Filomat 34:8 (2020), 2511–2531
https://doi.org/10.2298/FIL2008511E

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Bayesian and Non-Bayesian Estimation of Four-Parameter of Bivariate Discrete Inverse Weibull Distribution with Applications to Model Failure Times, Football and Biological Data

M. S. Eliwa[a], M. El-Morshedy[b,a]

[a]*Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt.*
[b]*Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia.*

Abstract. In this paper we have considered one model, namely the bivariate discrete inverse Weibull distribution, which has not been considered in the statistical literature yet. The proposed model is a discrete analogue of Marshall-Olkin inverse Weibull distribution. Some of its important statistical properties are studied. Maximum likelihood and Bayesian mmethods are used to estimate the model parameters. A detailed simulation study is carried out to examine the bias and mean square error of maximum likelihood and Bayesian estimators. Finally, three real data sets are analyzed to illustrate the importance of the proposed model.

1. Introduction

The Weibull (W) distribution and its inverse model have the ability to model failure rates which are quite common in reliability and biological studies. Therefore, more modifications and extensions of W distribution and its inverse model have been presented in the statistical literature to describe various phenomena in different fields, see for example, Silva et al. (2010), Almalki and Yuan (2013), Jehhan et al. (2018), El-Morshedy and Eliwa (2019), Shakhatreh et al. (2019), Salah el al. (2020), Tahir et al. (2020), El-Morshedy et al. (2020a), and references cited therein.

In many practical situations, it is important to consider different bivariate distributions that could be used to model bivariate data. Therefore, many bivariate distributions have been reported in the literature to discuss and analyze such these data. For instance, Olkin and Liu (2003), Wang and Rennolls (2007), Erdem and Shi (2011), Li et al. (2012), Eliwa and El-Morshedy (2019, 2020), Eliwa et al. (2020a), El-Morshedy et al. (2020b), and references cited therein.

While continuous random quantities are commonly observed in practice, discrete random variables can also be encountered frequently for many different practical reasons. For example, in lifetime modeling,

2010 *Mathematics Subject Classification*. 60E05; 62E10; 62N05
Keywords. Bivariate discrete distributions; Joint hazard rate function; Bayesian estimation; Likelihood ratio test statistic; Simulation.
Received: 29 July 2019; Revised: 30 December 2019; Accepted: 07 March 2020
Communicated by Biljana Popović
Corresponding author: M. El-Morshedy
ORCID: M. S. Eliwa: https://orcid.org/0000-0001-5619-210X, M. El-Morshedy: https://orcid.org/0000-0002-7585-5519
Email addresses: mseliwa@mans.edu.eg (M. S. Eliwa), m.elmorshedy@psau.edu.sa; mah_elmorshedy@mans.edu.eg (M. El-Morshedy)

field failures are often collected and listed daily, weekly, and so forth, Jazi et al. (2010), Nekoukhou and Bidram (2015), Eliwa et al. (2020b), El-Morshedy et al. (2020c), and references cited therein. Moreover, discrete bivariate models can also be useful for many other disciplines. Thus, several bivariate discrete distributions have been proposed and studied in the literature mainly to analyze bivariate discrete data. See for example, Wu et al. (2003), Yuen et al. (2006), Morata (2009), Lee and Cha (2015), Nekoukhou and Kundu (2017), Kundu and Nekoukhou (2018), Jia et al. (2019), El-Morshedy et al. (2020d), and references cited therein.

In this regard, we propose a flexible bivariate discrete distribution, in the so-called the bivariate discrete inverse Weibull (BDsIW) distribution. This model can be usefully applied not only by statisticians, but also by data analysis in many different disciplines, such as sports, engineering and medical applications. The BDsIW distribution can be obtained from 3-independent discrete inverse Weibull (DsIW) distributions by using the maximization method as suggested by Lee and Cha (2015).

The BDsIW model can be considered as a natural discrete analogue of the Marshall-Olkin bivariate inverse Weibull distribution. The bivariate discrete inverse exponential (BDsIE) and bivariate discrete inverse Rayleigh (BDsIR) distributions can be obtained as special cases. Since the joint cumulative distribution function (CDF), joint reliability function (RF), and joint probability mass function (PMF) of the BDsIW model can be expressed as closed forms, it is recommended to utilize the BDsIW model instead of all other competitive models. In the applied fields, especially in the field of modeling, the BDsIW model could be useful in the following cases:

1. Maintenance model or stress model as proposed by Kundu and Gupta (2009).
2. Modeling skewed data sets, especially the right skewed heavy tail data sets.
3. In physics and reliability analysis, the BDsIW model can be applied in modeling the failure time data. As shown in Table 6, the BDsIW model showed its superiority against the bivariate Poisson (BPo), bivariate binomial (BBi), bivariate Poisson Holgate (BPoH), independent bivariate Poisson (IBPo), BDsIE and BDsIR distributions.
4. In sports field, the BDsIW model can be applied in modeling the football data. As shown in Table 10, the BDsIW model provides better fits than other many well-known competitive models.
5. In the medicine field, the BDsIW model can be applied in modeling the nasal drainage severity data. As shown in Table 14, the BDsIW model showed its superiority against many well-known competitive models.

2. The BDsIW Distribution

Jazi *et al.* (2010) introduced the DsIW distribution. The PMF of the DsIW distribution can be written as

$$\pi(x;\theta,\zeta) = \theta^{(x+1)^{-\zeta}} - \theta^{x^{-\zeta}};\ \ x \in \mathbb{N}_0, \tag{1}$$

where $0 < \theta < 1$ and $\zeta > 0$. Clearly, the discrete exponential (DsE) and discrete Rayleigh (DsR) distributions can be obtained as special cases for $\zeta = 1$ and $\zeta = 2$, respectively. Suppose $W_1 \sim \mathrm{DsIW}(\theta_1,\zeta)$, $W_2 \sim \mathrm{DsIW}(\theta_2,\zeta)$ and $W_3 \sim \mathrm{DsIW}(\theta_3,\zeta)$ and they are independently distributed. If $X_d = \max(W_d, W_3); d = 1, 2$, then we can say that the bivariate vector $\mathbf{X} = (X_1, X_2)$ has a BDsIW distribution with the parameter vector $\Psi = (\theta_1,\theta_2,\theta_3,\zeta)^T$. We will denote this discrete bivariate distribution by $\mathrm{BDsIW}(\theta_1,\theta_2,\theta_3,\zeta)$. If $\mathbf{X} \sim \mathrm{BDsIW}(\theta_1,\theta_2,\theta_3,\zeta)$, then the joint CDF of $\mathbf{X}$ is given by

$$\begin{aligned} F_{X_1,X_2}(x_1,x_2;\Psi) &= \theta_1^{(x_1+1)^{-\zeta}}\theta_2^{(x_2+1)^{-\zeta}}\theta_3^{(x_3+1)^{-\zeta}} \\ &= F_{\mathrm{DsIW}}(x_1;\theta_1,\zeta)\ F_{\mathrm{DsIW}}(x_2;\theta_2,\zeta)\ F_{\mathrm{DsIW}}(x_3;\theta_3,\zeta) \\ &= \begin{cases} F_{\mathrm{DsIW}}(x_1;\theta_1\theta_3,\zeta)\ F_{\mathrm{DsIW}}(x_2;\theta_2,\zeta) & ;\ 0 < x_1 < x_2 < \infty \\ F_{\mathrm{DsIW}}(x_1;\theta_1,\zeta)\ F_{\mathrm{DsIW}}(x_2;\theta_2\theta_3,\zeta) & ;\ 0 < x_2 < x_1 < \infty \\ F_{\mathrm{DsIW}}(x;\theta_1\theta_2\theta_3,\zeta) & ;\ 0 < x_1 = x_2 = x < \infty, \end{cases} \end{aligned} \tag{2}$$

where $x_1 \in \mathbb{N}_0$, $x_2 \in \mathbb{N}_0$ and $x_3 = \min\{x_1, x_2\}$. The marginal CDFs of the BDsIW distribution can be represented as

$$F_{X_d}(x_d; \theta_d, \theta_3, \zeta) = P[\max(W_d, W_3) \leq x_d] = F_{\mathrm{DsIW}}(x_d; \theta_d\theta_3, \zeta)\,'; d = 1, 2. \tag{3}$$

The corresponding joint PMF to Equation (2) is given by

$$f_{X_1,X_2}(x_1, x_2; \Psi) = \begin{cases} f_1(x_1, x_2; \Psi) & ;\ 0 < x_1 < x_2 < \infty \\ f_2(x_1, x_2; \Psi) & ;\ 0 < x_2 < x_1 < \infty \\ f_3(x; \Psi) & ;\ 0 < x_1 = x_2 = x < \infty, \end{cases} \tag{4}$$

where

$$f_1(x_1, x_2; \Psi) = f_{\mathrm{DsIW}}(x_1; \theta_1\theta_3, \zeta)\, f_{\mathrm{DsIW}}(x_2; \theta_2, \zeta),$$
$$f_2(x_1, x_2; \Psi) = f_{\mathrm{DsIW}}(x_1; \theta_1, \zeta)\ f_{\mathrm{DsIW}}(x_2; \theta_2\theta_3, \zeta)$$

and

$$f_3(x; \Psi) = F_{\mathrm{DsIW}}(x; \theta_2, \zeta)\, f_{\mathrm{DsIW}}(x; \theta_1\theta_3, \zeta) - F_{\mathrm{DsIW}}(x - 1; \theta_2\theta_3, \zeta)\, f_{\mathrm{DsIW}}(x; \theta_1, \zeta).$$

The expressions $f_1(x_1, x_2; \Psi)$, $f_2(x_1, x_2; \Psi)$ and $f_3(x; \Psi)$ for $x_1, x_2 \in \mathbb{N}_0$ can be easily obtained by using the relation

$$f_{X_1,X_2}(x_1, x_2; \Psi) = F(x_1, x_2; \Psi) - F(x_1 - 1, x_2; \Psi) - F(x_1, x_2 - 1; \Psi) + F(x_1 - 1, x_2 - 1; \Psi). \tag{5}$$

The plots of the joint PMF are displayed in Figure 1.

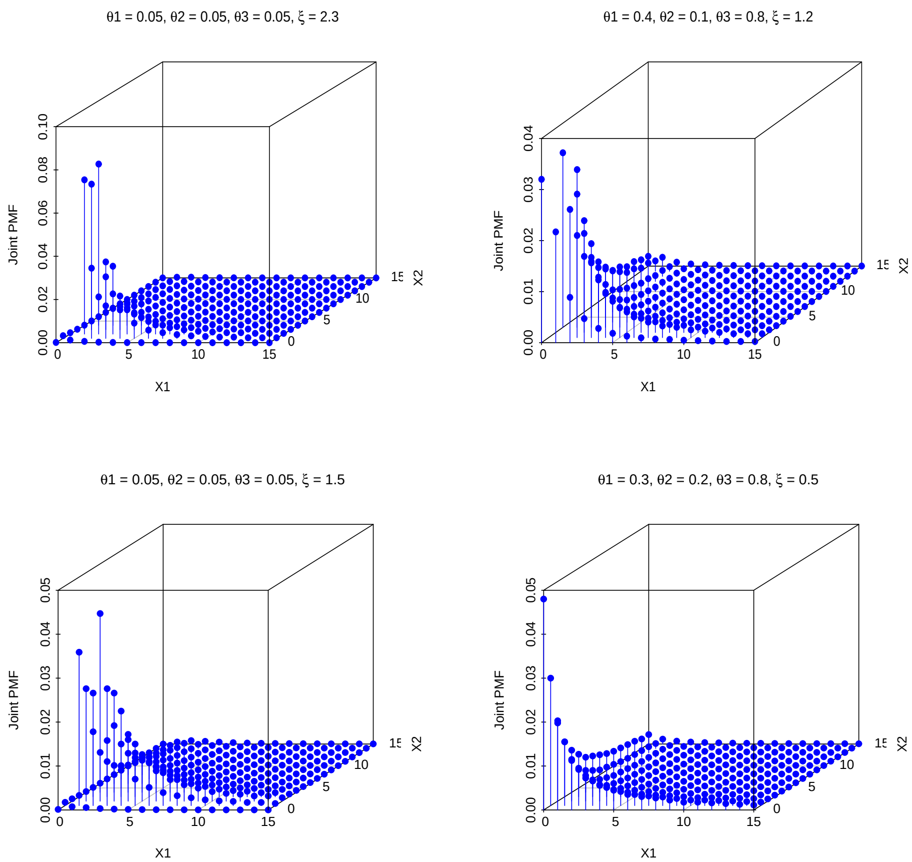

Figure 1. The scatter plots of the joint PMF.

From the plots, it is seen that the joint PMF has a long right tail as compared to its left tail. Thus, it can be used to model skewed data set.

3. Statistical Properties

3.1. The joint RF and joint hazard rate function (HRF)

Assume $\mathbf{X} \sim \mathrm{BDsIW}(\theta_1, \theta_2, \theta_3, \zeta)$, then the joint RF of $\mathbf{X}$ can be expressed as

$$R_{X_1,X_2}(x_1, x_2; \Psi) = 1 - F_{X_1}(x_1; \theta_1\theta_3, \zeta) \; - F_{X_2}(x_2; \theta_2\theta_3, \zeta) + F_{X_1,X_2}(x_1, x_2; \Psi)$$
$$= \begin{cases} R_1(x_1, x_2; \Psi) & ;\ 0 < x_1 < x_2 < \infty \\ R_2(x_1, x_2; \Psi) & ;\ 0 < x_2 < x_1 < \infty \\ R_3(x; \Psi) & ;\ 0 < x_1 = x_2 = x < \infty, \end{cases} \tag{6}$$

where

$$R_1(x_1, x_2; \Psi) \; = 1 - F_{\mathrm{DsIW}}(x_1; \theta_1\theta_3, \zeta) - F_{\mathrm{DsIW}}(x_2; \theta_2\theta_3, \zeta) + F_{\mathrm{DsIW}}\,(x_1; \theta_1\theta_3, \zeta)\; F_{\mathrm{DsIW}}\,(x_2; \theta_2, \zeta)\,,$$
$$R_2(x_1, x_2; \Psi) \; = 1 - F_{\mathrm{DsIW}}(x_1; \theta_1\theta_3, \zeta) - F_{\mathrm{DsIW}}(x_2; \theta_2\theta_3, \zeta) + F_{\mathrm{DsIW}}\,(x_1; \theta_1, \zeta)\; F_{\mathrm{DsIW}}\,(x_2; \theta_2\theta_3, \zeta)$$

and

$$R_3(x; \Psi) \; = 1 - F_{\mathrm{DsIW}}(x; \theta_1\theta_3, \zeta) - F_{\mathrm{DsIW}}(x; \theta_2\theta_3, \zeta) + F_{\mathrm{DsIW}}\,(x; \theta_1\theta_2\theta_3, \zeta)\,.$$

The plots of the joint RF are displayed in Figure 2.

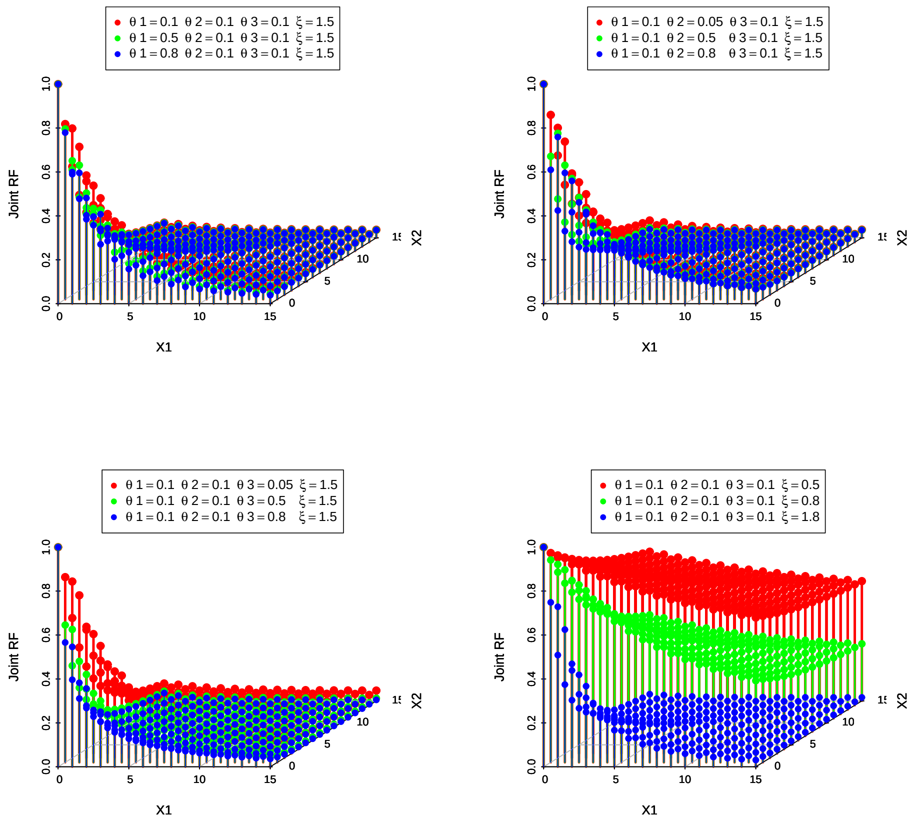

Figure 2. The scatter plots of the joint RF.

The plots indicate that the joint RF decreases in value when the values of $\theta_1, \theta_2, \theta_3$ and ζ increase. The joint HRF of $\mathbf{X}$ is given by

$$r_{X_1,X_2}(x_1, x_2; \Psi) = \begin{cases} r_1(x_1, x_2; \Psi) & ;\ 0 < x_1 < x_2 < \infty \\ r_2(x_1, x_2; \Psi) & ;\ 0 < x_2 < x_1 < \infty \\ r_3(x; \Psi) & ;\ 0 < x_1 = x_2 = x < \infty, \end{cases} \tag{7}$$

where $r_j(x_1, x_2; \Psi) = \frac{f_j(x_1,x_2;\Psi)}{R_j(x_1-1,x_2-1;\Psi)}$; $j = 1, 2, 3$. Figure 3 shows the joint HRF plots of the BDsIW distribution for different parameter values.

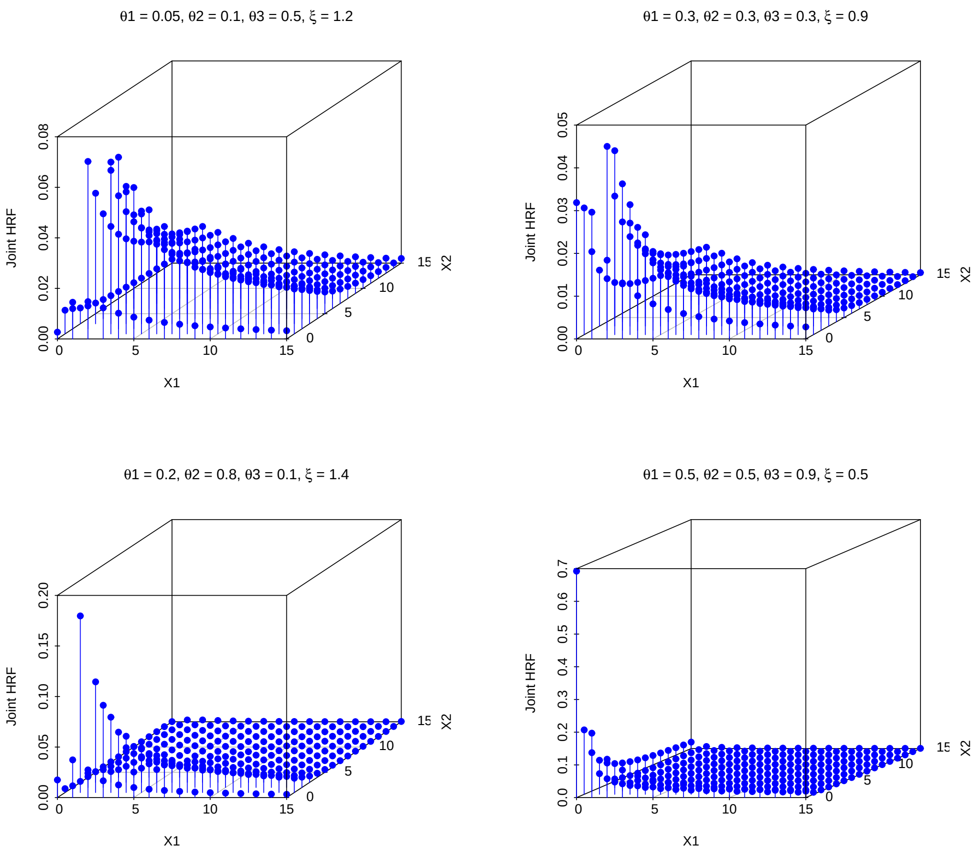

Figure 3. The scatter plots of the joint HRF.

From the plots, it is clear that the joint HRF can take different shapes depending on the parameter values, which make it convenient to discuss different shapes of hazard rate in practice.

3.1.1. The conditional HRF

Assume $\mathbf{X} \sim \text{BDsIW}(\theta_1, \theta_2, \theta_3, \zeta)$ and $X_1 < X_2$. Then, the HRF of the conditional distribution X_1 given $X_2 > x_2$ is given by

$$r^*(X_1|X_2 > x_2) = \frac{\zeta\,(x_1 + 1)^{-\zeta-1}}{R_1(x_1, x_2; \Psi)} \left\{F_{\text{DsIW}}(x_2; \theta_2, \zeta) - 1\right\} F_{\text{DsIW}}(x_1; \theta_1\theta_3, \zeta) \ln(\theta_1\theta_3), \tag{8}$$

while the HRF of the conditional distribution X_2 given $X_1 > x_1$ is given by

$$r^{*}(X_2|X_1 > x_1) = \frac{\zeta\,(x_2+1)^{-\zeta-1}\,F_{\mathrm{DsIW}}\,(x_2;\theta_2,\zeta)}{R_1(x_1,x_2;\Psi)}\left\{F_{\mathrm{DsIW}}\,(x_1;\theta_1\theta_3,\zeta)\ln(\theta_2) - F_{\mathrm{DsIW}}\,(x_2;\theta_3,\zeta)\ln(\theta_2\theta_3)\right\}. \quad (9)$$

Similarly, when $X_2 < X_1$, then

$$r^{**}(X_1|X_2 > x_2) = \frac{\zeta\,(x_1+1)^{-\zeta-1}\,F_{\mathrm{DsIW}}\,(x_1;\theta_1,\zeta)}{R_2(x_1,x_2;\Psi)}\left\{F_{\mathrm{DsIW}}\,(x_2;\theta_2\theta_3,\zeta)\ln(\theta_1) - F_{\mathrm{DsIW}}\,(x_1;\theta_3,\zeta)\ln(\theta_1\theta_3)\right\} \quad (10)$$

and

$$r^{**}(X_2|X_1 > x_1) = \frac{\zeta\,(x_2+1)^{-\zeta-1}}{R_2(x_1,x_2;\Psi)}\left\{F_{\mathrm{DsIW}}\,(x_1;\theta_1,\zeta)-1\right\}F_{\mathrm{DsIW}}\,(x_2;\theta_2\theta_3,\zeta)\ln(\theta_2\theta_3). \quad (11)$$

3.1.2. The joint HRF of a parallel system

Consider a parallel system consists of 2-component. Then, we can defined the joint HRF as a vector which is useful to measure the total life span of a 2-component as follows

$$r(\underline{x}) = (r(x), r_{12}(x_1|x_2), r_{21}(x_2|x_1)), \quad (12)$$

where $r(x)$ gives the HRF of the system using the information that the 2-component has survived beyond x, $r_{12}(x_1|x_2)$ gives the HRF span of the first component given that it has survived to an age x_1 and the other has failed at x_1. Similar argument holds for $r_{21}(x_2|x_1)$ (see Cox, 1972). Thus, if $\mathbf{X} \sim \mathrm{BDsIW}(\theta_1,\theta_2,\theta_3,\zeta)$, then

$$\begin{aligned} r(x)|_{X=\min(X_1,X_2)} = {} & \frac{F_{\mathrm{DsIW}}\,(x-1;\theta_3,\zeta)}{R_3(x;\Psi)}\left[-F_{\mathrm{DsIW}}\,(x-1;\theta_1,\zeta) - F_{\mathrm{DsIW}}\,(x-1;\theta_2,\zeta) + F_{\mathrm{DsIW}}\,(x-1;\theta_1\theta_2,\zeta)\right] \\ & + \frac{F_{\mathrm{DsIW}}\,(x;\theta_3,\zeta)}{R_3(x;\Psi)}\left[F_{\mathrm{DsIW}}\,(x;\theta_1,\zeta) + F_{\mathrm{DsIW}}\,(x;\theta_2,\zeta) - F_{\mathrm{DsIW}}\,(x;\theta_1\theta_2,\zeta)\right], \end{aligned}$$

$$r_{12}(x_1|x_2)|_{X_1>X_2} = -\zeta\,(x_1+1)^{-\zeta-1}\left[1 - F_{\mathrm{DsIW}}\,(x_1;\theta_1,\zeta)\right]^{-1}\ln(\theta_1)$$

and

$$r_{21}(x_2|x_1)|_{X_1<X_2} = -\zeta\,(x_2+1)^{-\zeta-1}\left[1 - F_{\mathrm{DsIW}}\,(x_2;\theta_2,\zeta)\right]^{-1}\ln(\theta_2).$$

3.2. Stress-strength probability

Assume $\mathbf{X} \sim \mathrm{BDsIW}(\theta_1,\theta_2,\theta_3,\zeta)$. Then, the stress-strength probability can be expressed as

$$\mathbf{P}[X_1 < X_2] = \sum_{x=0}^{\infty}(\theta_1\theta_3)^{(x+1)^{-\zeta}}\left[(\theta_2\theta_3)^{(x+1)^{-\zeta}} - (\theta_2\theta_3)^{x^{-\zeta}}\right] \quad (13)$$

and

$$\mathbf{P}[X_1 > X_2] = \sum_{x=0}^{\infty}(\theta_2\theta_3)^{(x+1)^{-\zeta}}\left[(\theta_1\theta_3)^{(x+1)^{-\zeta}} - (\theta_1\theta_3)^{x^{-\zeta}}\right]. \quad (14)$$

Unfortunately, the stress-strength probability does not have a closed-form expression, but it can be easily computed numerically in any symbolic software (e.g. Maple, Matlab or Mathematica). Table 1 reports some numerical values of $\mathbf{P}[X_1 < X_2]$ as an example for different values of the parameters θ_1, θ_2 and θ_3.

Table 1. Some numerical values of $\mathbf{P}[X_1 < X_2]$ at $\zeta = 0.5$.

θ_1	$\theta_2 \downarrow$ $\theta_3 \rightarrow$	**0.2**	**0.4**	**0.6**	**0.8**	**0.99**
0.3	**0.3**	0.5025	0.5131	0.5248	0.5387	0.5543
	0.6	0.4382	0.4286	0.4245	0.4257	0.4322
	0.9	0.3916	0.3641	0.3451	0.3340	0.3311
0.6	**0.3**	0.5752	0.6157	0.6586	0.7068	0.7585
	0.6	0.5131	0.5387	0.5746	0.6231	0.6818
	0.9	0.4671	0.4782	0.5063	0.5532	0.6164
0.9	**0.3**	0.6275	0.6943	0.7663	0.8476	0.9351
	0.6	0.5689	0.6278	0.7041	0.8014	0.9150
	0.9	0.5248	0.5746	0.6524	0.7619	0.8972

Regarding Table 1, it is clear that:

1. For fixed values of θ_1, θ_3 and ζ with $\theta_2 \to 1$, $\mathbf{P}[X_1 < X_2]$ decreases.
2. For fixed values of θ_2, θ_3 and ζ with $\theta_1 \to 1$, $\mathbf{P}[X_1 < X_2]$ increases.

3.3. *The coefficient of median correlation*

Since the median of X_1 and X_2 can be expressed as

$$M_{X_d} = \left\{ \log \frac{\theta_d \theta_3}{U} \right\}^{\frac{1}{\zeta}} - 1;\ d = 1, 2, \tag{15}$$

where U has a uniform $U(0,1)$ distribution, the coefficient of median correlation between X_1 and X_2 can be proposed as

$$M_{X_1,X_2} = \begin{cases} 4F_{\mathrm{DsIW}}\left(M_{X_1}; \theta_1\theta_3, \zeta\right)\ F_{\mathrm{DsIW}}\left(M_{X_2}; \theta_2, \zeta\right) - 1 & ;\ x_1 < x_2 \\ 4F_{\mathrm{DsIW}}\left(M_{X_1}; \theta_1, \zeta\right)\ F_{\mathrm{DsIW}}\left(M_{X_2}; \theta_2\theta_3, \zeta\right) - 1 & ;\ x_2 \le x_1, \end{cases} \tag{16}$$

where $M_{X_1,X_2} = 4F_{X_1,X_2}(M_{X_1}, M_{X_2}) - 1$ (see Domma, 2009).

3.4. *The joint probability generating function (PGF), bivariate skewness and kurtosis*

If the bivariate vector $\mathbf{X} \sim \mathrm{BDsIW}(\theta_1, \theta_2, \theta_3, \zeta)$, then the PGF of X_1 and X_2 can be written as infinite mixtures,

$$\begin{aligned} G_{X_1,X_2}(y_1, y_2) &= \sum_{j,i=0}^{\infty} \mathbf{P}\left[X_1 = i, X_2 = j\right] y_1^i y_2^j \\ &= \sum_{j=0}^{\infty} \sum_{i=0}^{j-1} \left[(\theta_1\theta_3)^{(i+1)^{-\zeta}} - (\theta_1\theta_3)^{i^{-\zeta}} \right] \left[\theta_2^{(j+1)^{-\zeta}} - \theta_2^{j^{-\zeta}} \right] y_1^i y_2^j \\ &+ \sum_{j=0}^{\infty} \sum_{i=j+1}^{\infty} \left[\theta_1^{(i+1)^{-\zeta}} - \theta_1^{i^{-\zeta}} \right] \left[(\theta_2\theta_3)^{(j+1)^{-\zeta}} - (\theta_2\theta_3)^{j^{-\zeta}} \right] y_1^i y_2^j \\ &+ \sum_{i=0}^{\infty} \theta_2^{(i+1)^{-\zeta}} \left[(\theta_1\theta_3)^{(i+1)^{-\zeta}} - (\theta_1\theta_3)^{i^{-\zeta}} \right] y_1^i y_2^i \\ &- \sum_{i=0}^{\infty} (\theta_2\theta_3)^{(i+1)^{-\zeta}} \left[\theta_1^{(i+1)^{-\zeta}} - \theta_1^{i^{-\zeta}} \right] y_1^i y_2^i; \quad |y_1|, |y_2| < 1. \end{aligned} \tag{17}$$

Hence, different moments and product moments of the BDsIW distribution can be obtained, as infinite series, using the joint PGF. The correlation of X_1 and X_2 is the number defined by

$$\rho_{X_1,X_2} = \frac{\mathbf{E}(X_1X_2) - \mathbf{E}(X_1)\mathbf{E}(X_2)}{\sqrt{\mathbf{Var}(X_1)\mathbf{Var}(X_2)}};\ 0 \le \rho_{X_1,X_2} \le 1, \tag{18}$$

where $\mathbf{E}(X_i) = \sum_{x_i=0}^{\infty} x_i\, \pi(x_i; \theta_i\theta_3, \zeta)$ and $\mathbf{Var}(X_i) = \sum_{x_i=0}^{\infty} x_i^2\, \pi(x_i; \theta_i\theta_3, \zeta) - [\mathbf{E}(X_i)]^2$ for $i = 1, 2$. Using Mardia's (1970) measures of bivariate skewness and kurtosis, we get

$$\begin{aligned}\text{Skewness} = \frac{1}{(1-\rho^2)^3}[&\Upsilon_{30}^2 + \Upsilon_{03}^2 + 3\left(1 + 2\rho^2\right)\left(\Upsilon_{12}^2 + \Upsilon_{21}^2\right) - 2\rho^3\Upsilon_{30}\Upsilon_{03} \\ &+ 6\rho\{\Upsilon_{30}(\rho\Upsilon_{12} - \Upsilon_{21}) + \Upsilon_{03}(\rho\Upsilon_{21} - \Upsilon_{12}) - (2+\rho^2)\Upsilon_{21}\Upsilon_{12}\}]\end{aligned} \tag{19}$$

and

$$\text{Kurtosis} = \frac{\Upsilon_{40} + \Upsilon_{04} + 2\Upsilon_{22} + 4\rho\left(\rho\Upsilon_{22} - \Upsilon_{13} - \Upsilon_{31}\right)}{(1-\rho^2)^2}, \tag{20}$$

where $\Upsilon_{wq} = \frac{\mathbf{E}(X_1^w X_2^q)}{\left[\sqrt{\mathbf{Var}(X_1)}\right]^w\left[\sqrt{\mathbf{Var}(X_2)}\right]^q}$. Unfortunately, we cannot get closed-form expressions, but it can be easily computed numerically in any symbolic software (e.g. Maple, Matlab and Mathematica). Table 2 lists some numerical values for both skewness and kurtosis based on different values of the parameters θ_1, θ_2 and θ_3.

Table 2. Some numerical values of skewness (kurtosis) at $\zeta = 1.5$.

θ_1	$\theta_2 \downarrow\ \theta_3 \rightarrow$	**0.3**	**0.6**	**0.9**
0.3	**0.3**	0.2568(1.3680)	0.2499(1.3909)	0.2523(1.3998)
	0.6	0.2549(1.2697)	0.3204(1.2378)	0.4019(1.4967)
	0.9	0.2639(2.0010)	0.2631(1.9975)	0.2398(1.0093)
0.9	**0.3**	0.3690(0.9997)	0.3297(1.6667)	0.4331(1.9761)
	0.6	0.4778(2.3336)	0.2997(1.9987)	0.2368(1.9999)
	0.9	0.0969(1.3251)	0.0984(1.2369)	0.1069(1.1230)

Regarding to Table 2, it is clear that the BDsIW distribution can be used to model skewed data with platykurtic.

3.5. Other properties

1. Assume $(X_{i1}, X_{i2}) \sim \text{BDsIW}(\theta_{i1}, \theta_{i2}, \theta_{i3}, \zeta)$ for $i = 1, 2, ..., n$ and they are independently distributed. If $Z_s = \max(X_{1s}, X_{2s}, ..., X_{ns})$; $s = 1, 2$. Then, $(X_{i1}, X_{i2}) \sim \text{BDsIW}\left(\prod_{i=1}^{n}\theta_{i1}, \prod_{i=1}^{n}\theta_{i2}, \prod_{i=1}^{n}\theta_{i3}, \zeta\right)$.
2. If the bivariate vector $\mathbf{X} \sim \text{BDsIW}(\theta_1, \theta_2, \theta_3, \zeta)$. Then,

a) $\max\{X_1, X_2\} \sim \text{DsIW}(\theta_1\theta_2\theta_3, \zeta)$.

b) The conditional CDF of X_1 given $X_2 \le x_2$, is given by

$$F_{X_1|X_2=x_2}(x_1 \mid x_2) = \begin{cases} \frac{F_{\text{DsIW}}(x_1;\theta_1\theta_3,\zeta)}{F_{\text{DsIW}}(x_2;\theta_3,\zeta)} & \text{if } 0 < x_1 < x_2 < \infty \\ F_{\text{DsIW}}(x_1;\theta_1,\zeta) & \text{if } 0 < x_2 < x_1 < \infty \\ F_{\text{DsIW}}(x;\theta_1,\zeta) & \text{if } 0 < x_1 = x_2 = x < \infty. \end{cases} \tag{21}$$

c) The conditional PMF of X_1 given $X_2 = x_2$, is given by

$$f_{X_1|X_2=x_2}(x_1 \mid x_2) = \begin{cases} f^{(1)}_{X_1|X_2=x_2}(x_1 \mid x_2) & \text{if } 0 < x_1 < x_2 < \infty \\ f^{(2)}_{X_1|X_2=x_2}(x_1 \mid x_2) & \text{if } 0 < x_2 < x_1 < \infty \\ f^{(3)}_{X_1|X_2=x_2}(x_1 \mid x) & \text{if } 0 < x_1 = x_2 = x < \infty, \end{cases} \tag{22}$$

where

$$f^{(1)}_{X_1|X_2=x_2}(x_1 \mid x_2) = \frac{f_{\mathrm{DsIW}}(x_1; \theta_1\theta_3, \zeta)\, f_{\mathrm{DsIW}}(x_2; \theta_2, \zeta)}{f_{\mathrm{DsIW}}(x_2; \theta_2\theta_3, \zeta)},$$

$$f^{(2)}_{X_1|X_2=x_2}(x_1 \mid x_2) = f_{\mathrm{DsIW}}(x_1; \theta_1, \zeta)$$

and

$$f^{(3)}_{X_1|X_2=x_2}(x_1 \mid x) = \frac{F_{\mathrm{DsIW}}(x; \theta_2, \zeta)\, f_{\mathrm{DsIW}}(x; \theta_1\theta_3, \zeta) - F_{\mathrm{DsIW}}(x-1; \theta_2\theta_3, \zeta)\, f_{\mathrm{DsIW}}(x; \theta_1, \zeta)}{f_{\mathrm{DsIW}}(x; \theta_2\theta_3, \zeta)}.$$

The proofs for the previous properties are quite standard, and therefore the details are avoided.

4. Estimation Methods

4.1. Maximum likelihood estimation (MLE) and asymptotic confidence intervals

In this section, we use the maximum likelihood method to estimate the unknown parameters $\theta_1, \theta_2, \theta_3$ and ζ of the BDsIW distribution. Suppose that, we have a sample of size n, of the form $\{(x_{11}, x_{21}), (x_{12}, x_{22}), ..., (x_{1n}, x_{2n})\}$ from the proposed model. We use the following notations: $I_1 = \{x_{1j} < x_{2j}\}$, $I_2 = \{x_{2j} < x_{1j}\}$, $I_3 = \{x_{1j} = x_{2j} = x_j\}$, $I = I_1 \cup I_2 \cup I_3$, $|I_1| = n_1$, $|I_2| = n_2$, $|I_3| = n_3$ and $n = \sum_{k=1}^{3} n_k$. Based on the observations, the likelihood function can be expressed as

$$l(X_1, X_2 | \, \theta_1, \theta_2, \theta_3, \zeta) = \prod_{j=1}^{n_1} f_1(x_{1j}, x_{2j}) \prod_{j=1}^{n_2} f_2(x_{1j}, x_{2j}) \prod_{j=1}^{n_3} f_3(x_j). \tag{23}$$

The log-likelihood function becomes

$$\begin{aligned} L(X_1, X_2 | \, \theta_1, \theta_2, \theta_3, \zeta) = {} & \sum_{j=1}^{n_1} \ln\left(\Phi_1(x_{1j}; \theta_1\theta_3, \zeta)\right) + \sum_{j=1}^{n_1} \ln\left(\Phi_1(x_{2j}; \theta_2, \zeta)\right) \\ & + \sum_{j=1}^{n_2} \ln\left(\Phi_1(x_{1j}; \theta_1, \zeta)\right) + \sum_{j=1}^{n_2} \ln\left(\Phi_1(x_{2j}; \theta_2\theta_3, \zeta)\right) \\ & + \sum_{j=1}^{n_3} \ln\left([\theta_2]^{(x_j+1)^{-\zeta}} \Phi_1(x_j; \theta_1\theta_3, \zeta) - [\theta_2\theta_3]^{(x_j)^{-\zeta}} \Phi_1(x_j; \theta_1, \zeta)\right), \end{aligned} \tag{24}$$

where $\Phi_1(x; \theta, \zeta) = \theta^{(x+1)^{\zeta}} - \theta^{x^{\zeta}}$. The first partial derivatives of Equation (24) with respect to $\theta_1, \theta_2, \theta_3$ and ζ are called the normal equations (see, **Appendix**), where $\Phi_2(x; \theta, \zeta) = x^{-\zeta}\theta^{x^{-\zeta}-1}$ and $\Phi_3(x; \theta, \zeta) = x^{-\zeta}\theta^{x^{-\zeta}-1}\ln(x)\ln(\theta)$. The MLEs of the parameters $\theta_1, \theta_2, \theta_3$ and ζ can be obtained by solving the system of four non-linear equations. These equations cannot be solved analytically; therefore an iterative procedure like Newton Raphson is required to solve them numerically. For the asymptotic confidence interval (CI), The normal approximation of the MLE can be used to construct asymptotic CIs for the parameters $\theta_1, \theta_2, \theta_3$ and ζ, when the sample size is large enough. A two sided $(1-\alpha)100\%$ CIs for $\theta_1, \theta_2, \theta_3$ and ζ are given by $(\widehat{\theta_1} \pm z_{\alpha/2}\sqrt{\mathbf{Var}(\widehat{\theta_1})})$, $(\widehat{\theta_2} \pm z_{\alpha/2}\sqrt{\mathbf{Var}(\widehat{\theta_2})})$, $(\widehat{\theta_3} \pm z_{\alpha/2}\sqrt{\mathbf{Var}(\widehat{\theta_3})})$ and $(\widehat{\zeta} \pm z_{\alpha/2}\sqrt{\mathbf{Var}(\widehat{\zeta})})$ respectively, where $\mathbf{Var}(\widehat{\theta_1})$, $\mathbf{Var}(\widehat{\theta_2})$, $\mathbf{Var}(\widehat{\theta_3})$ and $\mathbf{Var}(\widehat{\zeta})$ are the asymptotic variances of $\widehat{\theta_1}$, $\widehat{\theta_2}$, $\widehat{\theta_3}$ and $\widehat{\zeta}$ respectively.

4.2. Bayesian estimation (BSE) and credible intervals

In order to obtain the Bayesian estimators of the unknown parameters, it is necessary to obtain the likelihood function for the model. Considering the assumptions in the previous subsection, the likelihood function under bivariate complete samples in Equation (24). Consider the BSE under the assumption that the non-negative parameters $\delta_1, \delta_2, \delta_3$ and ζ are independently distributed, which have the gamma prior distribution, where $\theta_i = e^{-\delta_i}$; $i = 1, 2, 3$. Thus, $\pi(\delta_i) \propto \delta_i^{a_i-1} e^{b_i\delta_i}$ and $\pi(\zeta) \propto \zeta^{a_4-1} e^{b_4\zeta}$. All the hyper parameters a_k and b_k are assumed to be known and non-negative where $k = 1, 2, 3, 4$. The posterior distribution can be expressed as follows

$$G(\delta_1, \delta_2, \delta_3, \zeta | X_1, X_2) = \frac{\pi(\delta_1)\pi(\delta_2)\pi(\delta_3)\pi(\zeta)L(X_1, X_2| \delta_1, \delta_2, \delta_3, \zeta)}{\int\limits_{\zeta}\int\limits_{\delta_3}\int\limits_{\delta_2}\int\limits_{\delta_1}\pi(\delta_1)\pi(\delta_2)\pi(\delta_3)\pi(\zeta)L(X_1, X_2| \delta_1, \delta_2, \delta_3, \zeta)\, d\delta_1 d\delta_2 d\delta_3 d\zeta}. \tag{25}$$

Equation (25) can be expressed in a simple form as

$$G(\delta_1, \delta_2, \delta_3, \zeta | X_1, X_2) \propto \pi(\delta_1)\pi(\delta_2)\pi(\delta_3)\pi(\zeta)L(X_1, X_2| \delta_1, \delta_2, \delta_3, \zeta). \tag{26}$$

Thus, the Bayesian estimators of the parameters $\delta_1, \delta_2, \delta_3$ and ζ can be expressed as

$$\widehat{\delta_1} \propto \int\limits_{\delta_1} \delta_1 G(\delta_1, \delta_2, \delta_3, \zeta | X_1, X_2)\, d\delta_1, \ \widehat{\delta_2} \propto \int\limits_{\delta_2} \delta_2 G(\delta_1, \delta_2, \delta_3, \zeta | X_1, X_2)\, d\delta_2,$$
$$\widehat{\delta_3} \propto \int\limits_{\delta_3} \delta_3 G(\delta_1, \delta_2, \delta_3, \zeta | X_1, X_2)\, d\delta_3 \text{ and } \widehat{\zeta} \propto \int\limits_{\zeta} \zeta G(\delta_1, \delta_2, \delta_3, \zeta | X_1, X_2)\, d\zeta.$$

It is not possible to compute analytically the solution for these equations. So that, Markov Chain Monte Carlo (MCMC) approach used to approximate these equations. For the credible intervals, using MCMC techniques, Bayes credible intervals of the parameters $\theta_1, \theta_2, \theta_3$ and ζ can be obtained as

1. Start with initial values $(\theta_1^0, \theta_2^0, \theta_3^0, \zeta^0)$.
2. Generate posterior sample for $\theta_1, \theta_2, \theta_3$ and ζ.
3. Repeat step 2 M times and obtain $(\theta_{11}, \theta_{21}, \theta_{31}, \zeta_1)$, $(\theta_{12}, \theta_{22}, \theta_{32}, \zeta_2)$, ..., $(\theta_{1M}, \theta_{2M}, \theta_{3M}, \zeta_M)$.
4. Arrange $\theta_{1i}, \theta_{2i}, \theta_{3i}$ and ζ_i, in ascending order as

$$\theta_{1[1]}, \theta_{1[2]}, ..., \theta_{1[M]},\ \theta_{2[1]},\ \theta_{2[2]}, ..., \theta_{2[M]}, \theta_{3[1]},\ \theta_{3[2]}, ..., \theta_{3[M]} \text{ and } \zeta_{[1]}, \zeta_{[2]}, ..., \zeta_{[M]}.$$

5. A two-sided $(1-\alpha)100\%$ credible intervals for the unknown parameters $\theta_1, \theta_2, \theta_3$ and ζ are given by

$$(\theta_{1[M_{\alpha/2}]}, \theta_{1[M_{1-\alpha/2}]}),\ (\theta_{2[M_{\alpha/2}]}, \theta_{2[M_{1-\alpha/2}]}),\ (\theta_{3[M_{\alpha/2}]}, \theta_{3[M_{1-\alpha/2}]}) \text{ and } (\zeta_{[M_{\alpha/2}]}, \zeta_{[M_{1-\alpha/2}]}).$$

5. Simulation Results

In this section, the MLE and BSE approaches are used to estimate the parameters $\theta_1, \theta_2, \theta_3$ and ζ of the BDsIW distribution. The population parameters are generated using software **R** package. The sampling distributions are obtained for different sample sizes $n = 50, 100, 150, 250, 350, 400$ from $N = 1000$ replications. This study presents an assessment of the properties for both MLE and BSE techniques in terms of bias and mean square error (MSE). A general form to generate a bivariate vector **X** from the BDsIW distribution is first to generate the value **Y** from the continuous BIW distribution and then to discretize this value to obtain **X**. The MLEs and BSEs are reported in Tables 3 and 4 for BDsIW(0.8, 0.4, 0.4, 0.5) and BDsIW(0.6, 0.25, 0.3, 0.9). For the Bayesian simulation, we assume that all the hyper parameters are equal to 0.01.

Table 3. The bias and MSE values for the BDsIW$(0.8, 0.4, 0.4, 0.5)$.

	Size	θ_1		θ_2		θ_3		ζ	
Method	**n**	**bias**	**MSE**	**bias**	**MSE**	**bias**	**MSE**	**bias**	**MSE**
BSE	**50**	0.044	0.026	0.066	0.027	0.043	0.028	0.041	0.031
	100	0.034	0.024	0.060	0.021	0.036	0.021	0.036	0.026
	150	0.025	0.023	0.049	0.015	0.027	0.016	0.022	0.021
	250	0.019	0.022	0.034	0.012	0.020	0.011	0.018	0.013
	350	0.012	0.015	0.018	0.008	0.013	0.009	0.009	0.007
	400	0.007	0.009	0.008	0.004	0.009	0.005	0.003	0.002
MLE	**50**	0.046	0.029	0.069	0.028	0.045	0.029	0.044	0.036
	100	0.035	0.027	0.065	0.024	0.039	0.022	0.038	0.029
	150	0.027	0.022	0.052	0.016	0.036	0.018	0.028	0.024
	250	0.017	0.018	0.049	0.015	0.029	0.010	0.019	0.019
	350	0.016	0.016	0.033	0.011	0.016	0.009	0.011	0.012
	400	0.009	0.013	0.017	0.010	0.012	0.007	0.008	0.008

Table 4. The bias and MSE values for the BDsIW$(0.6, 0.25, 0.3, 0.9)$.

	Size	θ_1		θ_2		θ_3		ζ	
Method	**n**	**bias**	**MSE**	**bias**	**MSE**	**bias**	**MSE**	**bias**	**MSE**
MLE	**50**	0.041	0.034	0.027	0.032	0.029	0.019	0.025	0.017
	100	0.036	0.022	0.025	0.021	0.022	0.015	0.018	0.015
	150	0.027	0.017	0.020	0.013	0.018	0.013	0.012	0.013
	250	0.013	0.011	0.014	0.009	0.016	0.011	0.009	0.010
	350	0.011	0.008	0.012	0.007	0.014	0.010	0.007	0.008
	400	0.006	0.005	0.009	0.006	0.012	0.008	0.004	0.003
BSE	**50**	0.039	0.029	0.026	0.033	0.019	0.017	0.021	0.016
	100	0.032	0.023	0.025	0.028	0.016	0.015	0.019	0.015
	150	0.028	0.015	0.018	0.022	0.015	0.012	0.011	0.014
	250	0.011	0.014	0.017	0.017	0.013	0.011	0.010	0.010
	350	0.010	0.012	0.011	0.015	0.010	0.008	0.009	0.003
	400	0.009	0.004	0.010	0.004	0.008	0.006	0.003	0.001

From Tables 3 and 4, the following observations can be made:

1. The magnitude of bias always decreases to zero as $n \to \infty$.
2. The MSEs always decrease to zero as $n \to \infty$. This shows the consistency of the estimators.
3. The estimators of $\theta_1, \theta_2, \theta_3$ and ζ are slightly positive biased.
4. The MLE and BSE can be used quite effectively for data analysis purposes.

6. Data Analysis

In this section, we explain the experimental importance of the BDsIW distribution using three applications to real data sets. In each data, we shall compare the fits of the BDsIW distribution with some competitive models like: bivariate Poisson (BPo), bivariate binomial (BBi), bivariate Poisson Holgate (BPoH), independent bivariate Poisson (IBPo), bivariate discrete inverse exponential (BDsIE) and bivariate discrete inverse Rayleigh (BDsIR) distributions. The tested distributions are compared using some criteria namely, the maximized log-likelihood (L), Akaike information criterion (AIC), corrected Akaike information criterion (CAIC), Bayesian information criterion (BIC), Hannan-Quinn information criterion (HQIC), and likelihood ratio test statistic (Λ).

6.1. *The first data set (I): Motor data*

This data is reported in Relia staff (2002), and it represents the failure times of a parallel system constituted by two identical motors in days. Figure 4 plots the failure times (x_1, x_2) of the *ith* system $i = 1, ..., 18$.

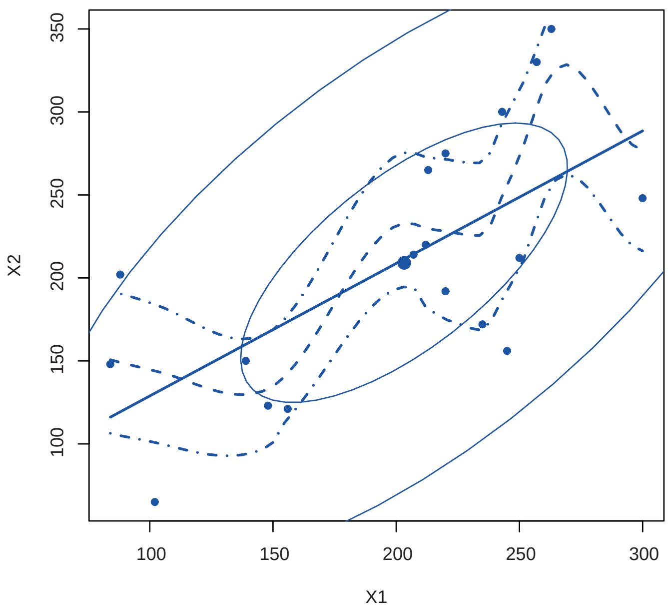

Figure 4. Failure times (in days) of motors A and B.

Before trying to analyze the data using the BDsIW distribution, we fit at first the marginals X_1 and X_2 separately as well as $\min(X_1, X_2)$. The MLEs of the parameters θ and ζ of the corresponding DsIW distribution for X_1, X_2 and $\min(X_1, X_2)$ are $(0.267, 0.259)$, $(0.011, 0.501)$ and $(0.099, 0.518)$ respectively. The $-L$ values are 144.077, 125.633 and 122.330, respectively. Moreover, the p-values ranged from 0.698 to 0.785. It is clear that the DsIW distribution fits the data for the marginals. Table 5 lists some descriptive statistics for data set I using the marginals of the BDsIW distribution.

Table 5. Summary statistics relative to data set I using the marginals.

	Mean		Variance	Skewness		Kurtosis	
Marginal ↓ Measure →	Statistics	Std.err	Statistics	Statistics	Std.err	Statistics	Std.err
X_1	199.000	15.116	4113.294	−0.591	0.536	−0.701	1.038
X_2	207.944	18.036	5855.585	0.181	0.536	−0.448	1.038

Regarding Table 5, it is clear that the variance is greater than the mean. Hence, this data represents over-dispersed data. Here, we fit the BDsIW distribution on this data. The MLEs, $-L$, AIC, CAIC, BIC, and

HQIC values are reported in Table 6.

Table 6. The MLEs, $-L$, AIC, CAIC, BIC, and HQIC values for data set I.

Model	MLEs	−L	AIC	CAIC	BIC	HQIC
BPo	$\widehat{\lambda}_1 = 189.1, \widehat{\lambda}_2 = 190.4, \widehat{\alpha}_1 = 0.165, \widehat{\alpha}_2 = 0.079$	472.05	952.11	955.18	955.67	952.59
BBi	$\widehat{p}_1 = 0.004, \widehat{p}_2 = 0.004, \widehat{n}_1 = 43652, \widehat{n}_2 = 43985,$ $\widehat{\alpha}_1 = 0.185, \widehat{\alpha}_2 = 0.081$	473.28	958.56	966.19	963.9	959.29
BPoH	$\widehat{a} = 199.01, \widehat{b} = 207.94, \widehat{d} = 81.10$	513.96	1033.92	1035.63	1036.59	1034.29
IBPo	$\widehat{\lambda}_1 = 1.876, \widehat{\lambda}_2 = 1.365$	370.39	744.78	745.58	746.56	745.03
BDsIE	$\widehat{\theta}_1 = 0.06, \widehat{\theta}_2 = 0.052, \widehat{\theta}_3 = 0.184$	378.47	762.94	764.65	765.61	763.31
BDsIR	$\widehat{\theta}_1 = 0.007, \widehat{\theta}_2 = 0.005, \widehat{\theta}_3 = 0.119$	519.61	1045.22	1046.93	1047.89	1045.59
BDsIW	$\widehat{\theta}_1 = 0.279, \widehat{\theta}_2 = 0.247, \widehat{\theta}_3 = 0.212, \widehat{\zeta} = 0.388$	**329.98**	**667.96**	**671.04**	**671.52**	**668.45**

According to Table 6, it is clear that the BDsIW distribution provides a better fit than other tested distributions, because it has the smallest values among $-L$, AIC, CAIC, BIC and HQIC. The profiles of the log-likelihood function for its parameters are unimodal functions. So, the likelihood equations have a unique solution. **The approximate 95% two sided confidence interval (CI)** of the parameters $\theta_1, \theta_2, \theta_3$ and ζ are given respectively as $[0.200, 0.301]$, $[0.199, 0.322]$, $[0.195, 0.267]$ and $[0.266, 0.420]$. Since, the BDsIE and BDsIR distributions can be obtained as special cases from the BDsIW distribution. Hence, we want to perform the following two tests:

Test 1: $H_{01} : \zeta = 1$ (BDsIE) against $H_{11} : \zeta \neq 1$ (BDsIW).

Test 2: $H_{02} : \zeta = 2$ (BDsIR) against $H_{12} : \zeta \neq 2$ (BDsIW).

The Λ, degree of freedom (d.f) and p-values for the BDsIE and BDsIR distributions are reported in Table 7 using data set I.

Table 7. The Λ, d.f and p-values using data set I.

Model	$H_\circ$	Λ	d.f.	p-values
BDsIE	$\zeta = 1$	96.98	1	< 0.01
BDsIR	$\zeta = 2$	379.26	1	< 0.01

Table 7 shows that H_{02} and H_{03} are rejected with 5% level of significance. So, we prefer the BDsIW distribution for analyzing this data. Recall Equations (18), (19) and (20), Table 8 summaries statistics for data set I using the BDsIW distribution.

Table 8. Summary statistics relative to data set I.

Model ↓ Measure →	Correlation	Skewness	Kurtosis
BDsIW	0.698	0.597	0.633

It is clear that the correlation between X_1 and X_2 is positive and strong. Also, the skewness and kurtosis are positive. Recall Equation (12), Figure 5 shows the hazard rate vector using data set I.

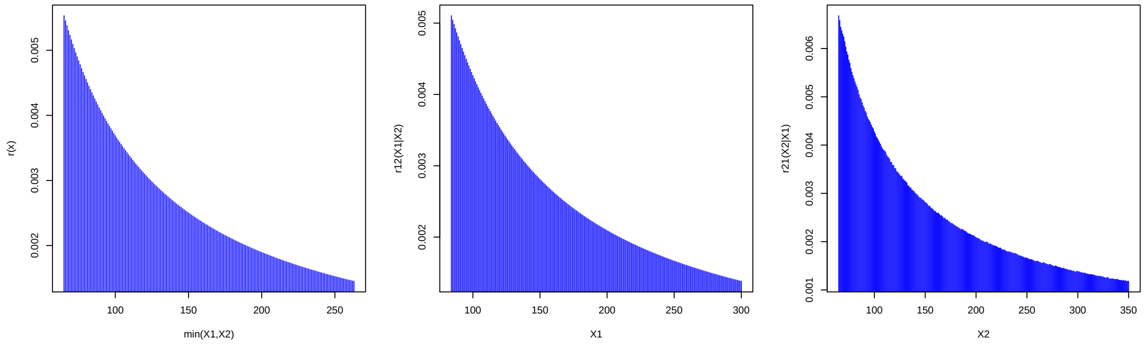

Figure 5. The hazard rate of min (X_1, X_2) in (left panel), $r_{12}(x_1|x_2)|_{X_1>X_2}$ in (middle panel) and $r_{21}(x_2|x_1)|_{X_1<X_2}$ in (right panel) using data set I.

It is clear that the hazard rate of min (X_1, X_2), $r_{12}(x_1|x_2)|_{X_1>X_2}$, and $r_{21}(x_2|x_1)|_{X_1<X_2}$ decreases.

6.2. *The second data set (II): Football data*

This data is reported in Lee and Cha (2015), and it represents a football match score in Italian football match (Serie A) during 1996 to 2011, between ACF Fiorentina(X_1) and Juventus(X_2). Figure 6 plots the football match score data.

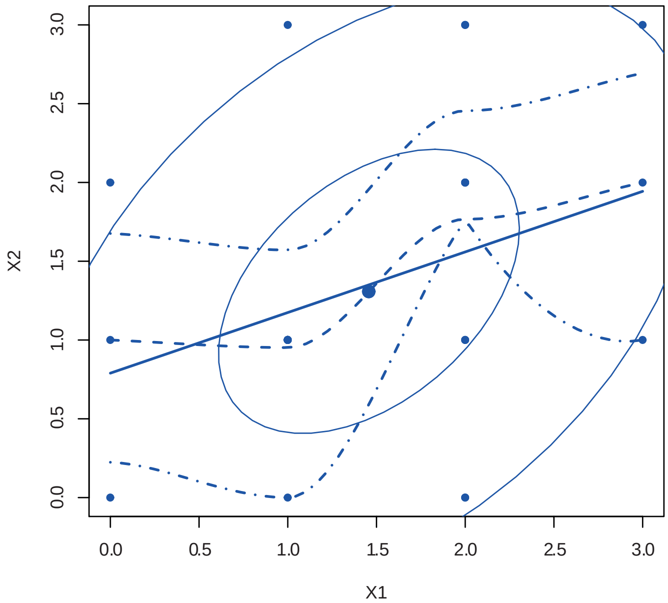

Figure 6. Scatter plot of the football match data.

Here, we fit at first the marginals X_1 and X_2 separately and the min(X_1, X_2). The MLEs of the parameters θ and ζ of the corresponding DsIW distribution for X_1, X_2 and min(X_1, X_2) are $(0.237, 2.798)$, $(0.095, 2.601)$ and $(0.310, 3.103)$ respectively. The $-L$ values are 30.86, 33.73 and 28.02 respectively. Moreover, the p-values ranged from 0.674 to 0.873. Figure 7 shows the estimated PMF plots for the marginals X_1, X_2 and $\min(X_1, X_2)$ using data set II.

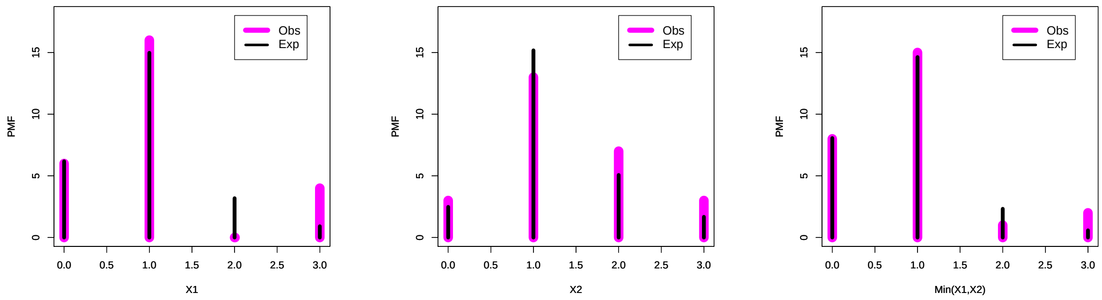

Figure 7. The estimated PMF for the marginals X_1, X_2 and $\min(X_1, X_2)$ using data set II.

Depending on Figure 7 and p-values, it is clear that the DsIW distribution fits the data for the marginals. Table 9 lists some descriptive statistics for data set II using the marginals of the BDsIW distribution.

Table 9. Summary statistics relative to data set II using the marginals of the BDsIW distribution.

	Mean		Variance	Skewness		Kurtosis	
Marginal ↓ Measure →	Statistics	Std.err	Statistics	Statistics	Std.err	Statistics	Std.err
X_1	1.500	0.149	0.672	0.000	0.427	−0.347	0.833
X_2	1.367	0.162	0.792	0.433	0.427	−0.374	0.833

Regarding Table 9, it is clear that the variance is smaller than the mean. Hence, this data represents under-dispersed data. Here, we fit the BDsIW distribution on this data. The MLEs, $-L$, AIC, CAIC, BIC, and HQIC values for the tested bivariate models are reported in Table 10.

Table 10. The MLEs, $-L$, AIC, CAIC, BIC, and HQIC values for data set II.

Model	MLEs	–L	AIC	CAIC	BIC	HQIC
BPo	$\widehat{\lambda_1} = 1.08, \widehat{\lambda_2} = 1.75, \widehat{\alpha_1} = 1.01, \widehat{\alpha_2} = 4.02$	62.52	133.05	134.94	138.07	134.49
BBi	$\widehat{p_1} = 0.012, \widehat{p_2} = 0.018, \widehat{n_1} = 4325, \widehat{n_2} = 4215,$ $\widehat{\alpha_1} = 0.174, \widehat{\alpha_2} = 0.138$	63.72	139.44	143.86	146.99	141.61
BPoH	$\widehat{a} = 1.08, \widehat{b} = 1.38, \widehat{c} = 0.70$	64.92	135.83	136.93	139.61	136.93
IBPo	$\widehat{\lambda_1} = 1.08, \widehat{\lambda_2} = 1.38$	67.60	139.21	139.72	141.72	139.92
BDsIE	$\widehat{\theta_1} = 0.669, \widehat{\theta_2} = 0.388, \widehat{\theta_3} = 0.514$	78.54	163.07	163.99	167.28	164.42
BDsIR	$\widehat{\theta_1} = 0.493, \widehat{\theta_2} = 0.212, \widehat{\theta_3} = 0.561$	64.10	134.21	135.29	137.98	135.29
BDsIW	$\widehat{\theta_1} = 0.420, \widehat{\theta_2} = 0.141, \widehat{\theta_3} = 0.587, \widehat{\zeta} = 2.738$	**61.96**	**131.82**	**133.82**	**136.95**	**133.37**

In Table 10, it is observed that the BDsIW distribution provides a better fit than other tested distributions, because it has the smallest values among $-L$, AIC, CAIC, BIC and HQIC. The profiles of the log-likelihood function for its parameters are unimodal functions. **The approximate 95% CI** of the parameters $\theta_1, \theta_2, \theta_3$ and ζ are given respectively as [0.171, 0.521], [0, 0.274], [0.181, 0.641] and [1.378, 3.495]. Here, we want to perform the following two tests:

Test 1: $H_{01} : \zeta = 1$ (BDsIE) against $H_{11} : \zeta \neq 1$ (BDsIW).
Test 2: $H_{02} : \zeta = 2$ (BDsIR) against $H_{12} : \zeta \neq 2$ (BDsIW).

The Λ, d.f and p-values for the BDsIE and BDsIR distributions are reported in Table 11 using data set II.

Table 11. The Λ, d.f and p-values for data set II.

Model	$H_\circ$	Λ	**d.f**	**p-values**
BDsIE	$\zeta = 1$	33.152	1	< 0.01
BDsIR	$\zeta = 2$	4.288	1	0.03

We can conclude that H_{01} and H_{02} are rejected with 5% level of significance. Hence, the BDsIE and BDsIR distributions cannot be used for this data set. So, we prefer the BDsIW distribution for analyzing this data. Figure 8 shows the estimated joint PMF for BDsIW, BDsIE, and BDsIR distributions, which support the results of Table 11.

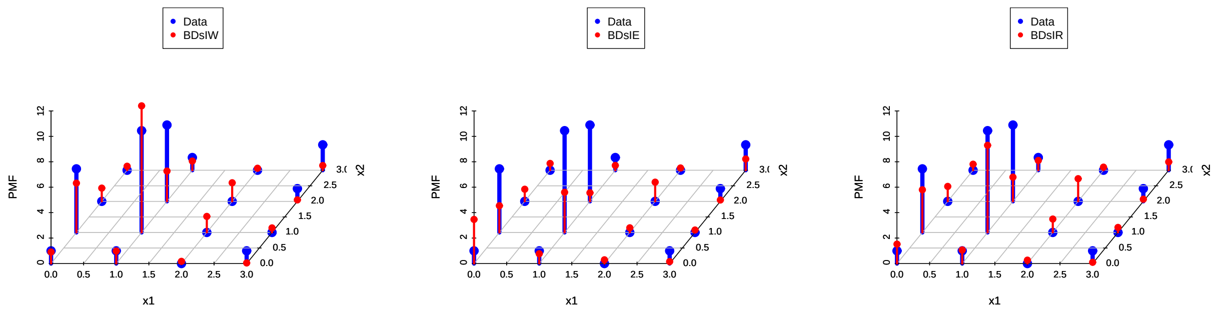

Figure 8. The estimated joint PMF for the BDsIW, BDsIE and BDsIR distributions using data set II.

Table 12 summaries statistics for data set II using the BDsIW distribution.

Table 12. Summary statistics relative to data set II.

Model $\downarrow$ **Measure** $\rightarrow$	**Correlation**	**Skewness**	**Kurtosis**
BDsIW	0.524	0.289	0.694

It is clear that the correlation between X_1 and X_2 is positive and strong. Further, the skewness and kurtosis are positive.

6.3. *The third data set (III): Nasal drainage severity score data*

This data is reported in Davis (2002), and it represents the efficacy of steam inhalation in the treatment of common cold symptoms. Figure 9 plots the Nasal drainage severity score data.

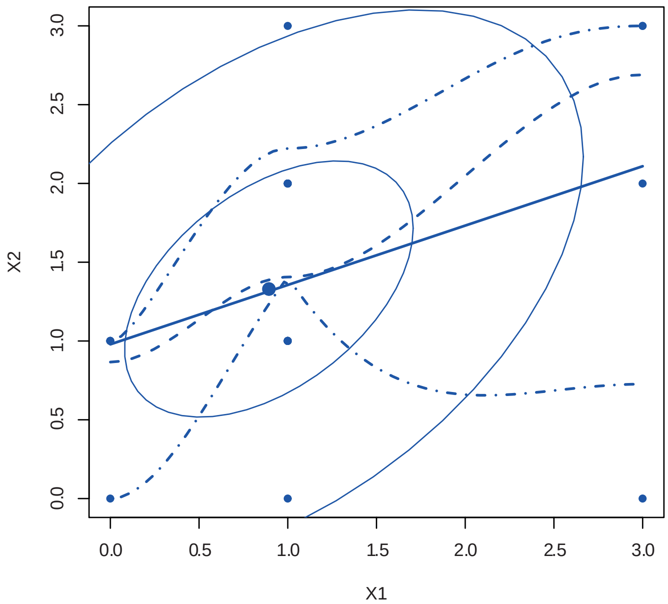

Figure 9. Scatter plot of the Nasal drainage severity score data.

We fit at first the marginals X_1 and X_2 separately and $\min(X_1, X_2)$. The MLEs of the parameters θ and ζ of the corresponding DsIW distribution for X_1, X_2 and $\min(X_1, X_2)$ are (0.065, 2.505), (0.115, 2.524) and (0.181, 2.699) respectively. The $-L$ values are 40.99, 39.83 and 36.68 respectively. Moreover, the p-values ranged from 0.728 to 0.896. Figure 10 shows the estimated PMF plots for the marginals X_1, X_2 and $\min(X_1, X_2)$ using data set III.

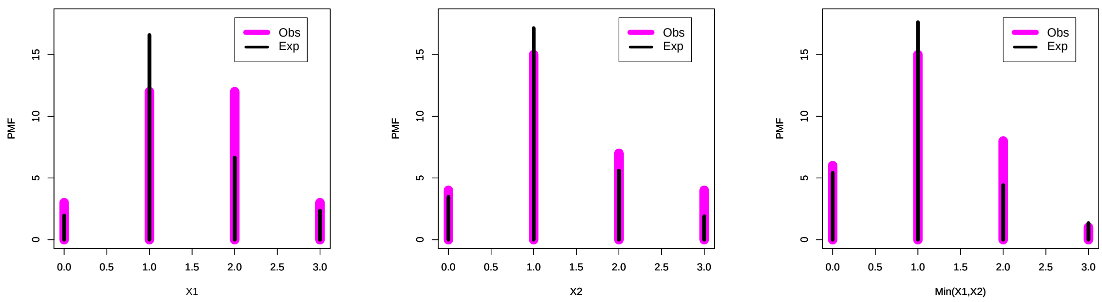

Figure 10. The estimated PMF for the marginals X_1, X_2 and $\min(X_1, X_2)$ using data set III.

From Figure 10 and p-values, it is clear that the DsIW distribution fits the data for the marginals. Table 13 lists some descriptive statistics for data set III using the marginals of the BDsIW distribution.

Table 13. Summary statistics relative to data set III using the marginals of the BDsIW distribution.

	Mean		Variance	Skewness		Kurtosis	
Marginal ↓ Measure →	Statistics	Std.err	Statistics	Statistics	Std.err	Statistics	Std.err
X_1	1.076	0.183	0.874	1.111	0.456	0.785	0.887
X_2	1.384	0.167	0.726	0.390	0.456	−0.238	0.887

Regarding Table 13, it is clear that the variance is smaller than the mean. Hence, this data represents under-dispersed data. Now, we fit BDsIW distribution on this data. The MLEs, $-L$, AIC, CAIC, BIC, and HQIC values for the tested bivariate models are reported in Table 14.

Table 14. The MLEs, $-L$, AIC, CAIC, BIC and HQIC values for data set III.

Model	MLEs	−L	AIC	CAIC	BIC	HQIC
BPo	$\widehat{\lambda}_1 = 0.262, \widehat{\alpha}_1 = 0.165, \widehat{\lambda}_2 = 0.405, \widehat{\alpha}_2 = 2.97$	77.66	163.33	164.93	168.93	164.66
BBi	$\widehat{p}_1 = 0.145, \widehat{p}_2 = 0.122, \widehat{n}_1 = 452, \widehat{n}_2 = 468,$ $\widehat{\alpha}_1 = 0.396, \widehat{\alpha}_2 = 0.748$	81.14	174.28	177.93	182.69	176.97
BPoH	$\widehat{a} = 1.369, \widehat{b} = 2.314, \widehat{c} = 0.396$	82.36	170.72	171.64	174.92	172.06
IBPo	$\widehat{\lambda}_1 = 1.499, \widehat{\lambda}_2 = 1.367$	80.66	165.32	165.76	168.12	166.21
BDsIE	$\widehat{\theta}_1 = 0.501, \widehat{\theta}_2 = 0.622, \widehat{\theta}_3 = 0.383$	92.48	190.96	191.88	195.16	192.30
BDsIR	$\widehat{\theta}_1 = 0.262, \widehat{\theta}_2 = 0.405, \widehat{\theta}_3 = 0.363$	78.66	163.32	164.24	167.52	164.66
BDsIW	$\widehat{\theta}_1 = 0.192, \widehat{\theta}_2 = 0.337, \widehat{\theta}_3 = 0.360, \widehat{\zeta} = 2.453$	**76.51**	**161.02**	**162.62**	**166.62**	**162.81**

From Table 14, it is clear that BDsIW distribution provides a better fit than other tested distributions. The profiles of the log-likelihood function for its parameters are unimodal functions. **The approximate 95% two sided CI** of the parameters $\theta_1, \theta_2, \theta_3$ and ζ are given respectively as $[0, 0.301]$, $[0.105, 0.492]$, $[0.116, 0.451]$ and $[1.120, 2.974]$. Here, we want to perform the following two tests:

Test 1: $H_{01} : \zeta = 1$ (BDsIE) against $H_{11} : \zeta \neq 1$ (BDsIW).

Test 2: $H_{02} : \zeta = 2$ (BDsIR) against $H_{12} : \zeta \neq 2$ (BDsIW).

Table 15 shows the Λ and p-values for BDsIE and BDsIR distributions using data set III.

Table 15. The Λ, d.f and p-values for data set III.

Model	$H_\circ$	Λ	d.f.	p-values
BDIE	$\zeta = 1$	31.94	1	< 0.01
BDIR	$\zeta = 2$	4.3	1	0.03

From Table 15, we can conclude that H_{01} and H_{02} are rejected with 5% level of significance. So, we prefer the BDsIW distribution for analyzing this data. Figure 11 shows the estimated joint PMF for the BDsIW, BDsIE and BDsIR distributions, which support the results of Table 15.

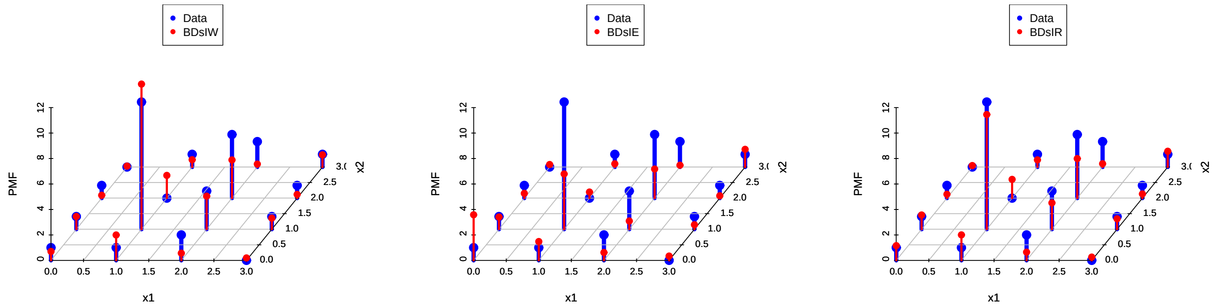

Figure 11. The estimated joint PMF for the BDsIW, BDsIE and BDsIR distributions using data set III.

Table 16 summaries statistics for data set III using the BDsIW distribution.

Table 16. Summary statistics relative to data set III.

Model ↓ Measure →	Correlation	Skewness	Kurtosis
BDsIW	0.403	0.699	0.989

It is clear that the correlation between X_1 and X_2 is positive and weak. Moreover, the skewness and kurtosis are positive. In Table 17 the results of BSE for the real data sets are listed.

Table 17. Posterior summaries for real data sets

Data	Parameters	Mean	MSE	MC Error	Median	Credible intervals
I	θ_1	0.272	0.198	0.0011	0.275	$[0.195, 0.299]$
	θ_2	0.248	0.201	0.0009	0.246	$[0.183, 0.256]$
	θ_3	0.210	0.199	0.0013	0.212	$[0.129, 0.256]$
	ζ	0.390	0.209	0.0028	0.389	$[0.266, 0.453]$
II	θ_1	0.410	0.214	0.0031	0.417	$[0.301, 0.502]$
	θ_2	0.140	0.247	0.0011	0.101	$[0.056, 0.188]$
	θ_3	0.577	0.176	0.0029	0.566	$[0.430, 0.679]$
	ζ	2.729	0.248	0.0030	2.726	$[2.302, 3.168]$
III	θ_1	0.192	0.064	0.0018	0.041	$[0.111, 0.231]$
	θ_2	0.339	0.107	0.0024	0.335	$[0.301, 0.399]$
	θ_3	0.358	0.105	0.0016	0.355	$[0.301, 0.413]$
	ζ	2.450	0.124	0.0010	2.446	$[2.103, 3.127]$

The results presented in Table 17 are very similar to the MLE results.

7. Conclusions

In this paper, we have proposed a flexible bivariate discrete distribution, in the so-called BDsIW distribution. The proposed model has the marginals, which are discrete inverse Weibull distributions. The joint CDF and joint PMF have simple forms. Therefore, the new bivariate discrete distribution can be easily used in practice for modeling bivariate discrete data. Some statistical properties have been discussed in detail. Moreover, the simulation results have indicated that the MLE and BSE work quite satisfactorily and its can be used to estimate the model parameters. Finally, three real data sets have been analyzed, and it is observed that the BDsIW distribution is better than the BPo, BBi, BPoH, IBPo, BDsIE and BDsIR distributions to analyze the data considered herein.

References

[1] Almalki, S. J., and Yuan, J. (2013). A new modified Weibull distribution. Reliability engineering and system safety, 111, 164-170.
[2] Cox, D. R., (1972). Regression models and life tables. Royal statistics society, 34, 187-220.
[3] Davis, C. S., (2002). Statistical methods for the analysis of repeated measures data. Springer-Verlag, New York.
[4] Domma, F., (2009). Some properties of the bivariate Burr type III distribution. Statistics. DOI: 10.1080/02331880902986547.
[5] Eliwa, M. S., Alhussain, Z. A., Ahmed. E. A., Salah, M. M., Ahmed, H. H., and El-Morshedy, M., (2020a). Bivariate Gompertz generator of distributions: statistical properties and estimation with application to model football data. Journal of the national science foundation of Sri Lanka, 48(2), 149-162.
[6] Eliwa, M. S., Alhussain, Z. A., and El-Morshedy, M., (2020b). Discrete Gompertz-G family of distributions for over- and under-dispersed data with properties, estimation, and applications. Mathematics, 8, 358.
[7] Eliwa, M. S., and El-Morshedy, M., (2019). Bivariate Gumbel-G family of distributions: statistical properties, Bayesian and non-Bayesian estimation with application. Annals of data science, 6(1), 39-60.
[8] Eliwa, M. S., and El-Morshedy, M., (2020). Bivariate odd Weibull-G family of distributions: Properties, Bayesian and non-Bayesian estimation with bootstrap confidence intervals and application. Journal of Taibah University for science, 14, 1–22.
[9] El-Morshedy, M., Alhussain, Z. A., Atta, D., Almetwally, E. M., and Eliwa, M. S., (2020b). Bivariate Burr X generator of distributions: properties and estimation methods with applications to complete and Type-II censored samples. Mathematics, 8(2), 264.
[10] El-Morshedy, M., and Eliwa, M. S., (2019). The odd flexible Weibull-H family of distributions: properties and estimation with applications to complete and upper record data. Filomat, 33(9), 2635-2652.
[11] El-Morshedy, M., Eliwa, M. S., and Afify, A. Z. (2020a). The odd Chen generator of distributions: properties and estimation methods with applications in medicine and engineering. Journal of the national science foundation of Sri Lanka, 48(2), 113-130.
[12] El-Morshedy, M., Eliwa, M. S., and Nagy, H., (2020c). A new two-parameter exponentiated discrete Lindley distribution: properties, estimation and applications. Journal of applied statistics, 47(2), 354-375.
[13] El-Morshedy, M., Eliwa, M. S., El-Gohary, A., and Khalil, A. A., (2020d). Bivariate exponentiated discrete Weibull distribution: statistical properties, estimation, simulation and applications. Mathematical sciences, 14, 29-42. https://doi.org/10.1007/s40096-019-00313-9.
[14] Erdem, E., and Shi, J. (2011). Comparison of bivariate distribution construction approaches for analysing wind speed and direction data. Wind energy, 14(1), 27-41.
[15] Jazi, M. A., Lai, C. D., and Alamatsaz, M. H. (2010). A discrete inverse Weibull distribution and estimation of its parameters. Statistical methodology, 7(2), 121-132.
[16] Jehhan, A., Mohamed, I., Eliwa, M. S., Al-mualim, S., and Yousof, H. M., (2018). The two-parameter odd Lindley Weibull lifetime model with properties and applications. International journal of statistics and probability, 7, 57-68.
[17] Jia, J. M., Yan, Z. Z., and Peng, X. Y. (2019). A new discrete extended weibull distribution. IEEE access, 7, 175474-175486.
[18] Kundu, D., and Gupta, R. D., (2009). Bivariate generalized exponential distribution. Journal of multivariate analysis, 100, 581-593.
[19] Kundu, D., and Nekoukhou, N., (2018). Univariate and bivariate geometric discrete generalized exponential distributions. Journal of statistical theory and practice. DOI:10.1080/15598608.2018.1441082.
[20] Lee, H., and Cha, J. H., (2015). On two general classes of discrete bivariate distributions. The American statistician, 69(3), 221-230.
[21] Li, Y., Hui, G., Zhao, Z., and Hu, Y. (2012). The bivariate distribution characteristics of spatial structure in natural K orean pine broad-leaved forest. Journal of vegetation science, 23(6), 1180-1190.
[22] Mardia, K. V., (1970). Measures of multivariate skewness and kurtosis with applications. Biometrika, 57(3), 519-530.
[23] Morata, L. B., (2009). A priori ratemaking using bivariate Poisson regression models. Insurance: mathematics and economics, 44, 135-141.
[24] Nekoukhou, N., and Kundu, D., (2017). Bivariate discrete generalized exponential distribution. Statistics, 51, 1143-1158.
[25] Nekoukhou, V., and Bidram, H. (2015). The exponentiated discrete Weibull distribution. Sort, 39, 127-146.
[26] Olkin, I., and Liu, R. (2003). A bivariate beta distribution. Statistics and probability letters, 62(4), 407-412.
[27] Relia Soft's R and D staff., (2002). Using QALT models to analyze system configurations with load sharing. Reliability Edge, 3, 1-4. Http://www.reliasoft.com/pubs/reliabilityedge_4q2002.pdf
[28] Salah, M. M., El-Morshedy, M., Eliwa, M. S., and Yousof, H. M. (2020). Expanded Fréchet model: mathematical properties, copula, different estimation methods, applications and validation testing. Mathematics, 8(11), 1949.
[29] Shakhatreh, M. K., Lemonte, A. J., and Moreno-Arenas, G. (2019). The log-normal modified Weibull distribution and its reliability implications. Reliability engineering and system safety, 188, 6-22.
[30] Silva, G. O., Ortega, E. M., and Cordeiro, G. M. (2010). The beta modified Weibull distribution. Lifetime data analysis, 16(3), 409-430.
[31] Tahir, M. H., Hussain, M. A., Cordeiro, G. M., El-Morshedy, M., and Eliwa, M. S. (2020). A new Kumaraswamy generalized family of distributions with properties, applications, and bivariate extension. Mathematics, 8(11), 1989.
[32] Wang, M., and Rennolls, K. (2007). Bivariate distribution modeling with tree diameter and height data. Forest science, 53(1), 16-24.
[33] Wu, X., and Yuen, K. C., (2003). A discrete-time risk model with interaction between classes of business. Insurance: mathematics and economics, 33, 117-133.
[34] Yuen, K. C., Guo, J., and Wu, X., (2006). On the first time of ruin in the bivariate compound Poisson model. Insurance: mathematics and economics, 38, 298-308.

Appendix

The normal equations obtaining from likelihood functions can be expressed as follows:

$$\frac{\partial L}{\partial \theta_1}=\sum_{j=1}^{n_1}\frac{\theta_3\Phi_2(x_{1j}+1;\theta_1\theta_3,\zeta)-\theta_3\Phi_2(x_{1j};\theta_1\theta_3,\zeta)}{\Phi_1(x_{1j};\theta_1\theta_3,\zeta)}+\sum_{j=1}^{n_2}\frac{\Phi_2(x_{1j}+1;\theta,\zeta)-\Phi_2(x_{1j};\theta,\zeta)}{\Phi_1(x_{1j};\theta_1,\zeta)}+$$
$$\sum_{j=1}^{n_3}\frac{\theta_3\theta_2^{(x_j+1)^{-\zeta}}\left[\Phi_2(x_j+1;\theta_1\theta_3,\zeta)-\Phi_2(x_j;\theta_1\theta_3,\zeta)\right]-(\theta_2\theta_3)^{(x_j)^{-\zeta}}\left[\Phi_2(x_j+1;\theta_1,\zeta)-\Phi_2(x_j;\theta_1,\zeta)\right]}{\theta_2^{(x_j+1)^{-\zeta}}\Phi_1(x_j;\theta_1\theta_3,\zeta)-(\theta_2\theta_3)^{(x_j)^{-\zeta}}\Phi_1(x_j;\theta_1,\zeta)},$$

$$\frac{\partial L}{\partial \theta_2}=\sum_{j=1}^{n_1}\frac{\Phi_2(x_{2j}+1;\theta_2,\zeta)-\Phi_2(x_{2j};\theta_2,\zeta)}{\Phi_1(x_{2j};\theta_2,\zeta)}+\sum_{j=1}^{n_2}\frac{\theta_3\Phi_2(x_{2j}+1;\theta_2\theta_3,\zeta)-\theta_3\Phi_2(x_{2j};\theta_2\theta_3,\zeta)}{\Phi_1(x_{2j};\theta_1,\zeta)}$$
$$+\sum_{j=1}^{n_3}\frac{\Phi_2(x_j+1;\theta_2,\zeta)\Phi_1(x_j;\theta_1\theta_3,\zeta)-\theta_3\Phi_2(x_j;\theta_2\theta_3,\zeta)\Phi_2(x_j;\theta_1,\zeta)}{\theta_2^{(x_j+1)^{-\zeta}}\Phi_1(x_j;\theta_1\theta_3,\zeta)-(\theta_2\theta_3)^{(x_j)^{-\zeta}}\Phi_1(x_j;\theta_1,\zeta)},$$

$$\frac{\partial L}{\partial \theta_3}=\sum_{j=1}^{n_1}\frac{\theta_1\left[\Phi_2(x_{1j}+1;\theta_1\theta_3,\zeta)-\Phi_2(x_{1j};\theta_1\theta_3,\zeta)\right]}{\Phi_1(x_{1j};\theta_1\theta_3,\zeta)}+\sum_{j=1}^{n_2}\frac{\theta_2\left[\Phi_2(x_{2j}+1;\theta_2\theta_3,\zeta)-\Phi_2(x_{2j};\theta_2\theta_3,\zeta)\right]}{\Phi_1(x_{2j};\theta_2\theta_3,\zeta)}+$$
$$\sum_{j=1}^{n_3}\frac{\theta_1\theta_2^{(x_j+1)^{-\zeta}}\left[\Phi_2(x_j+1;\theta_1\theta_3,\zeta)-\Phi_2(x_j;\theta_1\theta_3,\zeta)\right]-\theta_2\Phi_2(x_j;\theta_2\theta_3,\zeta)\Phi_1(x_j;\theta_1,\zeta)}{\theta_2^{(x_j+1)^{-\zeta}}\Phi_1(x_j;\theta_1\theta_3,\zeta)-(\theta_2\theta_3)^{(x_j)^{-\zeta}}\Phi_1(x_j;\theta_1,\zeta)}$$

and

$$\frac{\partial L}{\partial \zeta}=\sum_{j=1}^{n_1}\frac{\Phi_3(x_{1j};\theta_1\theta_3,\zeta)-\Phi_3(x_{1j}+1;\theta_1\theta_3,\zeta)}{\Phi_1(x_{1j};\theta_1\theta_3,\zeta)}+\sum_{j=1}^{n_1}\frac{\Phi_3(x_{2j};\theta_2,\zeta)-\Phi_3(x_{2j}+1;\theta_2,\zeta)}{\Phi_1(x_{2j};\theta_2,\zeta)}$$
$$+\sum_{j=1}^{n_2}\frac{\Phi_3(x_{1j};\theta_1,\zeta)-\Phi_3(x_{1j}+1;\theta_1,\zeta)}{\Phi_1(x_{1j};\theta_1,\zeta)}+\sum_{j=1}^{n_2}\frac{\Phi_3(x_{2j};\theta_2\theta_3,\zeta)-\Phi_3(x_{2j}+1;\theta_2\theta_3,\zeta)}{\Phi_1(x_{2j};\theta_2\theta_3,\zeta)}$$
$$+\sum_{j=1}^{n_3}\frac{\theta_2^{(x_j+1)^{-\zeta}}\left[\Phi_3(x_j;\theta_1\theta_3,\zeta)-\Phi_3(x_j+1;\theta_1\theta_3,\zeta)\right]-\Phi_1(x_j;\theta_1\theta_3,\zeta)\Phi_3(x_j+1;\theta_2,\zeta)}{\theta_2^{(x_j+1)^{-\zeta}}\Phi_1(x_j;\theta_1\theta_3,\zeta)-(\theta_2\theta_3)^{(x_j)^{-\zeta}}\Phi_1(x_j;\theta_1,\zeta)}$$
$$+\sum_{j=1}^{n_3}\frac{(\theta_2\theta_3)^{(x_j)^{-\zeta}}\left[\Phi_3(x_j+1;\theta_1,\zeta)-\Phi_3(x_j;\theta_1,\zeta)\right]-\Phi_1(x_j;\theta_1,\zeta)\Phi_3(x_j;\theta_2\theta_3,\zeta)}{\theta_2^{(x_j+1)^{-\zeta}}\Phi_1(x_j;\theta_1\theta_3,\zeta)-(\theta_2\theta_3)^{(x_j)^{-\zeta}}\Phi_1(x_j;\theta_1,\zeta)}.$$