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, M. El-Morshedy

Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt.

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 methods 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,
2. The BDsIW Distribution

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

\[ \pi(x; \theta, \zeta) = \theta^{x_1+1} \zeta - \theta^{x_2+1} \zeta; \quad x \in \mathbb{N}_0, \]  

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 \(W_1 \sim \text{DsIW}(\theta_1, \zeta)\), \(W_2 \sim \text{DsIW}(\theta_2, \zeta)\) and \(W_3 \sim \text{DsIW}(\theta_3, \zeta)\) and they are independently distributed. If \(X_d = \max(W_d, W_3)\); \(d = 1, 2\), then we can say that the bivariate vector \(X = (X_1, X_2)\) 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 \(\text{BDsIW}(\theta_1, \theta_2, \theta_3, \zeta)\). If \(X \sim \text{BDsIW}(\theta_1, \theta_2, \theta_3, \zeta)\), then the joint CDF of \(X\) is given by

\[
F_{X_1,X_2}(x_1, x_2; \Psi) = \begin{cases} 
\theta_1^{(x_1+1)\zeta} \theta_2^{(x_2+1)\zeta} \theta_3^{(x_3+1)\zeta} & ; 0 < x_1 < x_2 < \infty \\
F_{\text{DsIW}}(x_1; \theta_1, \zeta) F_{\text{DsIW}}(x_2; \theta_2, \zeta) F_{\text{DsIW}}(x_3; \theta_3, \zeta) & ; 0 < x_1 = x_2 = x < \infty \\
F_{\text{DsIW}}(x_1; \theta_1, \zeta) F_{\text{DsIW}}(x_2; \theta_2, \zeta) & ; 0 < x_1 < x_2 < \infty \\
F_{\text{DsIW}}(x; \theta_1, \zeta, \theta_2, \zeta) & ; 0 < x_1 = x_2 = x < \infty \\
\end{cases}
\]  

(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

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

The marginal CDFs of the BDsIW distribution can be represented as

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

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

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

where

\[
f_1(x_1, x_2; \Psi) = f_{DsIW}(x_1; \theta_1, \theta_3, \zeta) f_{DsIW}(x_2; \theta_2, \zeta),
\]

\[
f_2(x_1, x_2; \Psi) = f_{DsIW}(x_1; \theta_1, \zeta) f_{DsIW}(x_2; \theta_2 \theta_3, \zeta)
\]

and

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

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

\[
f_{X_1,X_2}(x_1, x_2; \Psi) = F(x_1, x_2; \Psi) - F(x_1 - 1, x_2; \Psi) - F(x_1, x_2_1; \Psi) + F(x_1 - 1, x_2_1; \Psi).
\] (5)

The plots of the joint PMF are displayed in Figure 1.
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 $\mathbf{X} \sim \text{BDSiW}(\theta_1, \theta_2, \theta_3, \zeta)$, then the joint RF of $\mathbf{X}$ can be expressed as

$$R_{X_1, X_2}(x_1, x_2; \Psi) = 1 - F_{X_1}(x_1; \theta_1 \theta_3, \zeta) - F_{X_2}(x_2; \theta_2 \theta_3, \zeta) + F_{X_1,X_2}(x_1, x_2; \Psi)$$

$$= \begin{cases} 
R_1(x_1, x_2; \Psi) & ; \ 0 < x_1 < x_2 < \infty \\
R_2(x_1, x_2; \Psi) & ; \ 0 < x_1 < \infty \\
R_3(x; \Psi) & ; \ 0 < x_1 = x_2 = x < \infty, 
\end{cases}$$

where

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

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

and

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

The plots of the joint RF are displayed in Figure 2.
The plots indicate that the joint RF decreases in value when the values of $\theta_1, \theta_2, \theta_3$ and $\zeta$ increase. The joint HRF of $X$ is given by

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

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

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

$$r^*(X_1|X_2 > x_2) = \frac{\zeta (x_1 + 1)^{-\zeta-1}}{R_1(x_1,x_2;\Psi)} \{F_{\text{DaIW}}(x_2;\theta_2,\zeta) - 1\} F_{\text{DaIW}}(x_1;\theta_1,\theta_3,\zeta) \ln (\theta_1\theta_3), \quad (8)$$
while the HRF of the conditional distribution \( X_2 \) given \( X_1 > x_1 \) is given by

\[
r'(x_2 | X_1 > x_1) = \frac{\zeta (x_2 + 1)^{-\zeta - 1} F_{DsiW}(x_2; \theta_2, \zeta)}{R_1(x_1, x_2; \Psi)} [F_{DsiW}(x_1; \theta_1\theta_3, \zeta) \ln (\theta_2) - F_{DsiW}(x_2; \theta_3, \zeta) \ln (\theta_2\theta_3)]. \tag{9}
\]

Similarly, when \( X_2 < X_1 \), then

\[
r''(x_2 | X_1 > x_1) = \frac{\zeta (x_1 + 1)^{-\zeta - 1} F_{DsiW}(x_1; \theta_1, \zeta)}{R_2(x_1, x_2; \Psi)} [F_{DsiW}(x_1; \theta_1\theta_3, \zeta) \ln (\theta_1) - F_{DsiW}(x_2; \theta_3, \zeta) \ln (\theta_1\theta_3)]. \tag{10}
\]

and

\[
r''(x_2 | X_1 < x_1) = \frac{\zeta (x_2 + 1)^{-\zeta - 1}}{R_2(x_1, x_2; \Psi)} [F_{DsiW}(x_1; \theta_1, \zeta) - 1] F_{DsiW}(x_2; \theta_2\theta_3, \zeta) \ln (\theta_2\theta_3). \tag{11}
\]

### 3.2. Stress-strength probability

Assume \( X \sim BDsiW(\theta_1, \theta_2, \theta_3, \zeta) \). Then, the stress-strength probability can be expressed as

\[
P[X_1 < X_2] = \sum_{i=0}^{\infty} (\theta_1\theta_3)^{(i+1)^{-\zeta}} \left[ (\theta_2\theta_3)^{(i+1)^{-\zeta}} - (\theta_2\theta_3)^{-\zeta} \right]. \tag{13}
\]

and

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

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...
Regarding Table 1, it is clear that:

1. For fixed values of $\theta_1$, $\theta_3$ and $\zeta$ with $\theta_2 \rightarrow 1$, $P[X_1 < X_2]$ decreases.
2. For fixed values of $\theta_2$, $\theta_3$ and $\zeta$ with $\theta_1 \rightarrow 1$, $P[X_1 < X_2]$ increases.

### 3.3. The coefficient of median correlation

Since the median of $X_1$ and $X_2$ can be expressed as

$$M_{x_i} = \left\{ \log \frac{\theta_i \theta_3}{U} \right\}^\dagger - 1; \ d = 1, 2,$$

(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

$$M_{X_1,X_2} = \left\{ \frac{4F_{DSW}(M_{X_1}; \theta_1 \theta_3, \zeta) F_{DSW}(M_{X_2}; \theta_2, \zeta) - 1}{F_{DSW}(M_{X_1}; \theta_1, \zeta) F_{DSW}(M_{X_2}; \theta_2 \theta_3, \zeta) - 1} \right\}; \ x_1 < x_2$$

$$M_{X_1,X_2} = \left\{ \frac{4F_{DSW}(M_{X_1}; \theta_1 \theta_3, \zeta) F_{DSW}(M_{X_2}; \theta_2, \zeta) - 1}{F_{DSW}(M_{X_1}; \theta_1, \zeta) F_{DSW}(M_{X_2}; \theta_2 \theta_3, \zeta) - 1} \right\}; \ x_2 \leq x_1,$$

(16)

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

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

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

$$G_{X_1,X_2}(y_1, y_2) = \sum_{j=0}^{\infty} P[X_1 = i, X_2 = j] y_1^i y_2^j$$

$$= \sum_{j=0}^{\infty} \sum_{i=0}^{j-1} \left[ (\theta_1 \theta_3)^{(i+1)-} - (\theta_1 \theta_3)^{i-} \right] \left[ (\theta_2)^{(i+1)-} - (\theta_2)^{i-} \right] y_1^i y_2^j$$

$$+ \sum_{j=0}^{\infty} \sum_{i=j+1}^{\infty} \left[ \left( \theta_1 \theta_3 \right)^{(i+1)-} - (\theta_1 \theta_3)^{i-} \right] \left[ (\theta_2 \theta_3)^{(i+1)-} - (\theta_2 \theta_3)^{i-} \right] y_1^i y_2^j$$

$$+ \sum_{j=0}^{\infty} \theta_1^{(j+1)-} \left[ (\theta_1 \theta_3)^{(j+1)-} - (\theta_1 \theta_3)^{j-} \right] y_1^j y_2^i$$

$$- \sum_{j=0}^{\infty} (\theta_2 \theta_3)^{(j+1)-} \left[ \theta_1^{(j+1)-} - \theta_1^{j-} \right] y_1^i y_2^j; \ |y_1|, |y_2| < 1.$$

(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

\[
\rho_{X_1,X_2} = \frac{\mathbf{E}(X_1X_2) - \mathbf{E}(X_1)\mathbf{E}(X_2)}{\sqrt{\text{Var}(X_1)\text{Var}(X_2)}}, \quad 0 \leq \rho_{X_1,X_2} \leq 1,
\]

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

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

and

\[
\text{Kurtosis} = \frac{\gamma_{40} + 2 \gamma_{04} + 4 \rho \left( \rho \gamma_{22} - \gamma_{13} - \gamma_{31} \right)}{(1 - \rho^2)^2},
\]

where \( \gamma_{ij} = \mathbf{E}(X_i X_j) \). 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 \( \theta_1, \theta_2 \) and \( \theta_3 \).

| \( \theta_1 \) | \( \theta_2 \) | \( \theta_3 \) | 0.3 | 0.6 | 0.9 |
|---|---|---|---|---|---|
| 0.3 | 0.3 | | 0.2568(1.3680) | 0.2499(1.3999) | 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) \( \max\{X_1, X_2\} \sim \text{DslW}(\theta_1 \theta_2 \theta_3, \zeta) \).

b) The conditional CDF of \( X_1 \) given \( X_2 \leq x_2 \), is given by

\[
F_{X_1|X_2=x_2}(x_1 | x_2) = \begin{cases} 
F_{\text{DslW}}(x_1; \theta_1, \theta_3, \zeta) & \text{if } 0 < x_1 < x_2 < \infty \\
F_{\text{DslW}}(x_2; \theta_2, \theta_3, \zeta) & \text{if } 0 < x_2 < x_1 < \infty \\
F_{\text{DslW}}(x; \theta_1, \zeta) & \text{if } 0 < x_1 = x_2 = x < \infty.
\end{cases}
\]
The log-likelihood function becomes

\[
\ell(X_1, X_2 | \theta_1, \theta_2, \theta_3, \zeta) = \prod_{j=1}^{n_1} f_1(x_{1j}, x_{2j}) \prod_{j=1}^{n_2} f_2(x_{1j}, x_{2j}) \prod_{j=1}^{n_3} f_3(x_{1j}),
\]

where

\[
\begin{align*}
 f_{X_1|X_2=x_2}^{(1)}(x_1 | x_2) &= \frac{F_{\text{DW}}(x_1; \theta_1 \theta_3, \zeta) F_{\text{DW}}(x_2; \theta_2, \zeta)}{F_{\text{DW}}(x_2; \theta_2 \theta_3, \zeta)} , \\
 f_{X_1|X_2=x_2}^{(2)}(x_1 | x_2) &= F_{\text{DW}}(x_1; \theta_1, \zeta) \\
 f_{X_1|X_2=x_2}^{(3)}(x_1 | x) &= \frac{F_{\text{DW}}(x; \theta_2, \zeta) f_{\text{DW}}(x; \theta_1 \theta_3, \zeta) - F_{\text{DW}}(x-1; \theta_2 \theta_3, \zeta) f_{\text{DW}}(x; \theta_1, \zeta)}{f_{\text{DW}}(x; \theta_2 \theta_3, \zeta)}.
\end{align*}
\]

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 \(\zeta\) of the BDSiW distribution. Suppose that, we have a sample of size \(n\), of the form \((x_{11}, x_{21}), (x_{12}, x_{22}), \ldots, (x_{1n}, x_{2n})\) 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 \(n = \sum_{j=1}^{3} n_j\). Based on the observations, the likelihood function can be expressed as

\[
\ell(X_1, X_2 | \theta_1, \theta_2, \theta_3, \zeta) = \prod_{j=1}^{n_1} f_1(x_{1j}, x_{2j}) \prod_{j=1}^{n_2} f_2(x_{1j}, x_{2j}) \prod_{j=1}^{n_3} f_3(x_{1j}).
\]

The log-likelihood function becomes

\[
L(X_1, X_2 | \theta_1, \theta_2, \theta_3, \zeta) = \sum_{j=1}^{n_1} \ln \Phi_1(x_{1j}; \theta_1 \theta_3, \zeta) + \sum_{j=1}^{n_2} \ln \Phi_1(x_{2j}; \theta_2, \zeta) \\
+ \sum_{j=1}^{n_1} \ln \Phi_1(x_{1j}; \theta_1, \zeta) + \sum_{j=1}^{n_2} \ln \Phi_1(x_{2j}; \theta_2 \theta_3, \zeta) \\
+ \sum_{j=1}^{n_3} \ln \left( \left[ \theta_2 \right]^{3} \Phi_1(x_j; \theta_1, \theta_3, \zeta) - \left[ \theta_2 \theta_3 \right]^{3} \Phi_1(x_j; \theta_1, \zeta) \right),
\]

where \(\Phi_1(x; \theta, \zeta) = \theta^{x+1} - \theta^z\). The first partial derivatives of Equation (24) with respect to \(\theta_1, \theta_2, \theta_3\) and \(\zeta\) are called the normal equations (see, Appendix), where \(\Phi_2(x; \theta, \zeta) = x^{-2} \theta^{x-1} \ln(x) \ln(x)\), and \(\Phi_3(x; \theta, \zeta) = x^{-2} \theta^{x-1} \ln(x) \ln(x)\). The MLEs of the parameters \(\theta_1, \theta_2, \theta_3\) and \(\zeta\) can be obtained by solving the system of 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 \(\zeta\), when the sample size is large enough. A two sided \((1-\alpha)100\%\) CIs for \(\theta_1, \theta_2, \theta_3\) and \(\zeta\) are given by

\[
\hat{\theta}_1 \pm z_{\alpha/2} \sqrt{\text{Var}\left(\hat{\theta}_1\right)}, \hat{\theta}_2 \pm z_{\alpha/2} \sqrt{\text{Var}\left(\hat{\theta}_2\right)}, \hat{\theta}_3 \pm z_{\alpha/2} \sqrt{\text{Var}\left(\hat{\theta}_3\right)}\) and \(\hat{\zeta} \pm z_{\alpha/2} \sqrt{\text{Var}\left(\hat{\zeta}\right)}\) respectively, where \(\text{Var}\left(\hat{\theta}_1\right), \text{Var}\left(\hat{\theta}_2\right), \text{Var}\left(\hat{\theta}_3\right)\) and \(\text{Var}\left(\hat{\zeta}\right)\) are the asymptotic variances of \(\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3\) and \(\hat{\zeta}\) 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 \( \zeta \) 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 \( \pi(\zeta) \propto \zeta^{a_3 - 1} e^{b_3 \zeta} \). 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

\[
G(\delta_1, \delta_2, \delta_3, \zeta | X_1, X_2) = \frac{\pi(\delta_1)\pi(\delta_2)\pi(\delta_3)\pi(\zeta)L(X_1, X_2 | \delta_1, \delta_2, \delta_3, \zeta)}{\int_{\delta_1, \delta_2, \delta_3, \zeta} \pi(\delta_1)\pi(\delta_2)\pi(\delta_3)\pi(\zeta)L(X_1, X_2 | \delta_1, \delta_2, \delta_3, \zeta) \, d\delta_1 d\delta_2 d\delta_3 d\zeta}.
\]

(25)

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

\[
G(\delta_1, \delta_2, \delta_3, \zeta | X_1, X_2) \propto \pi(\delta_1)\pi(\delta_2)\pi(\delta_3)\pi(\zeta)L(X_1, X_2 | \delta_1, \delta_2, \delta_3, \zeta).
\]

(26)

Thus, the Bayesian estimators of the parameters \( \delta_1, \delta_2, \delta_3 \) and \( \zeta \) can be expressed as

\[
\begin{align*}
\hat{\delta}_1 & \propto \int_{\delta_1} G(\delta_1, \delta_2, \delta_3, \zeta | X_1, X_2) \, d\delta_1, \\
\hat{\delta}_2 & \propto \int_{\delta_2} G(\delta_1, \delta_2, \delta_3, \zeta | X_1, X_2) \, d\delta_2, \\
\hat{\delta}_3 & \propto \int_{\delta_3} G(\delta_1, \delta_2, \delta_3, \zeta | X_1, X_2) \, d\delta_3 \text{ and } \hat{\zeta} \propto \int_{\zeta} G(\delta_1, \delta_2, \delta_3, \zeta | X_1, X_2) \, d\zeta.
\end{align*}
\]

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 \( \zeta \) 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 \( \zeta \).
3. Repeat step 2 \( M \) times and obtain \( (\theta_{11}, \theta_{21}, \theta_{31}, \zeta_1), (\theta_{12}, \theta_{22}, \theta_{32}, \zeta_2), \ldots, (\theta_{1M}, \theta_{2M}, \theta_{3M}, \zeta_M) \).
4. Arrange \( \theta_{1i}, \theta_{2j}, \theta_{3k}, \zeta_l \), in ascending order as

\[
\begin{align*}
\theta_{1[1]}, & \theta_{1[2]}, \ldots, \theta_{1[M]}, \\
\theta_{2[1]}, & \theta_{2[2]}, \ldots, \theta_{2[M]}, \\
\theta_{3[1]}, & \theta_{3[2]}, \ldots, \theta_{3[M]}, \text{ and } \zeta_{[1]}, \zeta_{[2]}, \ldots, \zeta_{[M]}.
\end{align*}
\]

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

\[
(\theta_{1[M/2]}, \theta_{1[M/2]}, \theta_{1[M/2]}), (\theta_{2[M/2]}, \theta_{2[M/2]}, \theta_{2[M/2]}), (\theta_{3[M/2]}, \theta_{3[M/2]}, \theta_{3[M/2]}), \text{ and } (\zeta_{M/2}, \zeta_{M/2}, \zeta_{M/2}).
\]

5. Simulation Results

In this section, the MLE and BSE approaches are used to estimate the parameters \( \theta_1, \theta_2, \theta_3 \) and \( \zeta \) 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.
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 $\zeta$ 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 (BPhoH), independent bivariate Poisson (BPo), 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 ($\Lambda$).
6.1. The first data set (I): Motor data

This data is reported in Relia sta (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 \(i\)th system \(i = 1, \ldots, 18\).

![Figure 4. Failure times (in days) of motors A and B.](image)

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 \(\theta\) and \(\zeta\) 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.

| Marginal ↓ Measure → | Mean Statistics | Mean Std.err | Variance Statistics | Variance Std.err | Skewness Statistics | Skewness Std.err | Kurtosis Statistics | Kurtosis Std.err |
|----------------------|----------------|-------------|---------------------|------------------|--------------------|------------------|--------------------|------------------|
| \(X_1\)              | 199.000        | 15.116      | 4113.294            | 700.978          | -0.591             | 0.536            | -0.701             | 1.038            |
| \(X_2\)              | 207.944        | 18.036      | 5855.585            | 1030.99          | 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   | \( \lambda_1 = 189.1, \lambda_2 = 190.4, \tilde{\alpha}_1 = 0.165, \tilde{\alpha}_2 = 0.079 \) | 472.05   | 952.11  | 955.18  | 955.67  | 952.59  |
| BBi   | \( p_1 = 0.004, p_2 = 0.004, n_1 = 43652, n_2 = 43985, \tilde{\alpha}_1 = 0.185, \tilde{\alpha}_2 = 0.081 \) | 473.28   | 958.56  | 962.19  | 963.9   | 959.29  |
| BPoH  | \( \tilde{\alpha} = 199.01, b = 207.94, \tilde{d} = 81.10 \) | 513.96   | 1033.92 | 1035.63 | 1036.59 | 1034.29 |
| IBPo  | \( \lambda_1 = 1.876, \lambda_2 = 1.365 \) | 370.39   | 744.78  | 745.58  | 746.56  | 745.03  |
| BDSIE | \( \theta_1 = 0.06, \theta_2 = 0.052, \theta_3 = 0.184 \) | 378.47   | 762.94  | 764.65  | 765.61  | 763.31  |
| BDSIR | \( \theta_1 = 0.007, \theta_2 = 0.005, \theta_3 = 0.119 \) | 519.61   | 1045.22 | 1046.93 | 1047.89 | 1045.59 |
| BDSIW | \( \theta_1 = 0.279, \theta_2 = 0.247, \theta_3 = 0.212, \zeta = 0.388 \) | 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 \( \zeta \) 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 \( \Lambda \), 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 \( \Lambda \), d.f and p-values using data set I.

| Model | \( \Lambda \) | d.f | p-values |
|-------|--------------|-----|----------|
| BDSIE | 96.98        | 1   | < 0.01   |
| BDSIR | 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.

| Measure → | Correlation | Skewness | Kurtosis |
|-----------|-------------|----------|----------|
| BDSIW     | 0.498       | 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.](image)

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.](image)

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 $\theta$ and $\zeta$ 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.](image)

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.

| Marginal | Measure | Mean Statistics | Variance Statistics | Skewness Statistics | Kurtosis Statistics |
|----------|---------|-----------------|---------------------|---------------------|---------------------|
| $X_1$    |         | 1.500           | 0.149               | 0.67/2              | 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.

| Model | MLEs | $-L$ | AIC | CAIC | BIC | HQIC |
|-------|------|------|-----|------|-----|------|
| BPo   | $\lambda_1 = 1.08, \lambda_2 = 1.75, \hat{\alpha}_1 = 1.01, \hat{\alpha}_2 = 4.02$ | 62.52 | 133.05 | 134.94 | 138.07 | 134.49 |
| BBi   | $\rho_1 = 0.012, \rho_2 = 0.018, n_1 = 4325, n_2 = 4215, \hat{\alpha}_1 = 0.174, \hat{\alpha}_2 = 0.138$ | 63.72 | 139.44 | 143.86 | 146.99 | 141.61 |
| BPoH  | $\hat{\alpha} = 1.08, b = 1.38, \hat{\zeta} = 0.70$ | 64.92 | 135.83 | 136.93 | 139.61 | 136.93 |
| IBPo  | $\lambda_1 = 1.08, \lambda_2 = 1.38$ | 67.60 | 139.21 | 139.72 | 141.72 | 139.92 |
| BDsIE | $\hat{\theta}_1 = 0.669, \hat{\theta}_2 = 0.388, \hat{\theta}_3 = 0.514$ | 78.54 | 163.07 | 163.99 | 167.28 | 164.42 |
| BDsIR | $\hat{\theta}_1 = 0.493, \hat{\theta}_2 = 0.212, \hat{\theta}_3 = 0.561$ | 64.10 | 134.21 | 135.29 | 137.98 | 135.29 |
| BDsIW | $\hat{\theta}_1 = 0.420, \hat{\theta}_2 = 0.141, \hat{\theta}_3 = 0.587, \hat{\zeta} = 2.738$ | 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 $\zeta$ are given respectively as $[0.171, 0.521], [0.0, 0.274], [0.181, 0.641]$ 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.

| Model  | H_0 | Λ   | d.f | p-values |
|--------|-----|-----|-----|----------|
| BDsIE  | ζ = 1 | 33.152 | 1   | < 0.01   |
| BDsIR  | ζ = 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.

| Model | Measure → | 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.](image)

We fit at first the marginals $X_1$ and $X_2$ separately and $\min(X_1, X_2)$. The MLEs of the parameters $\theta$ and $\zeta$ 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.](image)
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.

| Marginal ↓ Measure → | Mean Statistics | Variance Statistics | Skewness Statistics | Kurtosis Statistics |
|-----------------------|-----------------|---------------------|---------------------|-------------------|
|                       | Mean | Std.err |              | Mean | Std.err |              | Mean | 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   | $\lambda_1 = 0.262$, $\alpha_1 = 0.165$, $\lambda_2 = 0.405$, $\alpha_2 = 2.97$ | 77.66 | 163.33 | 164.93 | 168.93 | 164.66 |
| BBi   | $p_1 = 0.145$, $p_2 = 0.122$, $n_1 = 452$, $n_2 = 468$, $\alpha_1 = 0.396$, $\alpha_2 = 0.748$ | 81.14 | 174.28 | 177.93 | 182.69 | 176.97 |
| BPoH  | $\alpha = 1.369$, $b = 2.314$, $\xi = 0.396$ | 82.36 | 170.72 | 171.64 | 174.92 | 172.06 |
| IBPo  | $\lambda_1 = 1.499$, $\lambda_2 = 1.367$ | 80.66 | 165.32 | 165.76 | 168.12 | 166.21 |
| BDsIE | $\theta_1 = 0.501$, $\theta_2 = 0.622$, $\theta_3 = 0.383$ | 92.48 | 190.96 | 191.88 | 195.16 | 192.30 |
| BDsIR | $\theta_1 = 0.262$, $\theta_2 = 0.405$, $\theta_3 = 0.363$ | 78.66 | 163.32 | 164.24 | 167.52 | 164.66 |
| BDsIW | $\theta_1 = 0.192$, $\theta_2 = 0.337$, $\theta_3 = 0.360$, $\xi = 2.453$ | 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 $\xi$ 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: $H_{01} : \xi = 1$ (BDsIE) against $H_{11} : \xi \neq 1$ (BDsIW).
Test 2: $H_{02} : \xi = 2$ (BDsIR) against $H_{12} : \xi \neq 2$ (BDsIW).

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

Table 15. The $\Lambda$, d.f and p-values for data set III.

| Model | $H$ | $\Lambda$ | d.f. | p-values |
|-------|-----|-----------|------|---------|
| BDIE  | $\xi = 1$ | 31.94 | 1 | < 0.01 |
| BDIR  | $\xi = 2$ | 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. 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    | $\theta_1$ | 0.272 | 0.198| 0.0011   | 0.275  | [0.195, 0.299]                  |
|      | $\theta_2$ | 0.248 | 0.201| 0.0009   | 0.246  | [0.183, 0.256]                  |
|      | $\theta_3$ | 0.210 | 0.199| 0.0013   | 0.212  | [0.129, 0.256]                  |
|      | $\zeta$    | 0.390 | 0.209| 0.0028   | 0.389  | [0.266, 0.453]                  |
| II   | $\theta_1$ | 0.410 | 0.214| 0.0031   | 0.417  | [0.301, 0.502]                  |
|      | $\theta_2$ | 0.140 | 0.247| 0.0011   | 0.101  | [0.056, 0.188]                  |
|      | $\theta_3$ | 0.577 | 0.176| 0.0029   | 0.566  | [0.430, 0.679]                  |
|      | $\zeta$    | 2.729 | 0.248| 0.0030   | 2.726  | [2.302, 3.168]                  |
| III  | $\theta_1$ | 0.192 | 0.064| 0.0018   | 0.041  | [0.111, 0.231]                  |
|      | $\theta_2$ | 0.339 | 0.107| 0.0024   | 0.335  | [0.301, 0.399]                  |
|      | $\theta_3$ | 0.358 | 0.105| 0.0016   | 0.355  | [0.301, 0.413]                  |
|      | $\zeta$    | 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., Salam, M. M., Ahmed, H. H., and El-Morshey, 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-Morshey, 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-Morshey, 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-Morshey, M., (2020). Bivariate odd Gumbel-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-Morshey, 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), 262.

[10] El-Morshey, 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), 2653-2652.

[11] El-Morshey, 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-Morshey, 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-Morshey, 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., Almualim, 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:10.1080/15598680.2018.1441082.

[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 oean 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] Nekoukhov, V., and Bidram, H. (2015). The exponentiated discrete Weibull distribution. Sort, 39, 127-146.

[26] Ortega, E. M., and Cordeiro, G. M. (2005). 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.reliaso.com/pubs/reliabilityedge_jq2002.pdf

[28] Salah, M. M., El-Morshey, 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-Morshey, 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:

\[
\frac{\partial L}{\partial \theta_1} = \sum_{j=1}^{m} \frac{\partial_2 \Phi_2(x_{ij} + 1; \theta_1, \theta_3, \zeta) - \partial_2 \Phi_2(x_{ij}; \theta_1, \theta_3, \zeta)}{\Phi_1(x_{ij}; \theta_1, \theta_3, \zeta)} + \sum_{j=1}^{m} \frac{\partial_3 \Phi_3(x_{ij} + 1; \theta_1, \zeta) - \partial_3 \Phi_3(x_{ij}; \theta_1, \zeta)}{\Phi_1(x_{ij}; \theta_1, \theta_3, \zeta)} + \sum_{j=1}^{m} \frac{\partial_3 \Phi_3(x_{ij} + 1; \theta_1, \zeta) - \partial_3 \Phi_3(x_{ij}; \theta_1, \zeta)}{\Phi_1(x_{ij}; \theta_1, \theta_3, \zeta)}
\]

\[
\frac{\partial L}{\partial \theta_2} = \sum_{j=1}^{m} \frac{\partial_2 \Phi_2(x_{ij} + 1; \theta_2, \zeta) - \partial_2 \Phi_2(x_{ij}; \theta_2, \zeta)}{\Phi_1(x_{ij}; \theta_1, \theta_3, \zeta)} + \sum_{j=1}^{m} \frac{\partial_3 \Phi_3(x_{ij} + 1; \theta_2, \zeta) - \partial_3 \Phi_3(x_{ij}; \theta_2, \zeta)}{\Phi_1(x_{ij}; \theta_1, \theta_3, \zeta)} + \sum_{j=1}^{m} \frac{\partial_3 \Phi_3(x_{ij} + 1; \theta_2, \zeta) - \partial_3 \Phi_3(x_{ij}; \theta_2, \zeta)}{\Phi_1(x_{ij}; \theta_1, \theta_3, \zeta)}
\]

\[
\frac{\partial L}{\partial \theta_3} = \sum_{j=1}^{m} \frac{\partial_2 \Phi_2(x_{ij} + 1; \theta_3, \zeta) - \partial_2 \Phi_2(x_{ij}; \theta_3, \zeta)}{\Phi_1(x_{ij}; \theta_1, \theta_3, \zeta)} + \sum_{j=1}^{m} \frac{\partial_3 \Phi_3(x_{ij} + 1; \theta_3, \zeta) - \partial_3 \Phi_3(x_{ij}; \theta_3, \zeta)}{\Phi_1(x_{ij}; \theta_1, \theta_3, \zeta)} + \sum_{j=1}^{m} \frac{\partial_3 \Phi_3(x_{ij} + 1; \theta_3, \zeta) - \partial_3 \Phi_3(x_{ij}; \theta_3, \zeta)}{\Phi_1(x_{ij}; \theta_1, \theta_3, \zeta)}
\]

and

\[
\frac{\partial L}{\partial \zeta} = \sum_{j=1}^{m} \frac{\partial_3 \Phi_3(x_{ij}; \theta_1, \theta_3, \zeta) - \partial_3 \Phi_3(x_{ij} + 1; \theta_1, \theta_3, \zeta)}{\Phi_1(x_{ij}; \theta_1, \theta_3, \zeta)} + \sum_{j=1}^{m} \frac{\partial_3 \Phi_3(x_{ij}; \theta_2, \zeta) - \partial_3 \Phi_3(x_{ij} + 1; \theta_2, \zeta)}{\Phi_1(x_{ij}; \theta_1, \theta_3, \zeta)} + \sum_{j=1}^{m} \frac{\partial_3 \Phi_3(x_{ij}; \theta_3, \zeta) - \partial_3 \Phi_3(x_{ij} + 1; \theta_3, \zeta)}{\Phi_1(x_{ij}; \theta_1, \theta_3, \zeta)} + \sum_{j=1}^{m} \frac{\partial_3 \Phi_3(x_{ij}; \theta_3, \zeta) - \partial_3 \Phi_3(x_{ij} + 1; \theta_3, \zeta)}{\Phi_1(x_{ij}; \theta_1, \theta_3, \zeta)} + \sum_{j=1}^{m} \frac{\partial_3 \Phi_3(x_{ij}; \theta_3, \zeta) - \partial_3 \Phi_3(x_{ij} + 1; \theta_3, \zeta)}{\Phi_1(x_{ij}; \theta_1, \theta_3, \zeta)} + \sum_{j=1}^{m} \frac{\partial_3 \Phi_3(x_{ij}; \theta_3, \zeta) - \partial_3 \Phi_3(x_{ij} + 1; \theta_3, \zeta)}{\Phi_1(x_{ij}; \theta_1, \theta_3, \zeta)}\]