Electronic Research Archive

http://www.aimspress.com/journal/era

ERA, 33(8): 4468–4494.
DOI: 10.3934/era.2025203
Received: 24 June 2025
Revised: 13 July 2025
Accepted: 21 July 2025
Published: 06 August 2025

Research article

Copula-based analysis of asymmetrically distributed joint data under a risk profile using the discrete Type-I extreme value distribution

Hend S. Shahen[1], Mahmoud El-Morshedy[1,*], Mohamed S. Eliwa[2] and Mohamed F. Abouelenein[3]

[1] Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia

[2] Department of Statistics and Operations Research, College of Science, Qassim University, Saudi Arabia

[3] Department of Insurance and Risk Management, College of Business, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11432, Riyadh, Saudi Arabia

* **Correspondence:** Email: m.elmorshedy@psau.edu.sa.

Abstract: Copulas provide a flexible framework for building bivariate probability models that reflect specific dependency patterns. This work introduces a discrete form of the Type-I extreme value (Gumbel) distribution within a copula-based structure. Key mathematical and statistical characteristics are examined, including the joint probability mass function, survival function, hazard rate, conditional expectation, joint probability generating function, and dependency properties such as positive quadrant dependence and total positivity of order two. The bivariate discrete Gumbel model demonstrates strong performance in handling asymmetric data and proves particularly useful for capturing extreme and outlier observations. Its joint hazard rate function adds further flexibility, making it suitable for modeling a range of failure rate behaviors. Parameter estimation is carried out using the maximum likelihood method, and a thorough simulation study evaluates the bias and mean squared errors across various sample sizes. To illustrate its practical relevance, the model is applied to three different real-world datasets: football match outcomes, nasal drainage severity scores, and lens defects involving surface and interior faults.

Keywords: statistical model; copula approach; extreme value distribution; failure analysis; conditional expectation; simulation; data-driven approaches; decision-making

1. Introduction

Discrete probability distributions are essential tools in mathematics, offering practical ways to understand and analyze real-world situations. These models help describe how certain events happen, whether they involve the movement of populations, the timing of arrivals, or the behavior of financial markets. Their impact spans across multiple disciplines such as business, finance, sports analytics, healthcare, and economics, where they are often used to make informed decisions. For example, the binomial distribution is frequently applied in finance to estimate the chances of a particular outcome, while the Poisson distribution is commonly used to model the number of events occurring in a fixed period or space, such as customers arriving at a service counter. In engineering, these distributions help assess the risk of system failures, allowing designers to build more dependable and safer systems. Because of their versatility, discrete distributions have drawn considerable attention in academic research, leading to the development of various models tailored to specific types of data. One such model is the discrete Gumbel (DGu) distribution, a noteworthy extension of the well-known Type-I extreme value distribution (as discussed by Kotz and Nadarajah [1]). This model stands out for its ability to capture the behavior of extreme values in discrete data, making it a valuable addition to the toolkit for analyzing events that deviate significantly from the norm. The cumulative distribution function (CDF) and its corresponding probability mass function (PMF) of the DGu model are, respectively, formulated as

$$F_{DGu}(x;\alpha,p) = e^{-\alpha p^{(x+1)}};\ x \in \mathbb{Z}, \tag{1.1}$$

and

$$P_{DGu}(x;\alpha,p) = e^{-\alpha p^{(x+1)}} - e^{-\alpha p^{x}};\ x \in \mathbb{Z}, \tag{1.2}$$

where $0 < p < 1$ and $\alpha > 0$. Chakraborty et al. [2] highlighted several important characteristics of the DGu model. In practical scenarios, it is common to encounter bivariate datasets where the two variables originate from different sources or processes. In such cases, using a simple univariate approach may not suffice, and there is a clear need for bivariate extensions of discrete distributions to capture the joint behavior of both variables effectively.

A bivariate discrete distribution is particularly useful when examining how two distinct discrete random variables interact. It provides a way to model their joint behavior and helps in estimating the probability of events on the basis of their combined outcomes. However, the number of well-established and adaptable bivariate discrete models is still relatively limited. This has motivated researchers in recent years to propose new methods for constructing such models. Techniques range from marginal-based approaches and compounding to more flexible frameworks like copulas. Among these, copula functions have gained significant attention for their ability to describe the dependence structure between variables independently of their marginal distributions. The foundational concept, known as Sklar's theorem, is discussed extensively by Boubaker and Sghaier [3], where a variety of copula families are also introduced. These include the Marshall-Olkin, Farlie-Gumbel-Morgenstern, Gaussian, and the Cuadras-Auge copulas (CACo), among others. Each offers unique flexibility in capturing different types of dependency between variables. Further examples and applications of these copula families can be found in the works of Hutchinson [4], Eliwa and El-Morshedy [5], Nelsen [6], Ning [7], Almetwally et al. [8], El-Morshedy and Eliwa [9], El-Sherpieny et al. [10], Susam [11], and Obeidat et al. [12], among others. The bivariate discrete Gumbel (BDGu) model

presented in this study is built using the CACo, providing a robust framework for constructing bivariate discrete distributions that capture intricate dependency patterns. The CACo is particularly effective in modeling asymmetric dependence and tail behavior, making it an ideal match for the BDGu distribution, which is designed to handle skewness, kurtosis, and extreme values. Together, these components form a flexible and powerful modeling tool suitable for complex real-world data in various areas. The CACo itself is part of a family of bivariate models first introduced by Cuadras and Auge [13] and is known for their ability to model non-symmetric and non-linear dependence structures, i.e.,

$$C_{\xi}(k,l) = \min(k,l)\max(k,l)^{\delta}, \tag{1.3}$$

where $\delta = 1-\xi$, $0 \le \xi \le 1$, and $0 \le k, l \le 1$. The parameter ξ measures the degree of dependence.

The BDGu model offers several valuable features that make it a versatile tool for statistical modeling. Its joint probability mass function (PMF) and survival function (SF) can both be expressed in closed-form, allowing for straightforward computation. Depending on the parameter values, the joint hazard rate function (HRF) can take on a variety of shapes such as increasing, decreasing, unimodal, or even the classic bathtub form (surface), making the BDGu model suitable for analyzing different types of failure behaviors. These flexible shapes result from specific combinations of the model's parameters. The BDGu model can also be applied in practical contexts like maintenance planning or stress-strength analysis, and it performs particularly well when modeling extreme values or outliers. In addition to its modeling flexibility, the BDGu possesses two important statistical properties. The first is positive quadrant dependence (PQD), which means that large (or small) values of one variable tend to be associated with similarly large (or small) values of the other, which is an important feature when studying interdependent systems or network behavior. The second is second-order total positivity (TP2), a property that implies a consistent, monotonic relationship between joint probabilities. This enhances both the interpretability of the model and the reliability of inferences drawn from it, particularly in contexts involving risk assessment or decision-making under uncertainty.

The article is structured as follows. Section 2 introduces the BDGu model and its marginal distributions. Section 3 discusses the main statistical properties of the model. Section 4 focuses on parameter estimation using the maximum likelihood method. In Section 5, simulation studies are conducted to assess the model's performance. Section 6 highlights the practical application of the BDGu distribution using three real datasets and also examines finite-time synchronization in Kuramoto and fractional-order networks. Finally, Section 7 provides concluding remarks and outlines potential directions for future research.

2. The BDGu distribution: Theoretical formulation

Taking the dependency structure between the two random variables X_1 and X_2 into account, the BDGu distribution can be derived by applying the CACo copula to discrete Gumbel marginals. The joint PMF of the BDGu model is then given by

$$P_{BDGu}(x_1,x_2) = \begin{cases} P_{\Omega} & \text{if } \Omega = \{(x_1,x_2) : x_1 < x_2\} \\ P_{\Lambda} & \text{if } \Lambda = \{(x_1,x_2) : x_1 > x_2\} \\ P_{\Xi} & \text{if } \Xi = \{(x_1,x_2) : x_1 = x_2 = x\} \end{cases}, \tag{2.1}$$

where

$$P_{\Omega} = P_1(x_1, x_2) = \left[e^{-\alpha p^{(x_1+1)}} - e^{-\alpha p^{x_1}}\right]\left[e^{-\alpha\delta p^{(x_2+1)}} - e^{-\alpha\delta p^{x_2}}\right]$$
$$= P_{DGu}(x_1; \alpha, p)\, P_{DGu}(x_2; \alpha\delta, p),$$

$$P_{\Lambda} = P_2(x_1, x_2) = \left[e^{-\alpha\delta p^{(x_1+1)}} - e^{-\alpha\delta p^{x_1}}\right]\left[e^{-\alpha p^{(x_2+1)}} - e^{-\alpha p^{x_2}}\right]$$
$$= P_{DGu}(x_1; \alpha\delta, p)\, P_{DGu}(x_2; \alpha, p),$$

and

$$P_{\Xi} = P_3(x) = e^{-\alpha\delta p^{(x+1)}}\left[e^{-\alpha p^{(x+1)}} - e^{-\alpha p^{x}}\right] - e^{-\alpha p^{x}}\left[e^{-\alpha\delta p^{(x+1)}} - e^{-\alpha\delta p^{x}}\right]$$
$$= e^{-\alpha\delta p^{(x+1)}} P_{DGu}(x; \alpha, p) - e^{-\alpha p^{x}} P_{DGu}(x; \alpha\delta, p).$$

Figure 1 presents plots of the joint PMF under selected parameter settings. These visualizations illustrate that the BDGu model effectively captures asymmetry in the data across the entire domain $(-\infty, \infty)$. Additionally, the joint PMF exhibits a monoclinic, unimodal surface, reflecting the model's flexibility in representing complex distributional shapes.

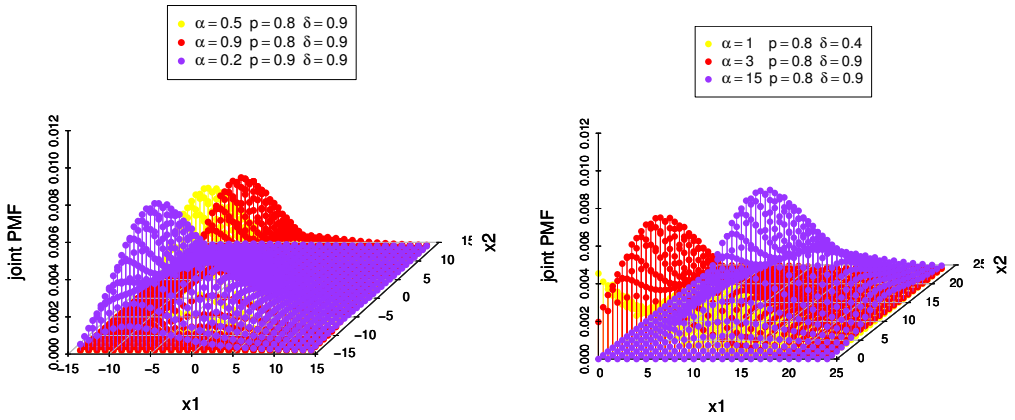

Figure 1. Schematic diagram of the joint PMF surface.

The joint CDF of the BDGu distribution can be proposed as follows:

$$F_{BDGu}(x_1, x_2) = e^{-\alpha\delta p^{(x_1+1)}} e^{-\alpha\delta p^{(x_2+1)}} e^{-\alpha(1-\delta)p^{(z+1)}};\ x_1, x_2 \in \mathbb{Z}, \tag{2.2}$$

where $z = \min\{x_1, x_2\}$. The marginal CDF of the BDGu distribution can be formulated as

$$F_{DGu}(x; \alpha, p) = e^{-\alpha p^{(x_i+1)}};\ x_i \in \mathbb{Z},\ i = 1,\ 2. \tag{2.3}$$

The corresponding joint SF of Eq (2.2) can be expressed as

$$S_{BDGu}(x_1, x_2) = \begin{cases} S_\Omega & \text{if } \Omega = \{(x_1, x_2) : x_1 < x_2\} \\ S_\Lambda & \text{if } \Lambda = \{(x_1, x_2) : x_1 > x_2\} \\ S_\Xi & \text{if } \Xi = \{(x_1, x_2) : x_1 = x_2 = x\} \end{cases}, \tag{2.4}$$

where

$$S_\Omega = S_1(x_1, x_2) = 1 - e^{-\alpha p^{(x_2+1)}} - e^{-\alpha p^{(x_1+1)}} \left[1 - e^{-\alpha\delta p^{(x_2+1)}}\right],$$

$$S_\Lambda = S_2(x_1, x_2) = 1 - e^{-\alpha p^{(x_1+1)}} - e^{-\alpha p^{(x_2+1)}} \left[1 - e^{-\alpha\delta p^{(x_1+1)}}\right],$$

and

$$S_\Xi = S_3(x) = 1 - 2e^{-\alpha p^{(x+1)}} + e^{-\alpha(\delta+1)p^{(x+1)}}.$$

Survival functions (SFs) are key statistical tools utilized to model the probability of survival over time under specific conditions. Their form varies depending on factors such as species traits and environmental influences, offering valuable insights across fields like ecology, economics, and public health. For more details, see Yang et al. [14–15]. Figure 2 displays plots of the joint SF for selected parameter values.

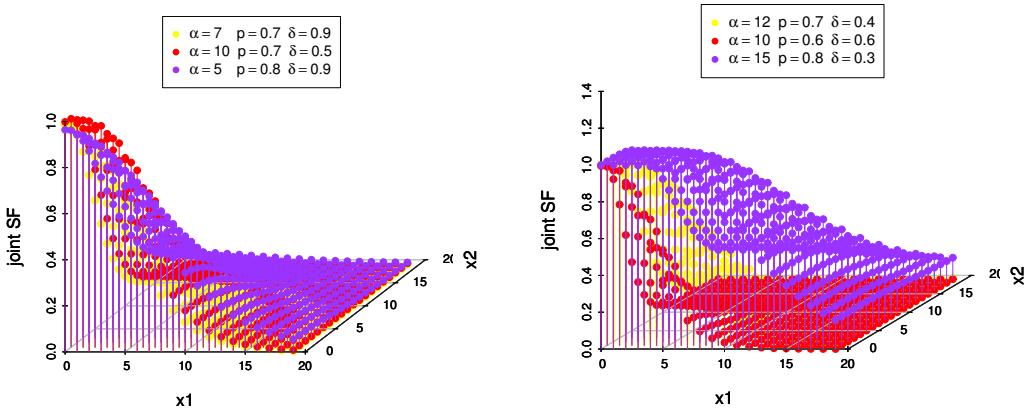

Figure 2. Schematic diagram of the joint SF surface.

The joint HRF, derived from the joint SF, measures the likelihood of an event occurring within a specific time frame, given that it has not occurred earlier. It is widely applied in probability, statistics, and risk assessment to evaluate event timing and potential outcomes. Mathematically, it is expressed as the ratio of the event's probability in a given period to the probability of its survival until that period. Figure 3 shows 3D plots of the joint HRF under selected parameter values.

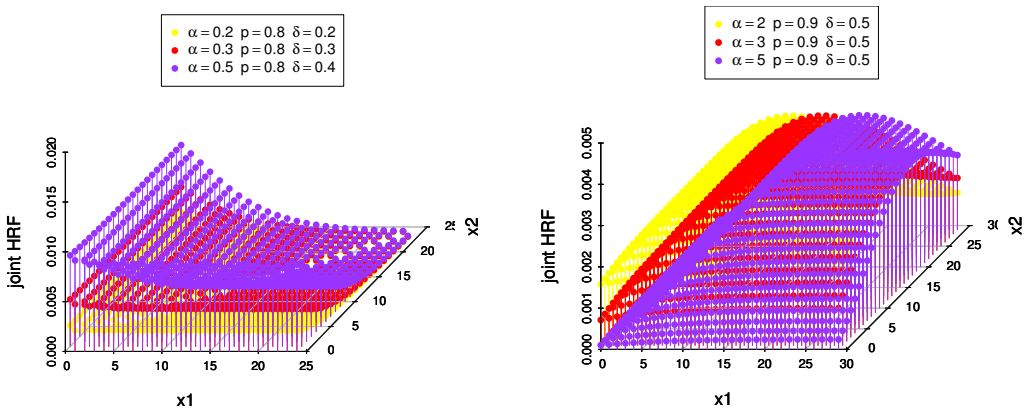

Figure 3. Schematic diagram of the joint HRF surface.

As observed, the joint HRF of the BDGu model is capable of modeling and analyzing various failure rate patterns, including increasing, decreasing, unimodal surfaces, and the classic bathtub shape.

3. Distribution properties

3.1. Positive quadrant dependence and total positivity of order two properties

Establishing such order relationships between the random variables Z_1 and Z_2 is crucial in various fields such as engineering, insurance, reliability, and finance. The variables Z_1 and Z are said to be positively quadrant dependent (PQD) if

$$F_{Z_1,Z_2}(z_1, z_2) \geq F_{Z_1}(z_1)F_{Z_2}(z_2), \tag{3.1}$$

for all z_1 and z_2. Let the bivariate random vector $\mathbf{X} = (X_1, X_2)$ has the BDGu distribution, X_1 and X_2 are positively quadrant dependent (PQD). It is easy to prove this property via Eqs (2.2) and (2.3). Since X_1 and X_2 are PQD, then for any increasing functions $F_{Z_1}(z_1)$ and $F_{Z_2}(z_2)$, we can get $Cov(F_{Z_1}(z_1), F_{Z_2}(z_2)) \geq 0$ (see, Nelsen [6]). According to the total positivity of order two (TP2) property, a discrete bivariate random vector $\mathbf{X}$ with the joint PMF $P(.,.)$ is said to have a TP2 property if

$$\frac{P(x_{11}, x_{21})\, P(x_{12}, x_{22})}{P(x_{12}, x_{21})\, P(x_{11}, x_{22})} \geq 1, \tag{3.2}$$

for $x_{11} < x_{12}$ and $x_{21} < x_{22}$. For the BDGu distribution, this model has been shown to generally satisfy the TP2 characteristic.

3.2. The conditional distribution

A conditional distribution describes the probability that an event has occurred, given that another event has already taken place. It plays a key role in statistics by improving predictive accuracy and is often estimated using Bayesian methods, which combine prior knowledge with observed data.

3.2.1. The conditional expectation

Consider that the bivariate random vector $\mathbf{X}$, which follows the BDGu distribution. The conditional PMF of X_1 given $X_2 = x_2$ can then be expressed as

$$P_{X_1|X_2=x_2}(x_1|x_2) = \begin{cases} P^*_{\Omega} & \text{if } \Omega = \{(x_1, x_2) : x_1 < x_2\} \\ P^*_{\Lambda} & \text{if } \Lambda = \{(x_1, x_2) : x_1 > x_2\} \\ P^*_{\Xi} & \text{if } \Xi = \{(x_1, x_2) : x_1 = x_2 = x\} \end{cases}, \tag{3.3}$$

where

$$P^*_{\Omega} = P_1(x_1|x_2) = \frac{P_{DGu}(x_1; \alpha, p) P_{DGu}(x_2; \alpha\delta, p)}{P_{DGu}(x_2; \alpha, p)},$$
$$P^*_{\Lambda} = P_2(x_1|x_2) = P_{DGu}(x_1; \alpha\delta, p),$$

and

$$P^*_{\Xi} = P_3(x) = \frac{e^{-\alpha\delta p^{(x+1)}} P_{DGu}(x; \alpha, p) - e^{-\alpha p^x} P_{DGu}(x; \alpha\delta, p)}{P_{DGu}(x; \alpha, p)}.$$

Similarly, the conditional PMF of X_2 given $X_1 = x_1$ can be derived. Figure 4 presents sketches of the conditional PMF for the BDGu distribution under different parameter settings, illustrating how the distribution adapts to varying dependence structures.

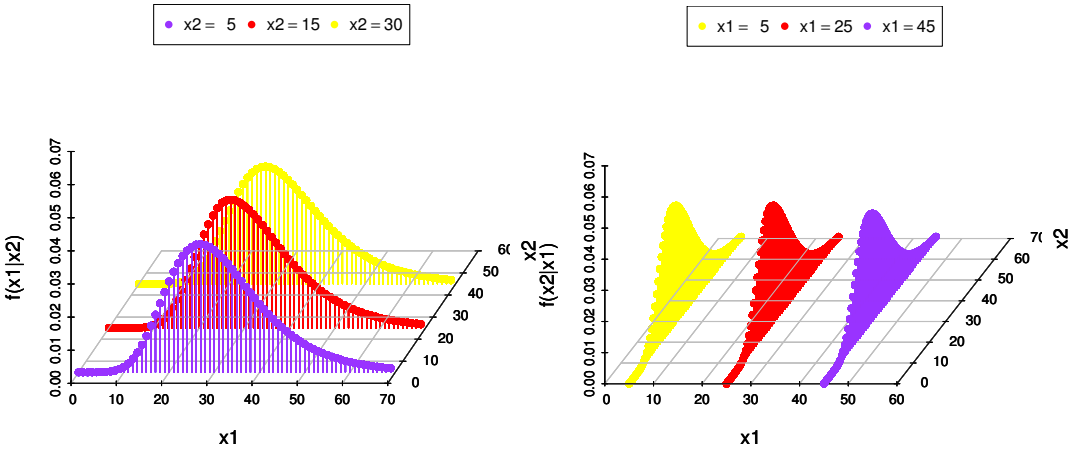

Figure 4. Some scatter plots of the conditional PMF for X_i given $X_j = x_j$; $i, j = 1, 2$.

According to the conditional PMF of X_1 given $X_2 = x_2$, the conditional expectation can be formulated as

$$\mathbf{E}(X_1|X_2 = x_2) = \sum_{x_1=0}^{x_2-1} \frac{x_1 \left[e^{-\alpha p^{(x_1+1)}} - e^{-\alpha p^{x_1}}\right]\left[e^{-\alpha\delta p^{(x_2+1)}} - e^{-\alpha\delta p^{x_2}}\right]}{\left[e^{-\alpha p^{(x_2+1)}} - e^{-\alpha p^{x_2}}\right]} + x_2 e^{-\alpha\delta p^{(x_2+1)}}$$
$$+ \sum_{x_1=x_2+1}^{\infty} x_1 \left[e^{-\alpha\delta p^{(x_1+1)}} - e^{-\alpha\delta p^{x_1}}\right] - \frac{x_2 e^{-\alpha p^{x_2}} \left[e^{-\alpha\delta p^{(x_2+1)}} - e^{-\alpha\delta p^{x_2}}\right]}{\left[e^{-\alpha p^{(x_2+1)}} - e^{-\alpha p^{x_2}}\right]}, \tag{3.4}$$

where $E(X_1|X_2 = x_2) = \sum_{x_1=0}^{\infty} x_1 P_{X_1|X_2=x_2}(x_1|x_2)$.

3.2.2. The conditional CDF

Assume that the random vector $\mathbf{X} = (X_1, X_2)$ follows the BDGu distribution. The conditional CDF of $X_1|X_2 = x_2$ is then given by

$$F_{X_1|X_2=x_2}(x_1|x_2) = \begin{cases} F_\Omega & \text{if } \Omega = \{(x_1, x_2) : x_1 < x_2\} \\ F_\Lambda & \text{if } \Lambda = \{(x_1, x_2) : x_1 > x_2\} \\ F_\Xi & \text{if } \Xi = \{(x_1, x_2) : x_1 = x_2 = x\} \end{cases}, \tag{3.5}$$

where

$$F_\Omega = F_1(x_1|x_2) = \frac{e^{-\alpha p^{(x_1+1)}}\left[e^{-\alpha\delta p^{(x_2+1)}} - e^{-\alpha\delta p^{x_2}}\right]}{\left[e^{-\alpha p^{(x_2+1)}} - e^{-\alpha p^{x_2}}\right]},$$

$$F_\Lambda = F_2(x_1|x_2) = e^{-\alpha\delta p^{(x_1+1)}},$$

and

$$F_\Xi = F_3(x_1|x_2) = e^{-\alpha\delta p^{(x+1)}},$$

Eq (3.5) can be derived as $F_{X_1|X_2=x_2}(x_1|x_2) = \frac{P(X_1\le x_1, X_2=x_2)}{P(X_2=x_2)}$. Similarly, the conditional CDF of $X_1|X_2 \le x_2$ can be expressed as

$$F_{X_1|X_2\le x_2}(x_1|x_2) = \begin{cases} e^{-\alpha p^{(x_1+1)}} e^{-\alpha(1-\delta)p^{(x_2+1)}} & \text{if } x_1 < x_2 \\ e^{-\alpha\delta p^{(x_1+1)}} & \text{if } x_2 < x_1 \\ e^{-\alpha\delta p^{(x+1)}} & \text{if } x_1 = x_2 = x \end{cases}. \tag{3.6}$$

3.3. Joint probability generating function

The joint probability generating function (JPGF) enables the derivation of key descriptive statistics, such as skewness for measuring asymmetry and the joint index of dispersion for assessing variability. These measures help in analyzing the shape of data, including the presence of extreme values. The JPGF of X_1 and X_2, say $G_{BDGu}(x_1, x_2)$, can be derived as follows:

$$\begin{aligned} G_{BDGu}(z_1, z_2) &= E(Z_1^{x_1} Z_2^{x_2}) = \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} \Pr(X_1 = i, X_2 = j)\, z_1^i z_2^j \\ &= \sum_{j=0}^{\infty}\sum_{i=0}^{j-1} \Pr(X_1 = i, X_2 = j)\, z_1^i z_2^j + \sum_{j=0}^{\infty}\sum_{i=j+1}^{\infty} \Pr(X_1 = i, X_2 = j)\, z_1^i z_2^j + \sum_{i=0}^{\infty} \Pr(X_1 = i, X_2 = i)\, z_1^i z_2^j \\ &= \sum_{i=0}^{\infty}\sum_{j=0}^{i-1} P_{BDGu}(i; \alpha, p)\, P_{BDGu}(j; \alpha\delta, p)\, z_1^i z_2^j + \sum_{i=0}^{\infty}\sum_{j=i+1}^{\infty} P_{BDGu}(i; \alpha\delta, p)\, P_{BDGu}(j; \alpha, p)\, z_1^i z_2^j \\ &+ \sum_{i=0}^{\infty}\left[e^{-\alpha\delta p^{i+1}} P_{BDGu}(i; \alpha, p) - e^{-\alpha p^i} P_{BDGu}(i; \alpha\delta, p)\right] z_1^i z_2^i, \end{aligned}$$

where $|z_1| < 1$ and $|z_2| < 1$. The marginal distributions of X_1 and X_2 can also be derived from the JPGF at $z_1 = 1$ and $z_2 = 1$, respectively. Furthermore, based on the JPGF, the skewness and kurtosis measures can be derived as follows:

$$\text{Skewness} = \frac{1}{(1-\rho^2)^3}[\Psi_{30}^2 + \Psi_{03}^2 + 3\left(1+2\rho^2\right)\left(\Psi_{12}^2 + \Psi_{21}^2\right) - 2\rho^3\Psi_{30}\Psi_{03} + 6\rho\{\Psi_{30}\left(\rho\Psi_{12} - \Psi_{21}\right) + \Psi_{03}\left(\rho\Psi_{21} - \Psi_{12}\right) - (2+\rho^2)\Psi_{21}\Psi_{12}\}], \tag{3.7}$$

and

$$\text{Kurtosis} = \frac{\Psi_{40} + \Psi_{04} + 2\Psi_{22} + 4\rho\left(\rho\Psi_{22} - \Psi_{13} - \Psi_{31}\right)}{(1-\rho^2)^2}, \tag{3.8}$$

where $\Psi_{ab} = \frac{E(X_1^a X_2^b)}{\left[\sqrt{Var(X_1)}\right]^a\left[\sqrt{Var(X_2)}\right]^b}$, and $\rho = \frac{E(X_1X_2)-E(X_1)E(X_2)}{\sqrt{Var(X_1)Var(X_2)}}$ $(0 \le \rho \ \le 1)$.

4. Maximum likelihood estimation

Maximum likelihood estimation (MLE) is a widely used method for estimating parameters of the new model by finding the values that maximize the likelihood of observing the given data. It is effective, easy to implement, and provides accurate results, even with complex or noisy datasets, making it valuable for data analysis and hypothesis testing. In this segment, the maximum likelihood is used to estimate the parameters of the BDGu model. Assume that $((x_{11}, x_{21}), ..., (x_{1n}, x_{2n}))$ is a random sample from BDGu model using the following notations: $W_1 = \{k;\ x_{1k} > x_{2k}\}$, $W_2 = \{k;\ x_{1k} < x_{2k}\}$, $W_3 = \{k;\ x_{1k} = x_{2k} = x_k\}$, $W = \cup_{n=1}^{3} W_n$, $|W_i| = n_i; i = 1, 2, 3$, and $\sum_{i=1}^{3} n_i = n$. Thus, the likelihood function can be proposed as follows:

$$l(\alpha, p, \delta) = \prod_{i=1}^{n_1} P_1(x_{1i}, x_{2i}) \prod_{i=1}^{n_2} P_2(x_{1i}, x_{2i}) \prod_{i=1}^{n_3} P_3(x_i). \tag{4.1}$$

The log-likelihood function can be expressed as

$$L(\alpha, p, \delta) = \sum_{i=1}^{n_1} \ln\left[\omega_3(x_{1i}, \alpha, p)\right] + \sum_{i=1}^{n_1} \ln\left[\omega_4(x_{2i}, \alpha, p, \delta)\right] + \sum_{i=1}^{n_2} \ln\left[\omega_4(x_{1i}, \alpha, p, \delta)\right] + \sum_{i=1}^{n_2} \ln\left[\omega_3(x_{2i}, \alpha, p)\right] + \sum_{i=1}^{n_3} \ln\left[e^{-\alpha\delta p^{(x_i+1)}}\omega_3(x_i, \alpha, p) - e^{-\alpha p^{x_i}}\omega_4(x_i, \alpha, p, \delta)\right].$$

Differentiating the log-likelihood with respect to α, p, and δ, and setting the results equal to zero, we have

$$\begin{aligned}\frac{\partial L}{\partial \alpha} = &\sum_{i=1}^{n_1} \frac{p^{x_{1i}}\left[-pe^{-\alpha p^{(x_{1i}+1)}} + e^{-\alpha p^{x_{1i}}}\right]}{e^{-\alpha p^{(x_{1i}+1)}} - e^{-\alpha p^{x_{1i}}}} + \sum_{i=1}^{n_1} \frac{\delta p^{x_{2i}}\left[-pe^{-\alpha\delta p^{(x_{2i}+1)}} + e^{-\alpha\delta p^{x_{2i}}}\right]}{e^{-\alpha\delta p^{(x_{2i}+1)}} - e^{-\alpha\delta p^{x_{2i}}}} \\ &+ \sum_{i=1}^{n_2} \frac{\delta p^{x_{1i}}\left[-pe^{-\alpha\delta p^{(x_{1i}+1)}} + e^{-\alpha\delta p^{x_{1i}}}\right]}{e^{-\alpha\delta p^{(x_{1i}+1)}} - e^{-\alpha\delta p^{x_{1i}}}} + \sum_{i=1}^{n_2} \frac{p^{x_{2i}}\left[-pe^{-\alpha p^{(x_{2i}+1)}} + e^{-\alpha p^{x_2 i}}\right]}{e^{-\alpha p^{(x_{2i}+1)}} - e^{-\alpha p^{x_2 i}}} \\ &+ \sum_{i=1}^{n_3}\left(\frac{p^{x_i}e^{-\alpha\delta p^{(x_i+1)}}\left[-pe^{-\alpha p^{(x_i+1)}} + e^{-\alpha p^{x_i}}\right] - \delta p^{(x_i+1)}e^{-\alpha\delta p^{(x_i+1)}}\omega_3(x_i, \alpha, p)}{e^{-\alpha\delta p^{(x_i+1)}}\omega_3(x_i, \alpha, p) - e^{-\alpha p^{x_i}}\omega_4(x_i, \alpha, p, \delta)}\right. \\ &\left. + \frac{-\delta p^{x_i}e^{-\alpha p^{x_i}}\left[-pe^{-\alpha\delta p^{(x_i+1)}} + e^{-\alpha\delta p^{x_i}}\right] + p^{x_i}e^{-\alpha p^{x_i}}\omega_4(x_i, \alpha, p, \delta)}{e^{-\alpha\delta p^{(x_i+1)}}\omega_3(x_i, \alpha, p) - e^{-\alpha p^{x_i}}\omega_4(x_i, \alpha, p, \delta)}\right),\end{aligned} \tag{4.2}$$

$$\frac{\partial L}{\partial p} = \sum_{i=1}^{n_1} \frac{\alpha p^{(x_{1i}-1)} \omega_1 (x_{1i}, \alpha, p)}{\omega_3 (x_{1i}, \alpha, p)} + \sum_{i=1}^{n_1} \frac{\alpha\delta p^{(x_{2i}-1)} \omega_2 (x_{2i}, \alpha, p, \delta)}{\omega_4 (x_{2i}, \alpha, p, \delta)}$$
$$+ \sum_{i=1}^{n_2} \frac{\alpha\delta p^{(x_{1i}-1)} \omega_2 (x_{1i}, \alpha, p, \delta)}{\omega_4 (x_{1i}, \alpha, p, \delta)} + \sum_{i=1}^{n_2} \frac{\alpha p^{(x_{2i}-1)} \omega_1 (x_{2i}, \alpha, p)}{\omega_3 (x_{2i}, \alpha, p, \delta)}$$
$$+ \sum_{i=1}^{n_3} \left(\frac{\alpha p^{(x_i-1)} e^{-\alpha\delta p^{(x_i+1)}} \omega_1 (x_i, \alpha, p) - \alpha\delta (x_i + 1) p^{x_i} e^{-\alpha\delta p^{(x_i+1)}} \omega_3 (x_i, \alpha, p)}{e^{-\alpha\delta p^{(x_i+1)}} \omega_3 (x_i, \alpha, p) - e^{-\alpha p^{x_i}} \omega_4 (x_i, \alpha, p, \delta)} \right.$$
$$\left. + \frac{-\alpha\delta p^{x_i} e^{-\alpha p^{x_i}} \omega_2 (x_i, \alpha, p, \delta) + \alpha p^{(x_i-1)} x_i e^{-\alpha p^{x_i}} \omega_4 (x_i, \alpha, p, \delta)}{e^{-\alpha\delta p^{(x_i+1)}} \omega_3 (x_i, \alpha, p) - e^{-\alpha p^{x_i}} \omega_4 (x_i, \alpha, p, \delta)} \right), \tag{4.3}$$

and

$$\frac{\partial L}{\partial \delta} = \sum_{i=1}^{n_1} \frac{\alpha p^{x_{2i}} \left[-p e^{-\alpha\delta p^{(x_{2i}+1)}} + e^{-\alpha\delta p^{x_{2i}}} \right]}{\omega_4 (x_{2i}, \alpha, p, \delta)} + \sum_{i=1}^{n_2} \frac{\alpha p^{x_{1i}} \left[-p e^{-\alpha\delta p^{(x_{1i}+1)}} + e^{-\alpha\delta p^{x_{1i}}} \right]}{\omega_4 (x_{1i}, \alpha, p, \delta)}$$
$$+ \sum_{i=1}^{n_3} \frac{-\alpha p^{(x_i+1)} e^{-\alpha\delta p^{(x_i+1)}} \omega_3 (x_i, \alpha, p) - \alpha p^{x_i} e^{-\alpha p^{x_i}} \left[-p e^{-\alpha\delta p^{(x_i+1)}} + e^{-\alpha\delta p^{x_i}} \right]}{e^{-\alpha\delta p^{(x_i+1)}} \omega_3 (x_i, \alpha, p) - e^{-\alpha p^{x_i}} \omega_4 (x_i, \alpha, p, \delta)}, \tag{4.4}$$

where $\omega_1 (x, \alpha, p) = -p (x + 1) e^{-\alpha p^{(x+1)}} + x e^{-\alpha p^x}$, $\omega_2 (x, \alpha, p, \delta) = -p (x + 1) e^{-\alpha\delta p^{(x+1)}} + x e^{-\alpha\delta p^x}$, $\omega_3 (x, \alpha, p) = e^{-\alpha p^{(x+1)}} - e^{-\alpha p^x}$, and $\omega_4 (x, \alpha, p, \delta) = e^{-\alpha\delta p^{(x+1)}} - e^{-\alpha\delta p^x}$. To perform MLE for the BDGu model in R, one can use an optimization routine such as **optim()** to maximize the log-likelihood function. It is important to apply parameter constraints (e.g., lower and upper bounds in **optim()**) to ensure valid estimates, particularly when parameters represent probabilities bounded between 0 and 1. Alternatively, more structured tools such as **maxLik::maxLik()** or **bbmle::mle2()** can also be employed for MLE. Both optimization approaches (**optim()** and **mle2() or maxLik()**) were applied in our paper to ensure robustness and consistency of the parameter estimation results.

5. Simulation: Estimated effectiveness

Simulation studies provide a practical framework for evaluating the performance of the estimator under varying conditions. In this segment, the accuracy and precision of the maximum likelihood estimates for BDGu parameters are assessed across different sample sizes ($n_1 = 25$, $n_2 = 50$, $n_3 = 100$, $n_4 = 150$, $n_5 = 250$, $n_6 = 400$, $n_7 = 500$) and parameter configurations through several testing Schemes: Scheme I: ($\alpha = 3$, $p = 0.8$, $\delta = 0.9$); Scheme II: ($\alpha = 0.9$, $p = 0.8$, $\delta = 0.7$); Scheme III: ($\alpha = 0.5$, $p = 0.8$, $\delta = 0.4$); and Scheme IV: ($\alpha = 0.2, p = 0.9, \delta = 0.9$). The numerical assessments are based on the bias and mean squared errors (MSE). First, we generated N = 10,000 samples of the BDGu model. The empirical results are listed in Tables 1 and 2 and are illustrated in Figures 5–8, receptively.

Table 1. Simulation results of Schemes I and II.

		Scheme I		Scheme II	
n	Parameter	Bias	MSE	Bias	MSE
25	α	0.13933664	0.06622938	0.01766736	0.03823454
	p	−0.27272883	0.12511903	0.03909087	0.02256482
	δ	0.09139273	0.06927388	−0.05754644	0.02602876
50	α	0.11240287	0.03473893	0.01539098	0.02067498
	p	−0.15584412	0.12102987	0.03888937	0.01399083
	δ	0.07788263	0.05532445	−0.05112874	0.02129498
100	α	0.11070256	0.02840887	0.01490983	0.01565773
	p	−0.12415967	0.03088948	0.02924364	0.01303877
	δ	0.04387392	0.02078477	−0.04667426	0.01733245
150	α	0.11020938	0.02318874	0.01139878	0.01018374
	p	−0.08789447	0.01643676	0.01777265	0.01144362
	δ	0.00877263	0.00962334	−0.03435697	0.01517388
250	α	0.10403998	0.01778273	0.00828377	0.00741847
	p	−0.07827784	0.01504984	0.00967464	0.01078999
	δ	0.00369093	0.00643553	−0.01289859	0.01224353
400	α	0.04487624	0.00973874	0.00092934	0.00078745
	p	−0.02344032	0.00625534	0.00324544	0.00746351
	δ	0.00094571	0.00152634	−0.00439851	0.00654742
500	α	0.00234433	0.00125364	0.00007374	0.00000274
	p	−0.00871635	0.00044353	0.00067458	0.00187646
	δ	0.00026384	0.00038925	−0.00098373	0.00022535

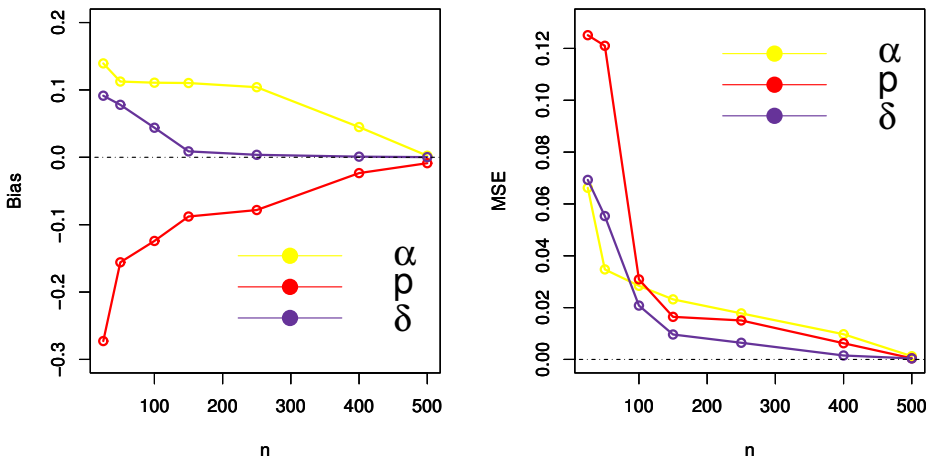

Figure 5. Simulation plots of Scheme I.

Table 2. Simulation results of Schemes III and IV.

		Scheme III		Scheme IV	
n	Parameter	Bias	MSE	Bias	MSE
25	α	−0.14233965	0.35460945	0.38765436	0.22686543
	p	0.26590344	0.52366456	0.21556742	0.12965497
	δ	0.21318549	0.46041184	0.26736458	0.16739904
50	α	−0.11859936	0.23637728	0.31455620	0.15668930
	p	0.14119450	0.21050775	0.13874796	0.08356462
	δ	0.12525842	0.14810338	0.20547267	0.12976456
100	α	−0.07509446	0.08837565	0.29106486	0.14709843
	p	0.13818564	0.19809375	0.01007456	0.00213563
	δ	0.11108965	0.11517505	0.16848929	0.09354768
150	α	−0.04614745	0.03314297	0.24857849	0.13044957
	p	0.11904294	0.16128564	0.00301176	0.00091764
	δ	0.10411745	0.10303645	0.11866493	0.04450254
250	α	−0.01202274	0.02027655	0.18427658	0.11917485
	p	0.10800864	0.13175636	0.00182087	0.00014448
	δ	0.060043974	0.07304685	0.02874695	0.00684931
400	α	−0.00517119	0.01213934	0.10622856	0.01880864
	p	0.02049460	0.01539466	0.00032970	0.00012487
	δ	0.03455097	0.03678198	0.00578947	0.00017465
500	α	−0.00374641	0.00719847	0.01034357	0.00348654
	p	0.00486754	0.00593947	0.00019788	0.00014728
	δ	0.00262835	0.00173611	0.00071545	0.00004382

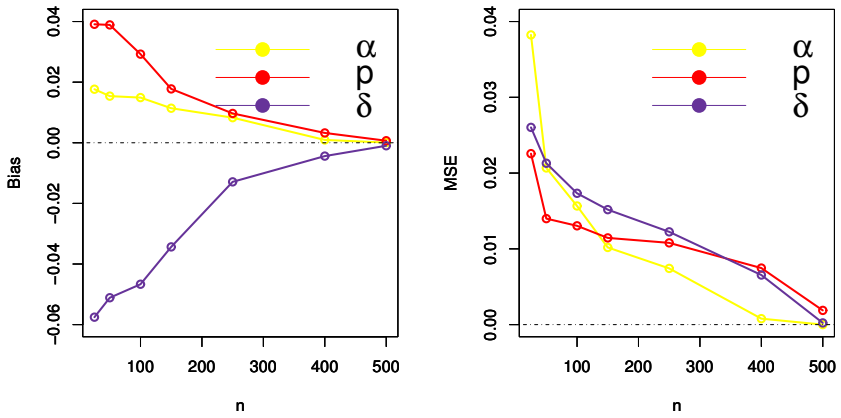

Figure 6. Simulation plots of Scheme II.

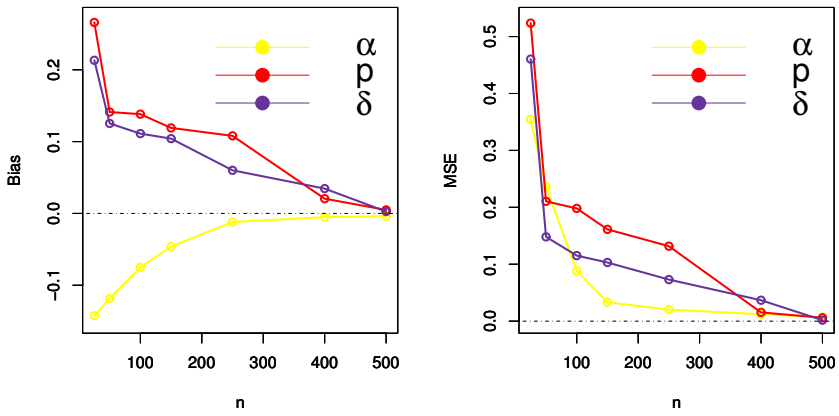

Figure 7. Simulation plots of Scheme III.

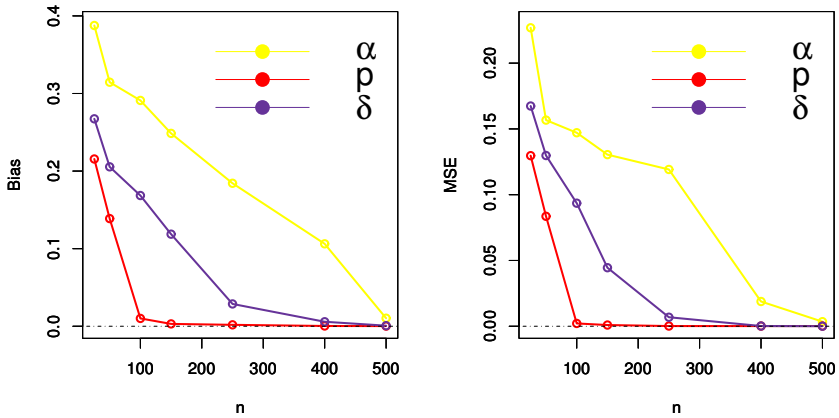

Figure 8. Simulation plots of Scheme IV.

The simulation results show that as the sample size increases, both the bias and MSE of the BDGu parameter estimates decrease, reflecting improved accuracy and reliability. This highlights the consistency and efficiency of the MLE method, confirming its effectiveness for the BDGu model, especially with larger datasets.

6. Empirical significance

In this segment, we illustrate the empirical significance of the BDGu model via three applications of real datasets. Bivariate discrete exponential (BDE), bivariate discrete inverse exponential (BDIE), bivariate discrete exponential (BDE), bivariate discrete inverse exponential (BDIE), bivariate Poisson with minimum operator (BPo_{min}), bivariate Poisson with three parameters (BPo-3P), bivariate Poisson with four parameters (BPo-4P), independent bivariate Poisson (IBPo), bivariate discrete Weibull

(BDW), bivariate discrete inverse Rayleigh (BDIR), bivariate binomial (BBi), and bivariate geometric (BGe) models. The tested models are compared utilizing some criteria namely, the negative maximized log-likelihood ($-L$), Akaike information criterion (AIC), corrected AIC (CAIC), Bayesian information criterion (BIC), and Hannan-Quinn information criterion.

6.1. Dataset I: Football data

The data are listed in Das [16], and report scores of Italian football matches (Serie A) during 1996 to 2011, between ACF Fiorentina (X_1) and Juventus (X_2). Figure 9 shows scatter and box plots of the football score data. As can be seen, there are some extreme points and outliers.

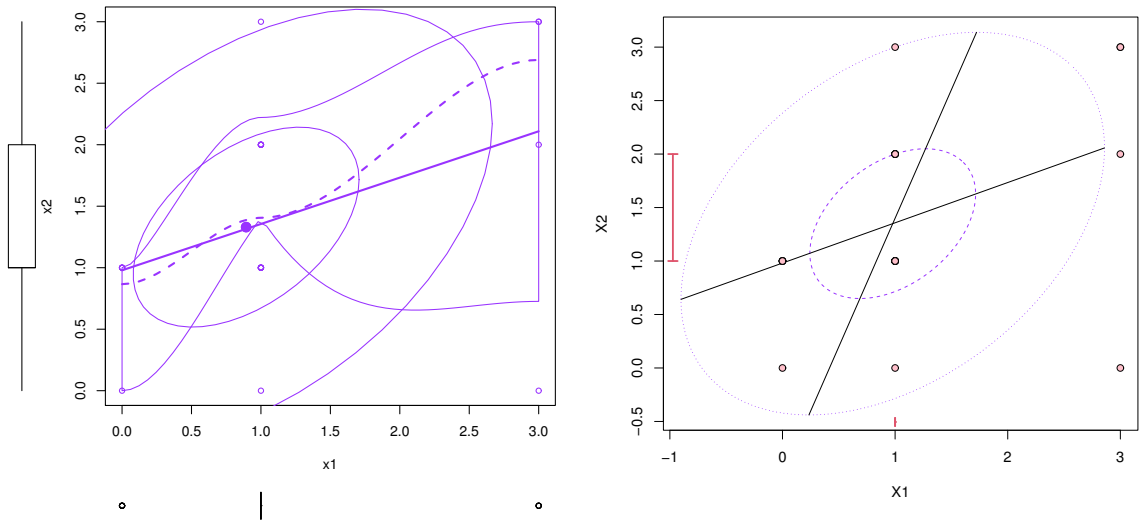

Figure 9. Scatter (left panel) and box (right panel) plots of Dataset I.

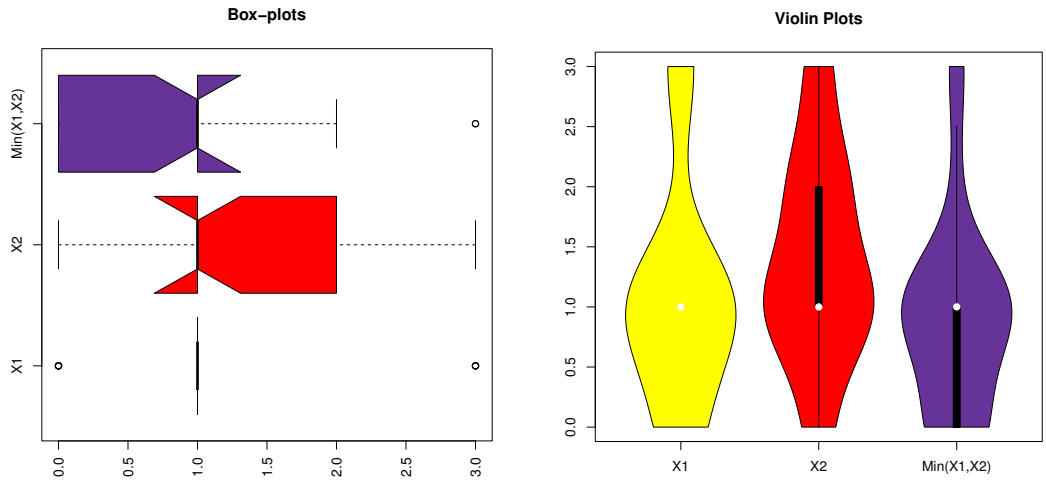

Figure 10. Non-parametric plots for margins in Dataset I.

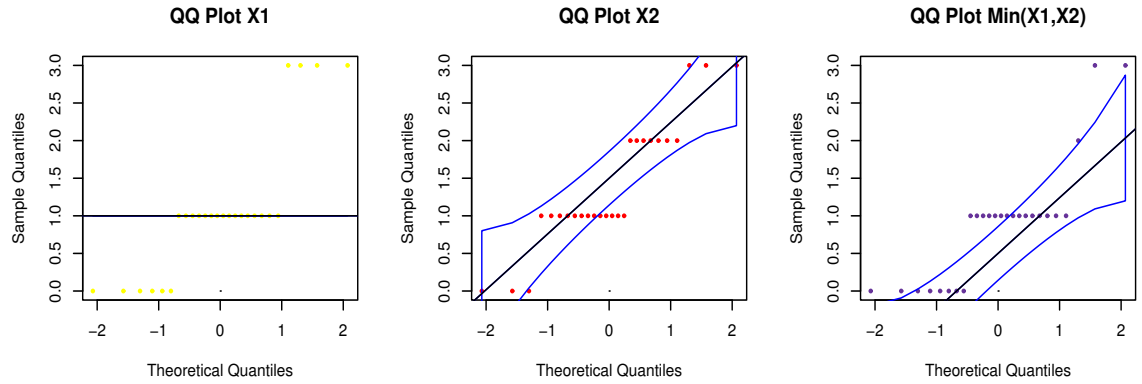

Figure 11. The quantile-quantile (QQ) plots for margins in Dataset I.

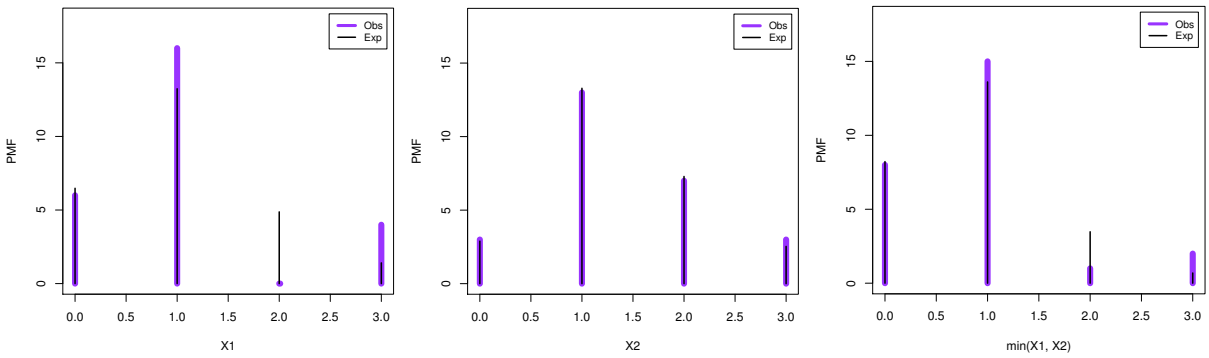

Figure 12. The estimated PMF for margins in Dataset I.

illustrative plots of the data margins can be shown in Figures 10 and 11 for different observation. After testing spread of data and its normality, we can now fit the BDGu model to Dataset I. But before doing this, the margins (X_1, X_2 and $\min(X_1, X_2)$) of the BDGu distribution for Dataset I must be tested. The MLEs of α and p of the corresponding DGu distribution for X_1, X_2 and $\min(X_1, X_2)$ are $(6.977, 0.199)$, $(10.166, 0.216)$, and $(7.592, 0.151)$, respectively. The $-L$ values are 31.78, 31.97, and 28.56, respectively. Moreover, the p-values lie between 0.551 to 0.845. Figure 12 shows the estimated PMF plots for the margins according to Dataset I, which proves the tentative results.

Now, the BDGu distribution is tested to fit Dataset I. The empirical/numerical results are listed in Table 3, and it is noted that the proposed model is the best among all the competing probability models.

Table 3. The MLEs and goodness-of-fit measures (G-O-F-Ms) of Dataset I.

Model	MLEs	−L	AIC	BIC	CAIC	HQIC
BDR	$\widehat{\delta}_1 = 0.790, \widehat{\delta}_2 = 0.872, \widehat{\delta}_3 = 0.905$	63.93	133.86	137.64	134.95	134.95
BDE	$\widehat{\delta}_1 = 0.652, \widehat{\delta}_2 = 0.812, \widehat{\delta}_3 = 0.713$	75.42	156.84	160.62	157.93	157.93
BDIE	$\widehat{\delta}_1 = 0.669, \widehat{\delta}_2 = 0.388, \widehat{\delta}_3 = 0.514$	78.54	163.08	166.86	164.17	164.17
BPo_{min}	$\widehat{\delta}_1 = 1.36, \widehat{\delta}_2 = 2.10, \widehat{\delta}_3 = 2.27$	64.23	134.46	138.23	135.55	135.54
BPo-3P	$\widehat{\eta}_1 = 1.08, \widehat{\eta}_2 = 1.38, \widehat{\eta}_3 = 0.70$	64.93	135.86	139.64	136.96	136.95
IBPo	$\widehat{\sigma}_1 = 1.08, \widehat{\sigma}_2 = 1.38$	67.62	139.25	141.76	139.77	139.97
BDW	$\widehat{\delta}_1 = 0.807, \widehat{\delta}_2 = 0.882, \widehat{\delta}_3 = 0.917, \widehat{\alpha} = 2.125$	63.91	135.82	140.85	137.73	137.27
BDIR	$\widehat{\delta}_1 = 0.493, \widehat{\delta}_2 = 0.212, \widehat{\delta}_3 = 0.561$	64.10	134.20	137.98	135.29	135.29
BDGu	$\widehat{\delta} = 0.757, \widehat{p} = 0.204, \widehat{\alpha} = 8.464$	62.74	131.48	135.26	132.57	132.57

Figure 13 shows the estimated joint PMF for the BDGu distribution and the other competitive models, which support the results of Table 3. Figures 14–16 show the profiles of the L function and contour plots, which indicate that the estimators are unique.

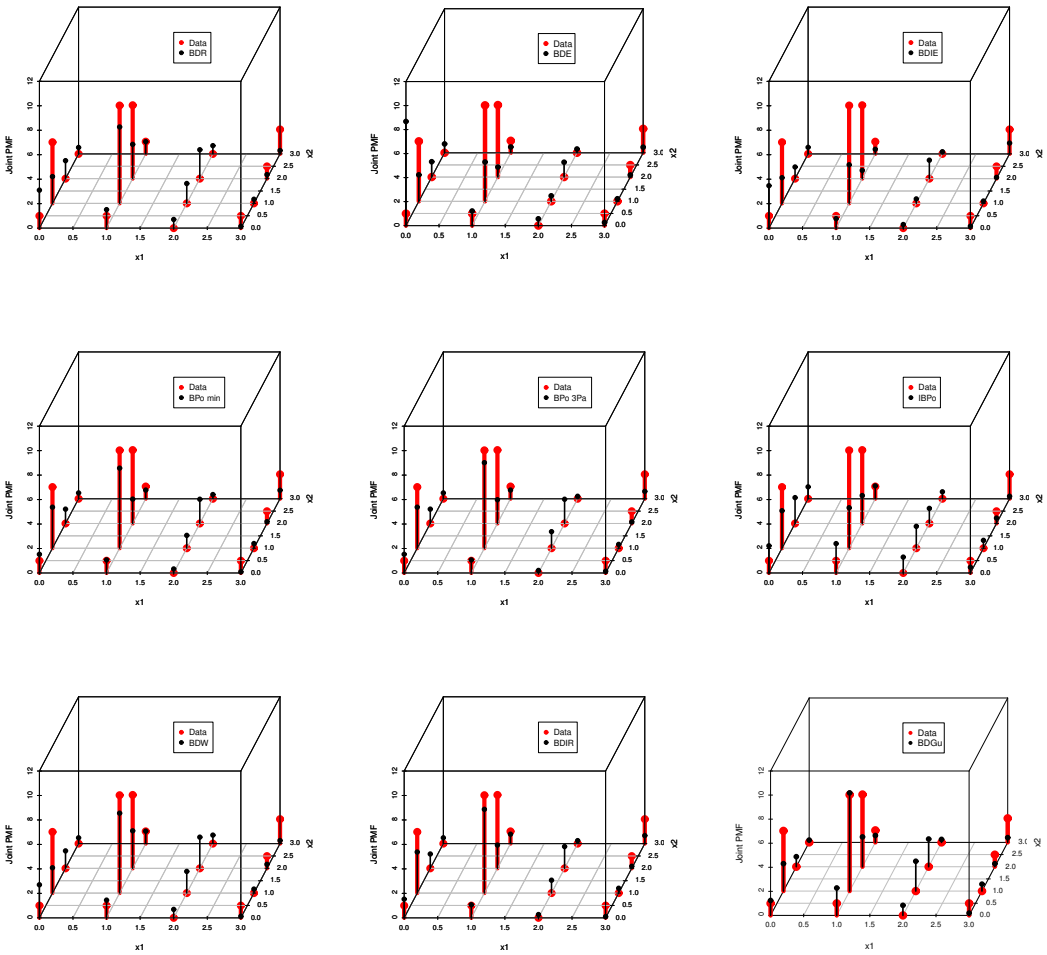

Figure 13. The estimated joint PMF using Dataset I.

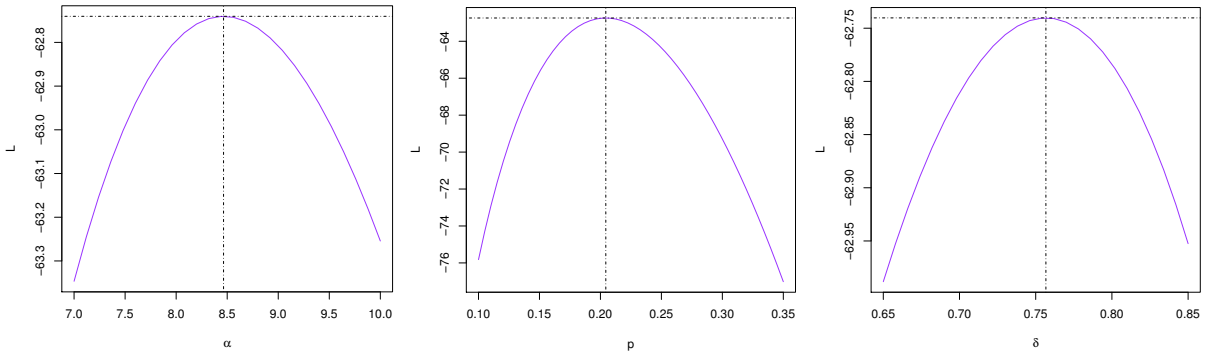

Figure 14. The profile likelihood of α, p, and δ according to Dataset I.

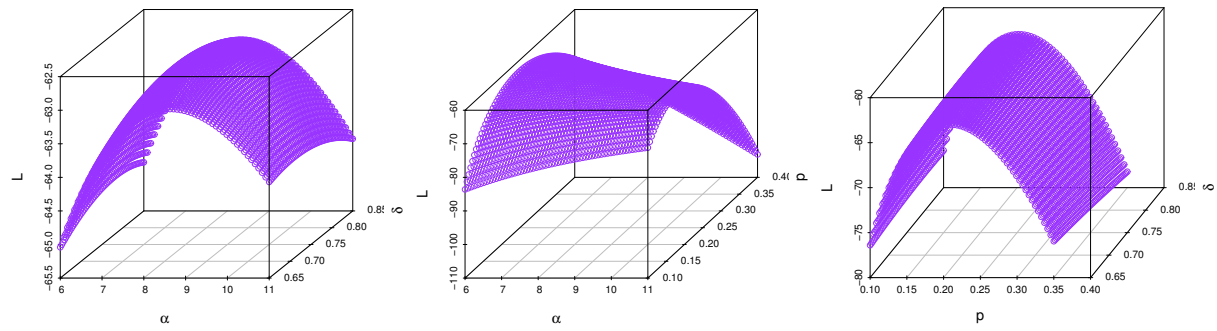

Figure 15. Surface profile likelihood of α, p, and δ based on Dataset I.

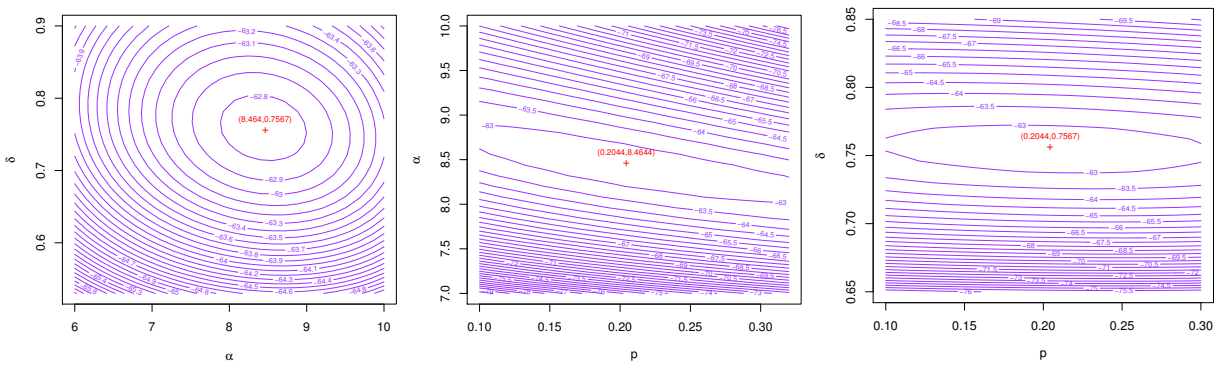

Figure 16. Contour diagrams of the estimators based on Dataset I.

6.2. *Dataset II: Nasal drainage severity score*

Theses data are listed in Davis [17], and it represents the efficacy of steam inhalation in the treatment of common cold symptoms. Figure 17 shows scatter and box plots of Dataset II. As can be observed, there are some extremes and outliers in the observations.

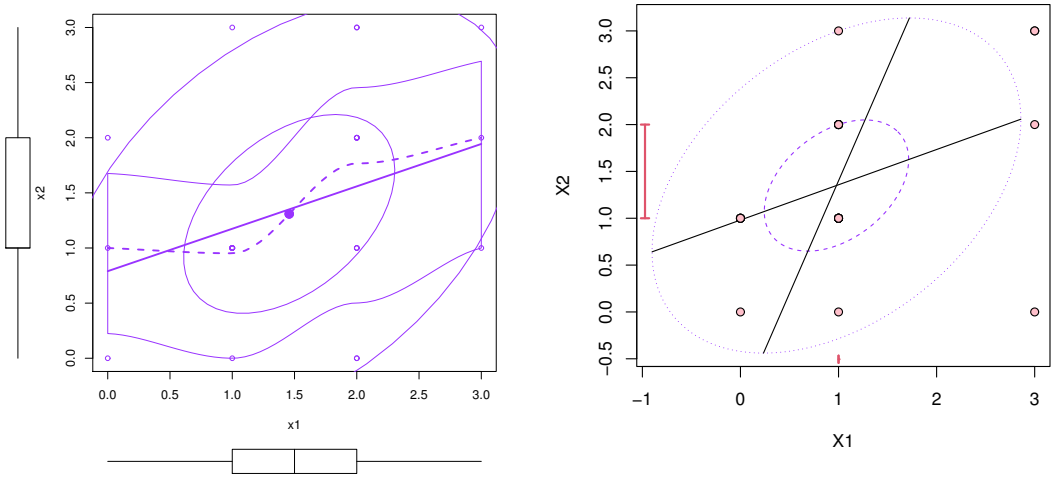

Figure 17. Scatter (left panel) and box (right panel) plots of Dataset II.

Display plots of the data margins can be shown in Figures 18 and 19 for diffuse monitoring. The MLEs of α and p for X_1, X_2, and $\min(X_1, X_2)$ are $(11.096, 0.229)$, $(8.845, 0.229)$, and $(9.043, 0.184)$, respectively. The $-L$ values are 37.31, 38.02, and 34.68, respectively. Further, the p-values lie between 0.601 to 0.923. Figure 20 shows the PMF estimated plots for the marginal regions in Dataset II, which prove the estimated results.

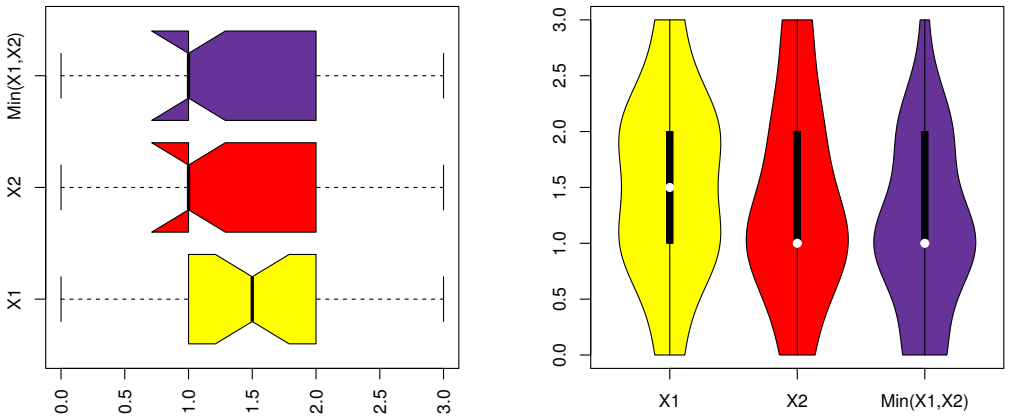

Figure 18. Non-parametric plots for margins in Dataset II.

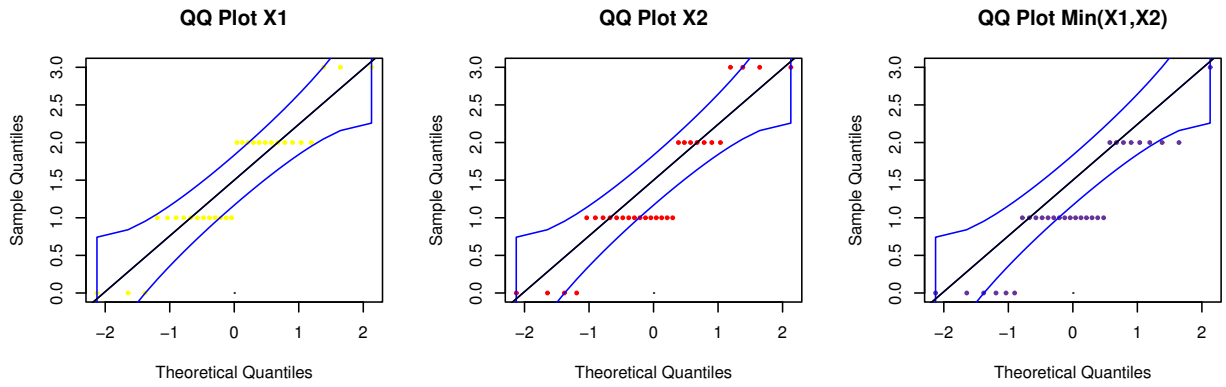

Figure 19. The QQ plots for margins in Dataset II.

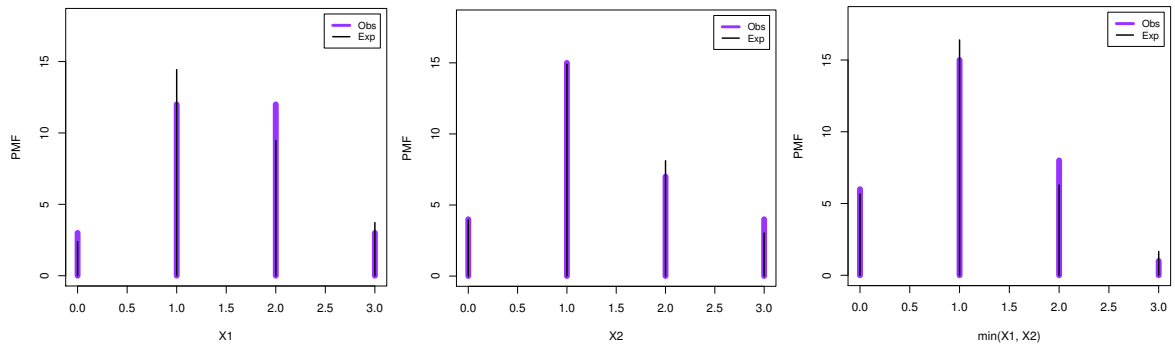

Figure 20. The estimated joint PMF using Dataset II.

Now, the BDGu distribution is tested to fit Dataset II. The empirical results are listed in Table 4, and it is noted that the presented model is the best among all competing distributions.

Table 4. The MLEs and G-O-F-Ms for Dataset II.

Model	MLEs	−L	AIC	BIC	CAIC	HQIC
BDE	$\widehat{\delta_1} = 0.846, \widehat{\delta_2} = 0.792, \widehat{\delta_3} = 0.693$	88.00	182.00	186.20	182.92	183.34
BDIE	$\widehat{\delta_1} = 0.501, \widehat{\delta_2} = 0.622, \widehat{\delta_3} = 0.383$	92.48	190.96	195.16	191.88	192.30
BPo-4P	$\widehat{\eta_1} = 0.262, \widehat{\rho_1} = 0.165, \widehat{\eta_2} = 0.405, \widehat{\rho_2} = 2.97$	77.66	163.33	168.93	164.93	164.66
IBPo	$\widehat{\sigma_1} = 1.499, \widehat{\sigma_2} = 1.367$	92.48	190.96	195.16	191.88	192.30
BDIW	$\widehat{\delta_1} = 0.192, \widehat{\delta_2} = 0.337, \widehat{\delta_3} = 0.360, \widehat{\mu} = 2.453$	76.51	161.02	166.62	162.62	162.81
BDIR	$\widehat{\delta_1} = 0.262, \widehat{\delta_2} = 0.405, \widehat{\delta_3} = 0.363$	78.66	163.32	167.52	164.24	164.66
BDGu	$\widehat{\delta} = 0.586, \widehat{p} = 0.240, \widehat{\alpha} = 9.260$	72.40	150.80	154.998	151.72	152.14

Figure 21 shows the estimated joint PMF for the BDGu model and the other competing distributions, which support the results of Table 4. Figures 22–24 show the profiles of the L function and contour plots.

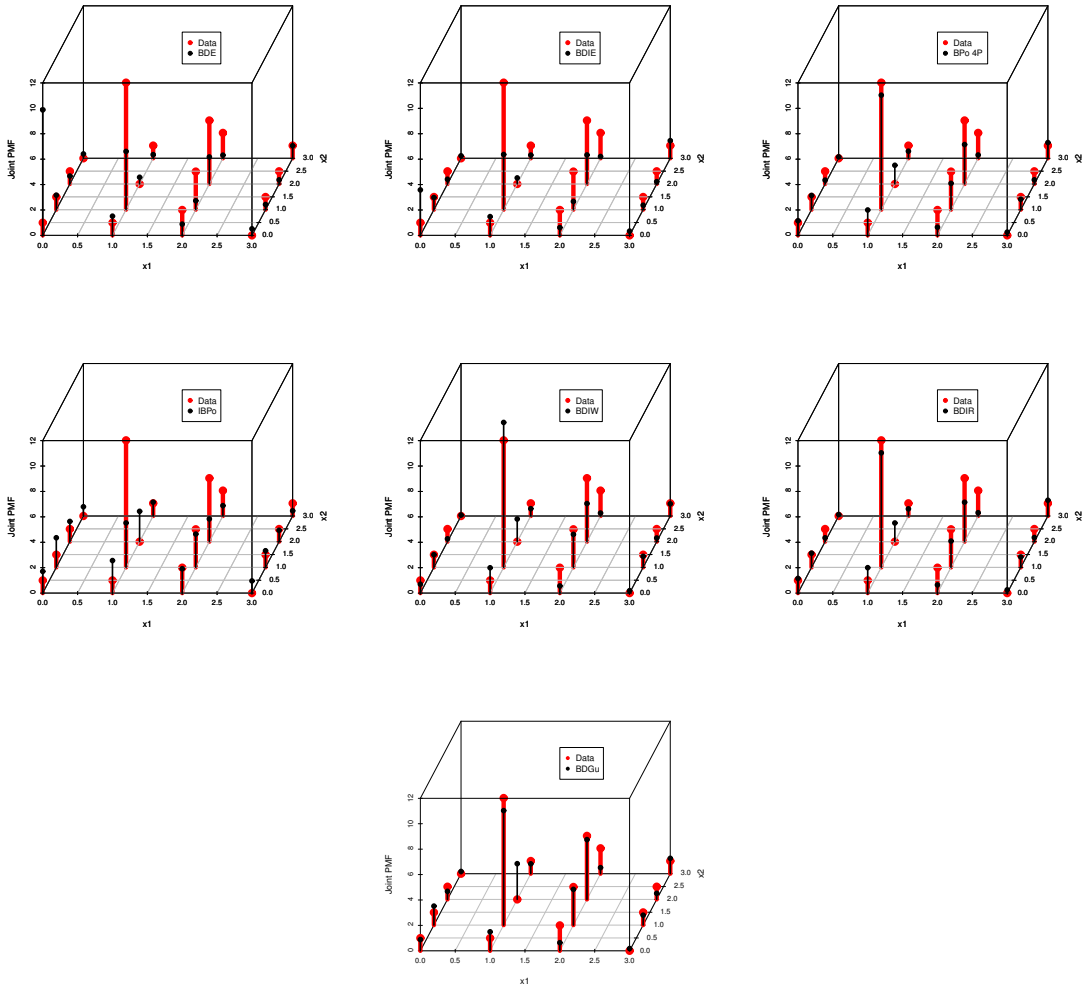

Figure 21. The estimated joint PMF using Dataset II.

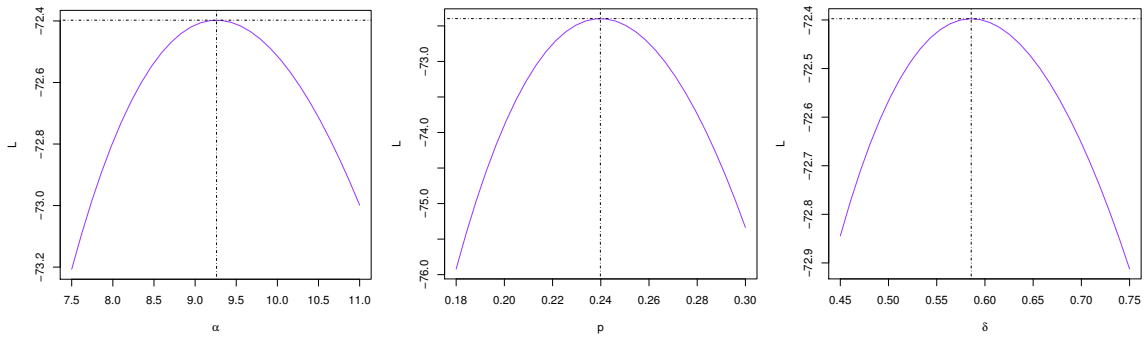

Figure 22. The profile likelihood of α, p, and δ according to Dataset II.

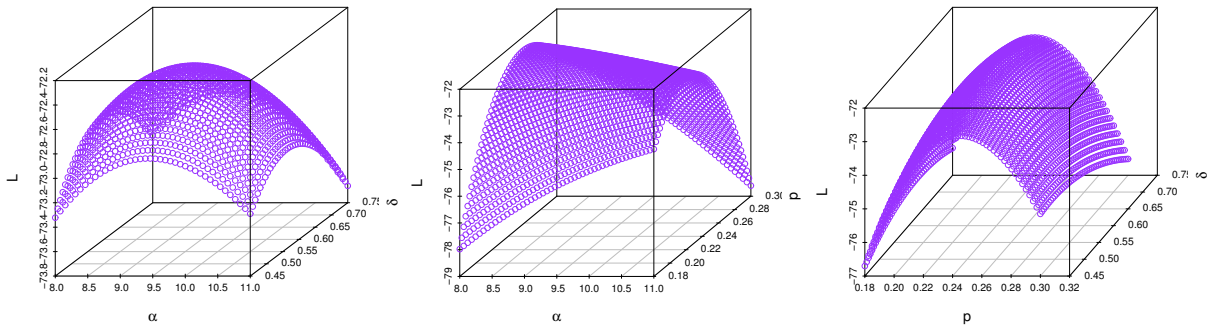

Figure 23. Surface profile likelihood of α, p, and δ based on Dataset II.

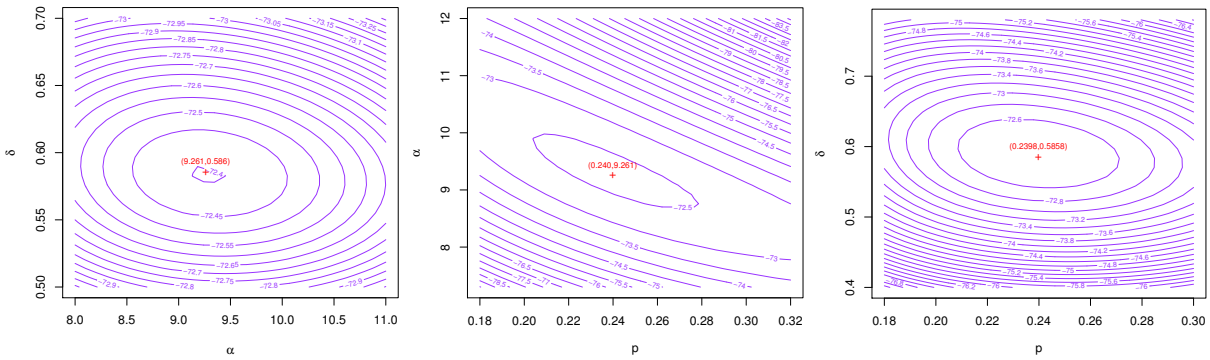

Figure 24. Contour diagrams of the model estimators based on Dataset II.

6.3. Dataset III: Surface and interior faults in lenses

The data is reported by Aitchison and Ho [18], and represents the counts of surface and interior faults in 100 lenses. Figure 25 shows scatter and box plots of Dataset III. As can be observed, there are some extremes and outliers in the observations.Visualization charts for data margins can be viewed in Figures 26 and 27 for diffuse observations. The MLEs of α, and p for X_1, X_2, and $\min(X_1, X_2)$ are $(4.586, 0.569)$, $(4.141, 0.552)$, and $(4.938, 0.380)$, respectively. The $-L$ values are 217.534, 214.796, and 164.578, respectively. Furthermore, the p-values lie between 0.425 to 0.652. Figure 28 shows the PMF plots estimated for the margins domains according to the Dataset III, substantiating the preliminary findings.

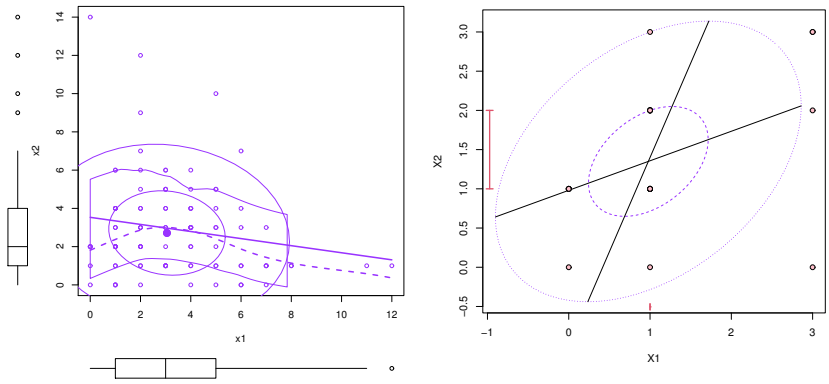

Figure 25. Scatter (left panel) and box (right panel) plots of Dataset III.

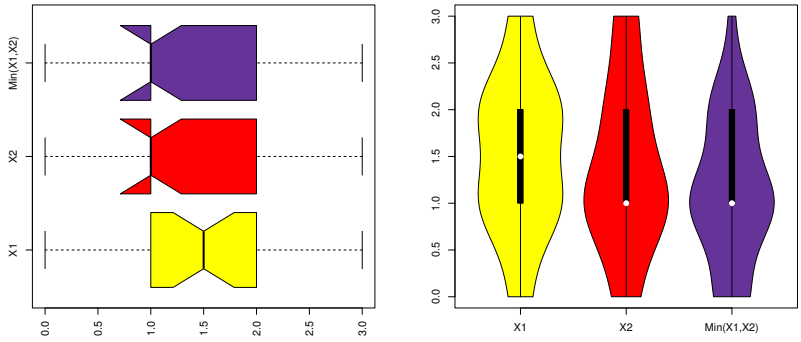

Figure 26. Non-parametric plots for margins in Dataset III.

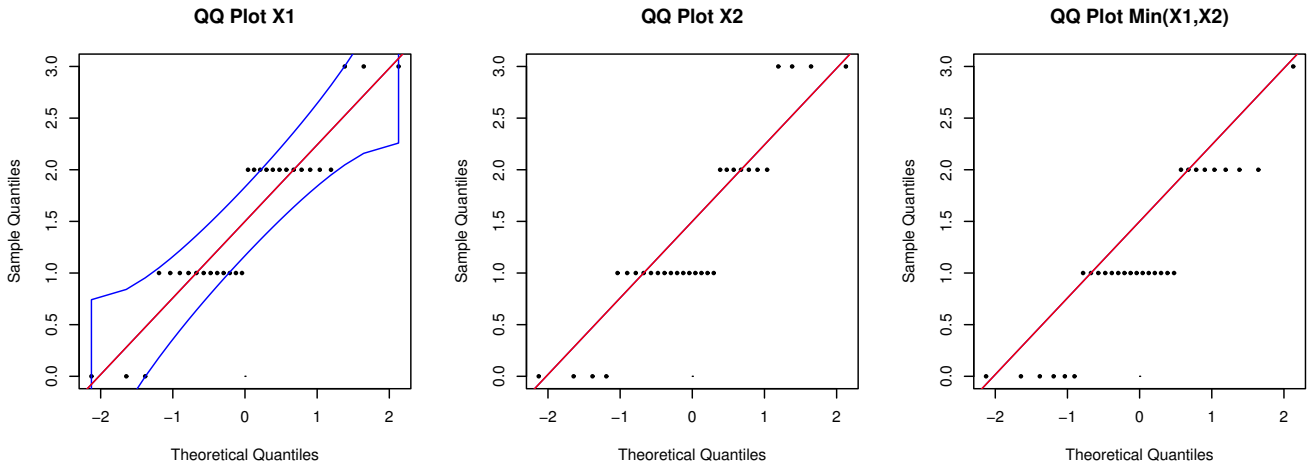

Figure 27. The QQ plots for margins in Dataset III.

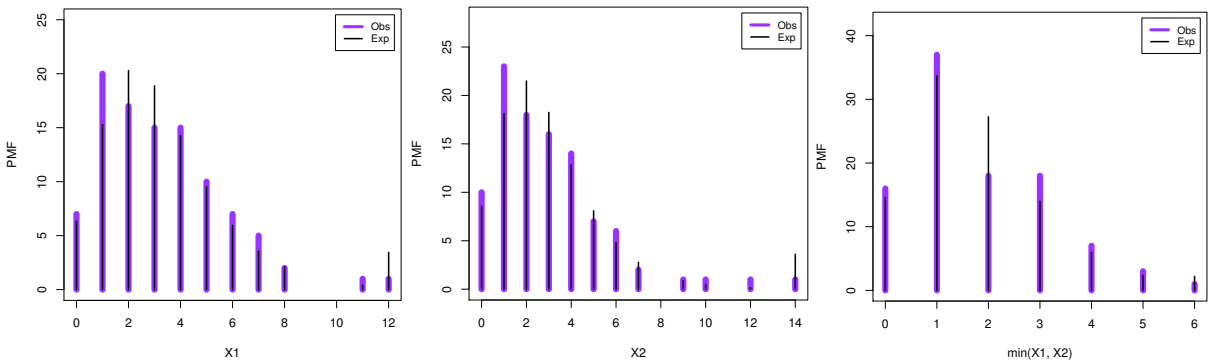

Figure 28. The estimated joint PMF using Dataset III.

The empirical results of the BDGu distribution are reported in Table 5, and it is found that the new model is the best among all tested distributions.

Table 5. The MLEs and G-O-F-Ms for Dataset III.

Model	MLEs	−L	AIC	BIC	CAIC	HQIC
BPo_{min}	$\widehat{\delta_1} = 3.25, \widehat{\delta_2} = 2.93, \widehat{\delta_3} = 4.16 \times 10^{-4}$	450.31	906.62	914.44	906.87	909.78
BBi	$\widehat{\eta_1} = 175, \widehat{\eta_2} = 255, \widehat{\delta_1} = 0.0157, \widehat{\delta_2} = 0.0099,$ $\widehat{\alpha_1} = 0.0237, \widehat{\alpha_2} = 0.0146$	439.02	890.04	905.67	890.94	896.37
BGe	$\widehat{\delta_1} = 0.201, \widehat{\delta_2} = 0.222, \widehat{\alpha_1} = 0.304, \widehat{\alpha_2} = 0.344$	445.21	898.42	908.84	898.84	902.64
BDIW	$\widehat{\theta_1} = 0.045, \widehat{\theta_2} = 0.078, \widehat{\theta_3} = 0.861, \widehat{\alpha} = 1.508$	450.32	908.63	919.05	909.05	912.85
BPo-4P	$\widehat{\delta_1} = 2.748, \widehat{\delta_2} = 2.550, \widehat{\alpha_1} = 4.162, \widehat{\alpha_2} = 3.703$	438.20	884.40	894.82	884.82	888.62
BDGu	$\widehat{\delta} = 0.9995, \widehat{p} = 0.5620, \widehat{\alpha} = 4.333$	433.03	872.06	879.88	872.31	875.22

Figures 29–31 show the profiles of the L function and contour plots.

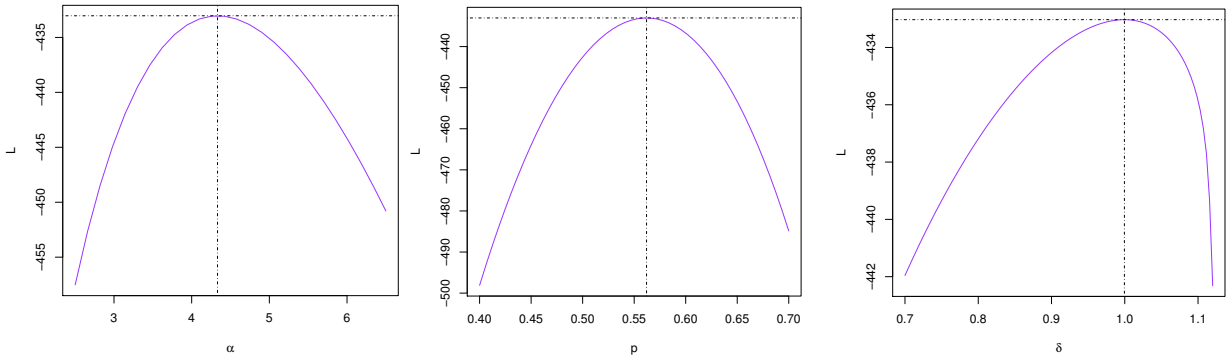

Figure 29. The profile likelihood of α, p, and δ according to Dataset III.

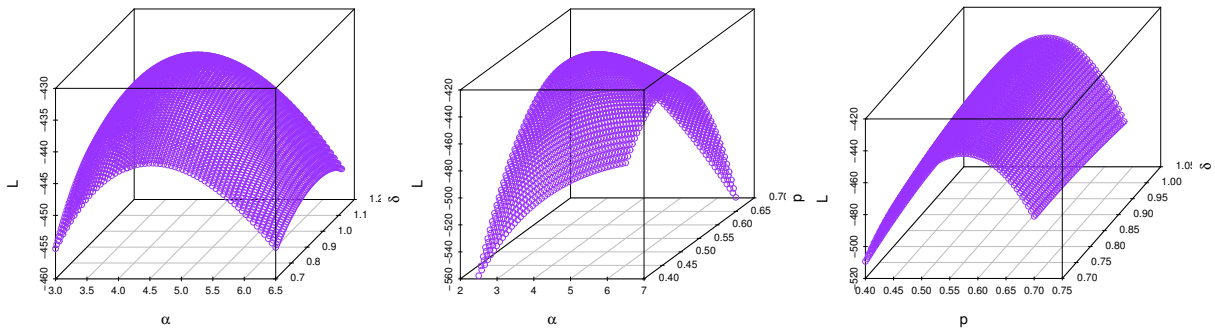

Figure 30. Surface profile likelihood of α, p, and δ based on Dataset III.

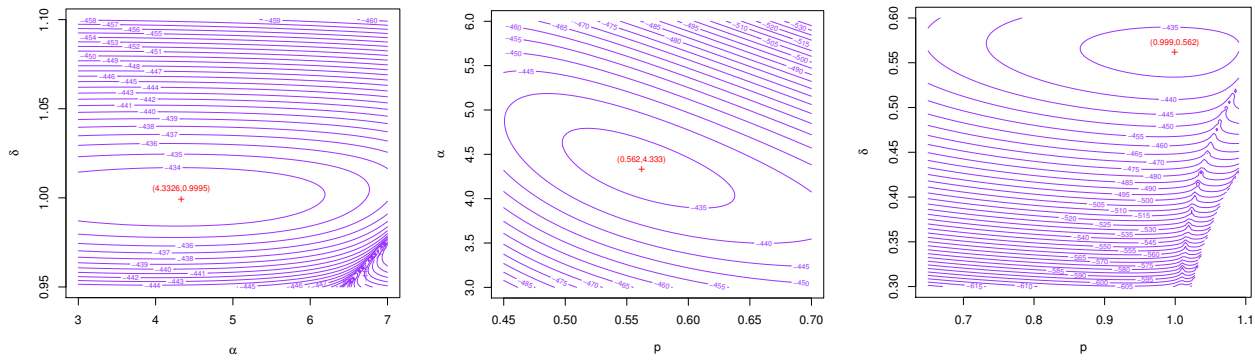

Figure 31. Contour diagrams of the model estimators based on Dataset III.

6.4. Execution of PQD and TP2 on Datasets I, II, and III

To showcase the versatility of the BDGu model in handling various types of bivariate discrete data, it was applied across three real-world datasets. In the context of football scores, the model effectively captured the positive dependence and variability between team performances, reflecting dynamic match conditions. For nasal secretion severity scores, it modeled co-occurring symptoms with asymmetric dependence and extreme values often seen in clinical data. In the case of lens defects, the model accurately represented the joint behavior of surface and interior failures, aiding in reliability analyses. The BDGu model's strengths lie in its ability to handle PQD and TP2, which together enhance its capability to model interdependent outcomes. PQD allows the model to reflect natural associations such as high scores, symptoms, or defect rates occurring together, while TP2 ensures a structured and interpretable relationship between variables. These properties make the BDGu model particularly effective for applications in sports analytics, medical diagnostics, and manufacturing reliability, where modeling joint behaviors, risk, and progression patterns is crucial.

6.5. BDGu-driven synchronization in Kuramoto and fractional networks

The BDGu distribution plays a significant role in modelling and analyzing systems where joint extremes and dependency structures are essential, particularly in synchronization dynamics and energy consumption analysis of complex networks such as Kuramoto-oscillator systems and multilayer fractional-order networks. In the context of fixed or prescribed-time synchronization and energy consumption in Kuramoto-oscillator networks, the BDGu model provides a powerful framework to capture the joint behaviour of extreme events, such as simultaneous energy spikes or synchronization delays across oscillators. The model's joint HRF can take various shapes depending on the estimated parameters, allowing it to flexibly describe increasing, decreasing, unimodal, or bathtub-shaped failure rates, all of which are highly relevant to practical network dynamics and maintenance scheduling. Moreover, in finite-time synchronization and energy consumption predictions for multilayer fractional-order networks, the BDGu model helps us identify not only when systems synchronize, but also the probabilistic structure of outlier behaviour or extreme energy use across layers or nodes. This is particularly useful for forecasting critical resource demands and designing more efficient, resilient control strategies.

One of the strengths of this distribution lies in its marginal flexibility, which allows it to be used independently to study the failure behaviour of individual subsystems. This supports applications in preventive maintenance, stress-strength reliability modelling, and the assessment of component-wise degradation in network nodes. Additionally, the BDGu model is particularly suited for modelling extremes and outliers in discrete data, making it an excellent candidate in systems where rare but high-impact synchronization failures or energy overloads may occur. This aligns well with real-world network scenarios, where such anomalies often dictate system-wide performance and reliability. From a statistical perspective, the BDGu model possesses two important properties: PQD, which captures the tendency of variables to move in the same direction, and TP2, which ensures monotonicity in joint probabilities making, it a robust tool for analyzing synchronization dynamics and energy consumption in complex networks.

7. Final remarks and future directions

This study presented and analyzed a new bivariate discrete distribution known as the BDGu, developed using the Cuadras-Auge copula framework. The model's theoretical foundations were carefully examined, followed by a detailed investigation of its statistical properties. The results showed that the BDGu distribution is well-suited for modeling asymmetric, dispersed, and heavy-tailed discrete data with diverse kurtosis patterns. Its joint hazard rate function proved capable of representing a wide range of failure behaviors, including increasing, decreasing, unimodal, and bathtub-shaped trends. The model also features two key dependence properties positive quadrant dependence and second-order total positivity which enhance its practical relevance in domains such as engineering, healthcare, and sports analytics. Maximum likelihood estimation was applied for parameter inference, and simulation studies confirmed the method's accuracy and consistency across varying sample sizes. Real-world applications using three distinct datasets further highlighted the BDGu model's strong performance compared with existing alternatives, based on statistical goodness-of-fit criteria. Looking ahead, future research will focus on extending the model to multivariate settings and exploring its applications in regression and time series analysis, as well as

implementing Bayesian and expectation-maximization approaches for deeper inferential insights.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/2025/01/33373).

Data availability statement

The data sets are available in the paper.

Conflicts of interest

The authors declare no conflict of interests.

References

1. S. Kotz, S. Nadarajah, *Extreme Value Distributions: Theory and Applications*, World Scientific, 2000. https://doi.org/10.1142/p191
2. S. Chakraborty, D. Chakravarty, J. Mazucheli, W. Bertoli, A discrete analog of Gumbel distribution: Properties, parameter estimation and applications, *J. Appl. Stat.*, **48** (2021), 712–737. https://doi.org/10.1080/02664763.2020.1744538
3. H. Boubaker, N. Sghaier, Markov-switching time-varying copula modeling of dependence structure between oil and GCC stock markets, Open *Open J. Stat.*, **6** (2016), 565–589. http://doi.org/10.4236/ojs.2016.64048
4. T. P. Hutchinson, C. D. Lai, *Continuous Bivariate Distributions Emphasising Applications*, Rumsby Scientific, 1990.
5. M. S. Eliwa, M. El-Morshedy, Bivariate Gumbel-G family of distributions: Statistical properties, Bayesian and non-Bayesian estimation with application, *Ann. Data Sci.*, **6** (2019), 39–60. https://doi.org/10.1007/s40745-018-00190-4
6. R. B. Nelsen, *An Introduction to Copulas*, Springer, 2007. https://link.springer.com/book/10.1007/0-387-28678-0
7. C. Ning, Dependence structure between the equity market and the foreign exchange market – a copula approach, *J. Int. Money Finance*, **29** (2010), 743–759. https://doi.org/10.1016/j.jimonfin.2009.12.002
8. E. M. Almetwally, H. Z. Muhammed, E. S. A. El-Sherpieny, Bivariate Weibull distribution: Properties and different methods of estimation, *Ann. Data Sci.*, **7** (2020), 163–193. https://doi.org/10.1007/s40745-019-00197-5

9. M. El-Morshedy, M. S. Eliwa, A bivariate probability generator for the odd generalized exponential model: Mathematical structure and data fitting, *Filomat*, **38** (2024), 1109–1133. https://doi.org/10.2298/FIL2403109E
10. E. S. A. El-Sherpieny, H. Z. Muhammed, E. M. Almetwally, Bivariate Chen distribution based on copula function: Properties and application of diabetic nephropathy, *J. Stat. Theory Pract.*, **16** (2022), 54. https://doi.org/10.1007/s42519-022-00275-7
11. S. O. Susam, A multi-parameter Generalized Farlie-Gumbel-Morgenstern bivariate copula family via Bernstein polynomial, *Hacettepe J. Math. Stat.*, **51** (2022), 1–14. https://doi.org/10.15672/hujms.993698
12. S. T. Obeidat, D. Das, M. S. Eliwa, B. Das, P. J. Hazarika, W. W. Mohammed, Two-dimensional probability models for the weighted discretized Fréchet–Weibull random variable with Min–Max operators: Mathematical theory and statistical goodness-of-fit analysis, *Mathematics*, **13** (2025), 625. https://doi.org/10.3390/math13040625
13. C. M. Cuadras, J. Augé, A continuous general multivariate distribution and its properties, *Commun. Stat. Theory Methods*, **10** (1981), 339–353. https://doi.org/10.1080/03610928108828042
14. S. Yang, D. Meng, A. Díaz, H. Yang, X. Su, A. M. P. D. Jesus, Probabilistic modeling of uncertainties in reliability analysis of mid- and high-strength steel pipelines under hydrogen-induced damage, *Int. J. Struct. Integrity*, **16** (2025), 39–59. https://doi.org/10.1108/IJSI-10-2024-0177
15. S. Yang, D. Meng, H. Yang, C. Luo, X. Su, Enhanced soft Monte Carlo simulation coupled with support vector regression for structural reliability analysis, *Proc. Inst. Civ. Eng. Transp.*, **2024** (2024), 1–16. https://doi.org/10.1680/jtran.24.00128
16. D. Das, T. S. Alshammari, K. A. Rashedi, B. Das, P. J. Hazarika, M. S. Eliwa, Discrete joint random variables in Fréchet-Weibull distribution: A comprehensive mathematical framework with simulations, goodness-of-fit analysis, and informed decision-making, *Mathematics*, **12** (2024), 3401. https://doi.org/10.3390/math12213401
17. C. S. Davis, *Statistical Methods for the Analysis of Repeated Measures Data*, Springer, 2002. https://link.springer.com/book/10.1007/b97287
18. J. Aitchison, C. H. Ho, The multivariate Poisson-log normal distribution, *Biometrika*, **76** (1989), 643–653. https://doi.org/10.1093/biomet/76.4.643