathbb{C}$ such that $|ph(z)| < \frac{\pi}{2}$, and let $\nu \in \mathbb{R}$ such that $\nu \pm \frac{1}{2} \notin -\mathbb{N}$. Then the following asymptotic expansion holds for $|z| \gg 1$:*

$$zK_{\nu+1}(z)L_\nu(z) \sim c_\nu \sqrt{\frac{z\pi}{2}} z^{\nu-1} e^{-z} + \frac{1}{2}. \tag{A1}$$

***Proof*.** In this proof, we refer to Olver et al. (2010, p.249, 252, 288, 293) for know results used. First we observe that $L_\nu(z) = M_\nu(z) + I_\nu(z)$. Also, the following asymptotic equivalences hold

$$K_\nu(z) \sim \sqrt{\frac{\pi}{2z}} e^{-z},$$

$$I_\nu(z) \sim \sqrt{\frac{1}{2\pi z}} e^{z},$$

and

$$M_\nu(z) \sim \frac{1}{\pi} \sum_{k \geq 0} c_\nu^k z^{\nu-2k-1},$$

where M_ν is the modified Struve function, I_ν is the modified Bessel function and $c_\nu^k = \frac{(-1)^{k+1}\Gamma(k+1/2)(1/2)^{\nu-2k-1}}{\Gamma(\nu+1/2-k)}$. Since $\sum_{k\geq 1} c_\nu^k z^{\nu-2k-1} = o(c_\nu^0 z^{\nu-1})$, then $\sum_{k\geq 0} c_\nu^k z^{\nu-2k-1} \sim c_\nu^0 z^{\nu-1}$ and by transitivity $M_\nu(z) \sim \frac{1}{\pi} c_\nu^0 z^{\nu-1}$. Therefore,

$$L_\nu(z) = M_\nu(z) + I_\nu(z) \sim \frac{1}{\pi} c_\nu^0 z^{\nu-1} + \sqrt{\frac{1}{2\pi z}} e^{z}.$$

Finally,

$$zK_{\nu+1}(z)L_\nu(z) \sim c_\nu^0 \sqrt{\frac{1}{2\pi}} z^{\nu-1/2} e^{-z} + \frac{1}{2} \sim \frac{1}{2}.$$

□

Corollary 2.

$$\lim_{x->+\infty} x\Big(K_0(x)L_{-1}(x)+K_1(x)L_0(x)\Big)=1.$$

Proof. This proof is straightforward from Proposition 1. □

Proposition 3. *Let $z\in\mathbb{C}$. Then:*

$$\int K_0(z)\,dz=\frac{\pi}{2}z(K_0(z)L_{-1}(z)+K_1(z)L_0(z))+C.$$

Proof. See expression 10.43.2 (`https://dlmf.nist.gov/10.43`) □

Lemma 2. *For $a>0$, we have*

$$\int_1^\infty K_0\Big(a\frac{x^2+1}{x}\Big)dx=\frac{\pi}{4a}e^{-2a}+\frac{\pi}{2}\Big[\frac{1}{2}-K_0(2a)L_{-1}(2a)-K_1(2a)L_0(2a)\Big].$$

Proof. Let $x=e^t$. Then

$$\begin{aligned}\int_1^\infty K_0\Big(a\frac{x^2+1}{x}\Big)dx &= \int_0^\infty e^tK_0(2a\cosh(t))dt,\\ &= \int_0^\infty \cosh(t)K_0(2a\cosh(t))dt+\\ &\quad \int_0^\infty \sinh(t)K_0(2a\cosh(t))dt.\end{aligned}$$

Using Proposition 2 with $z=a$ and the fact that $K_{1/2}(a)=\sqrt{\frac{\pi}{2a}}e^{-a}$ we find that

$$\int_0^\infty \cosh(t)K_0(2a\cosh(t))dt=\frac{\pi e^{-2a}}{4a}.$$

For the remaining integral, let $y=2a\cosh(t)$, yielding

$$\int_0^\infty \sinh(t)K_0\Big(2a\cosh(t)\Big)dt=\frac{1}{2a}\int_{2a}^\infty K_0(y)dy.$$

To compute the last integral, simply apply Proposition 3 and Corollary 2. □

Proposition 4. *Let erf denote the error function given by $erf(z)=\frac{2}{\sqrt{\pi}}\int_0^z e^{-t^2}dt$. Then*

(i) $erf'(z)=\frac{2}{\sqrt{\pi}}e^{-z^2}$

(ii) $lim_{z->+\infty}erf(z)=1$

(iii) $\int e^{-a^2x^2-\frac{b^2}{x^2}}dx=\frac{\sqrt{\pi}}{4a}[e^{2ab}erf(ax+\frac{b}{x})+e^{-2ab}erf(ax-\frac{b}{x})]+C$ *for* $z>0$ *and* $|ph(a)|<\pi/4$.

Proof. See Olver et al. (2010, Cap. 7). □

Appendix B. Simulation: Scenario 4

Let u_1 and u_2 denote the distribution functions of y_1 and y_2, respectively. Equation (3) can be rewritten as

$$f(y_1, y_2; \boldsymbol{\theta}) = f(y_1; \mu_1, \sigma_1^2) \times f(y_2; \mu_2, \sigma_2^2) \times c(u_1, u_2),$$

where $f(y_1; \mu_1, \sigma_1^2)$ and $f(y_2; \mu_2, \sigma_2^2)$ are the marginal density functions of y_1 and y_2, respectively, and $c(u_1, u_2)$ denotes the copula density function that captures the dependence structure between them. The following copulas are considered in this work: Clayton, Frank, Gumbel, and Gaussian.

1. **Clayton Copula:** The bivariate Clayton copula Durante & Sempi (2016) is defined as

$$c(u_1, u_2; \lambda) = \max\left\{\left(u_1^{-\lambda} + u_2^{-\lambda} - 1\right)^{-1/\lambda}, 0\right\},$$

   where the dependence parameter λ belongs to the interval $[-1, \infty)$, $\lambda \neq 0$. The case $\lambda \to 0$ corresponds to the independence copula.

2. **Frank Copula:** The bivariate Frank copula Durante & Sempi (2016) is given by

$$c(u_1, u_2; \lambda) = -\frac{1}{\lambda} \ln\left(1 + \frac{(\exp\{-\lambda u_1\} - 1)(\exp\{-\lambda u_2\} - 1)}{\exp\{-\lambda\} - 1}\right),$$

   where $\lambda \in \mathbb{R}$, $\lambda \neq 0$. As $\lambda \to 0$, the copula converges to the independence copula.

3. **Gumbel Copula:** The bivariate Gumbel copula Durante & Sempi (2016) is defined as

$$c(u_1, u_2; \lambda) = \exp\left(-\left[(-\ln u_1)^\lambda + (-\ln u_2)^\lambda\right]^{1/\lambda}\right),$$

   where $\lambda \in [1, \infty)$. The case $\lambda = 1$ corresponds to the independence copula.

4. **Gaussian Copula:** The bivariate Gaussian copula Durante & Sempi (2016) is expressed as

$$c(u_1, u_2; \lambda) = \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi\sqrt{1 - \lambda^2}} \exp\left(-\frac{s^2 - 2\lambda st + t^2}{2(1 - \lambda^2)}\right) dt\, ds,$$

   where $\Phi^{-1}(\cdot)$ denotes the quantile function (inverse CDF) of the standard normal distribution, and $\lambda \in (-1, 1)$ is the linear correlation coefficient.