Bivariate Inverse Weibull Distribution

Hiba Z. Muhammed

Department of Mathematical Statistics, Institute of Statistical Studies and Research,

Cairo University, Egypt.

Email: hiba_stat@yahoo.com, hiba_stat@cu.edu.eg.

Abstract

Recently it is observed that the inverse Weibull distribution can be used quite effectively to analyze lifetime data in one dimension. The main aim of this paper is to define a bivariate inverse Weibull distribution so that the marginals have inverse Weibull distributions. It is observed that the joint probability density function and the joint cumulative distribution function can be expressed in compact forms. Several properties of this distribution such as marginals, conditional distributions and product moments have been discussed. We obtained the maximum likelihood estimates for the unknown parameters of this distribution and their approximate variance- covariance matrix. We perform some simulations to see the performances of the MLEs. One data set has been re-analyzed and it is observed that the bivariate inverse Weibull distribution provides a better fit than the bivariate exponential distribution.

Keywords: *Bivariate Inverse Weibull distribution, product moments, Maximum Likelihood estimators.*

1 Introduction

It is well known that the Weibull probability density function (PDF) can be decreasing or unimodal, and the hazard function (HF) can be either decreasing or increasing depending on the shape parameter. Because of the flexibility of the PDF and HF, the Weibull distribution has been used quite extensively when the data indicate a monotone HF. But it cannot be used at all if the data indicate a non-monotone and unimodal HF. In many practical situations, it is often known a priori that the hazard rate cannot be monotone. It may happen that the course of a disease is such that the mortality reaches a peak after some finite period, and then declines slowly. For example, in a study of curability of breast cancer, Langlands et al. [1] found that the peak mortality occurred after about three years. Bennette [2] analyzed the data from the Veterans Administration lung cancer trial presented by Prentice [3] and showed that the empirical failure rates for both low and high performance status groups were unimodal in nature. It is important to analyze such data sets with the appropriate models. If the empirical studies indicate that the hazard function might be unimodal, then the inverse Weibull (IW) distribution may be an appropriate model.

The inverse Weibull distribution can be readily applied to a wide range of situations including applications in medicine, reliability and ecology. Keller et al. [4] obtained the inverse Weibull model by investigating failures of mechanical components subject to degradation phenomena such as the dynamic components of diesel engines; see for example Murthy et al. [5]. The physical failure process given by Erto and Rapone [6] also leads to the IW model. Erto and Rapone [6] showed that the IW model provides a good fit to survival data such as the times to breakdown of an insulating fluid subject to the action of constant tension; see also Nelson [7]. Calabria and Pulcini [8] provided an interpretation of the IW distribution in the context of a load-strength relationship for a component.

Recently, Aleem [9] introduced a bivariate inverse Weibull distribution using the Farlie- Gumbel- Morgenstern idea that proposed by Gumbel [10]. He obtained a bivariate inverse Weibull distribution whose marginals haven't inverse Weibull distributions. The main aim of this paper is to provide a bivariate inverse Weibull distribution whose marginals have inverse Weibull distributions using an idea similar to that of Theorem 3.2 proposed by Marshall and Olkin [11]. These authors introduced a multivariate exponential distribution whose marginals have exponential distributions and proposed a bivariate Weibull distribution.

On the basis of this idea, Sarhan and Balakrishnan [12] proposed a bivariate distribution that is more flexible than the bivariate exponential distribution. This distribution was generalized by Kundu and Gupta [13]. Recently, Kundu and Dey [14] have considered the maximum-likelihood estimation of the model parameters of the Marshall–Olkin bivariate Weibull distribution via the EM algorithm. Using the maximum instead of the minimum in the Marshall and Olkin scheme, Kundu and Gupta [13, 15] and Sarhan et al. [16] introduced the bivariate generalized exponential, bivariate proportional reversed hazard and bivariate generalized linear failure rate distributions, respectively.

The proposed Bivariate Inverse Weibull distribution (BIW) is constructed from three independent IW distributions using a maximization process. This new distribution is a singular distribution, and it can be used quit conveniently if there are ties in the data. The following interpretations are provided for the BIW model

Competing risks model: Assume a system has two components, labeled 1 and 2, and the survival time of component i is denoted by X_i, $i = 1, 2.$ It is considered that there are three independent causes of failures, which may affect the system. Only component 1 can fail due to cause 1, and similarly only component 2 can fail due to cause 2, while both the components can fail at the same time due to cause 3. Let U_i be the lifetime of cause i, $i = 1, 2, 3$. If U_1, U_2 and U_3 follow a IW distribution, then (X_1, X_2) follow the BIW model.

Shock model: Suppose there are three independent sources of shocks; say 1, 2 and 3. Suppose these shocks are affecting a system with two components, say 1 and 2. It is assumed that the shock from source 1 reaches the system and destroys component 1 immediately, the shock from source 2 reaches the system and destroys component 2 immediately, while if the shock from source 3 hits the system it destroys both components immediately. Let U_i denote the inter-interval times between the shocks in source i, $i = 1, 2, 3$, which follow the distribution IW. If X_1, X_2 denote the survival times of the components, then (X_1, X_2) follows the BIW model.

Stress Model: Suppose a system has two components. Each component is subject to individual independent stress say U_1 and U_2 respectively. The system has an overall stress U_3 which has been transmitted to both the components equally, independent of their individual stresses. Therefore, the observed stress at the two components are $X_1 = \max(U_1, U_3)$ and $X_2 = \max(U_2, U_3)$, respectively.

Maintenance Model: Suppose a system has two components and it is assumed that each component has been maintained independently and also there is an overall maintenance. Due to component maintenance, suppose the lifetime of the individual components is increased by U_i amount and because of the overall maintenance, the lifetime of each component is increased by U_3 amount. Therefore, the increased lifetimes of the two components are $X_1 = \max(U_1, U_3)$ and $X_2 = \max(U_2, U_3)$ respectively.

The paper is organized as follows: In Section 2, we introduce the BIW distribution and obtain representations for the cumulative distribution function (cdf) and probability density function (pdf). The conditional and marginal distributions of our bivariate model are presented in Section3. The maximum likelihood estimation, estimated variance-covariance matrix and asymptotic confidence intervals for BIW distribution are provided in Section 4. Simulation results and data analysis are presented in Section 5 and Section 6 respectively. Finally conclude the paper in Section 7.

2 Bivariate Inverse Weibull Distribution

A random variable with an Inverse Weibull distribution has a cdf and a pdf, for $x > 0$, in the following form

$$f_{IW}(x;\lambda,\alpha) = \alpha\,\lambda\,x^{-\alpha-1}\,e^{-\lambda x^{-\alpha}}\ ,\ F_{IW}(x;\lambda,\alpha) = e^{-\lambda x^{-\alpha}}$$

Respectively, where the quantities $\alpha > 0$ and $\lambda > 0$ are the shape and scale parameters respectively. From now on it will be denoted by $IW(\alpha,\lambda)$.

As in the Weibull distribution, the shape parameter α governs the shape of the PDF, the hazard function and the general properties of the IW distribution

The k^{th} $(k < \alpha)$ moment of X is

$$E(X^k) = \lambda^{\frac{k}{\alpha}}\ \Gamma\left(1 - \frac{k}{\alpha}\right).$$

Now, Suppose U_1, U_2 and U_3 are three independent random variables such that U_i ~ IW(λ_i, α) for $i = 1, 2, 3$. Define $X_1 = \max(U_1, U_3)$ and $X_2 = \max(U_2, U_3)$, then it is said that the bivariate vector (X_1, X_2) has bivariate inverse Weibull distribution with parameters $(\lambda_1, \lambda_2, \lambda_3, \alpha)$, denoted by $BIW(\lambda_1, \lambda_2, \lambda_3, \alpha)$. Then, the joint cdf of (X_1, X_2) is given as follows

$$F_{BIW}(x_1, x_2) = \prod_{i=1}^{3} F_{IW}(x_i; \lambda_i, \alpha)$$

where $x_3 = \min(x_1, x_2)$.

The following Proposition will provide the joint cumulative distribution function (cdf), joint probability density function (pdf), the marginal distributions and conditional probability density function

Proposition 2.1. If $(X_1, X_2) \sim BIW(\lambda_1, \lambda_2, \lambda_3, \alpha)$. Then, the joint cdf of (X_1, X_2) can be written as

$$F_{BIW}(x_1, x_2) = \begin{cases} F_1(x_1, x_2) & if \quad 0 < x_1 < x_2 < \infty \\ F_2(x_1, x_2) & if \quad 0 < x_2 < x_1 < \infty \\ F_3(x) & if \quad 0 < x_1 = x_2 = x < \infty, \end{cases} \tag{2.1}$$

Where

$F_1(x_1, x_2) = F_{IW}(x_1; \lambda_1 + \lambda_3, \alpha) F_{IW}(x_2; \lambda_2, \alpha)$

$F_2(x_1, x_2) = F_{IW}(x_1; \lambda_1, \alpha) F_{IW}(x_2; \lambda_2 + \lambda_3, \alpha)$

and $F_3(x) = F_{IW}(x\,; \lambda_1 + \lambda_2 + \lambda_3, \alpha)$

Proposition 2.2. If $(X_1, X_2) \sim BIW(\lambda_1, \lambda_2, \lambda_3, \alpha)$. Then, the joint pdf of (X_1, X_2) is given as

$$f_{BIW}(x_1, x_2) = \begin{cases} f_1(x_1, x_2) & if \quad 0 < x_1 < x_2 < \infty \\ f_2(x_1, x_2) & if \quad 0 < x_2 < x_1 < \infty \\ f_3(x) & if \quad 0 < x_1 = x_2 = x < \infty, \end{cases} \tag{2.2}$$

Where

$f_1(x_1, x_2) = f_{IW}(x_1; \lambda_1 + \lambda_3, \alpha) f_{IW}(x_2; \lambda_2, \alpha)$

$f_2(x_1, x_2) = f_{IW}(x_1; \lambda_1, \alpha) f_{IW}(x_2; \lambda_2 + \lambda_3, \alpha)$

and $f_3(x) = \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3} f_{IW}(x\,; \lambda_1 + \lambda_2 + \lambda_3, \alpha)$

Proof. The expressions for $f_1(.,.)$and $f_2(.,.)$ can be obtained simply by taking $\frac{\partial^2}{\partial x_1 \partial x_2} F_{X_1, X_2}(x_1, x_2)$ for $x_1 < x_2$ and $x_2 < x_1$ respectively. But $f_3(.)$cannot be obtained in the same way. Using the fact that

$$\int_0^{\infty}\int_0^{x_2} f_1(x_1, x_2)dx_1 dx_2 + \int_0^{\infty}\int_0^{x_1} f_2(x_1, x_2)dx_2 dx_1 + \int_0^{\infty} f_3(x)dx = 1,$$

$$\int_0^\infty \int_0^{x_2} f_1(x_1, x_2) dx_1 dx_2 = \frac{\lambda_2}{\lambda_1 + \lambda_2 + \lambda_3} \quad \text{and} \quad \int_0^\infty \int_0^{x_1} f_2(x_1, x_2) dx_2 dx_1 = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3}$$

Hence, we obtain

$$\int_0^\infty f_3(x) dx = \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}$$

Note that

$$\int_0^\infty f_3(x) dx = \alpha \lambda_3 \int_0^\infty e^{-(\lambda_1+\lambda_2+\lambda_3)x^{-\alpha}} \ x^{-\alpha-1} dx = \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}.$$

Therefore, the results follow.

It should be mentioned that the BIW distribution has both an absolute continuous part and a singular part, similar to Marshall and Olkin's bivariate exponential model. The function $f_{X_1,X_2}(.,.)$ may be considered to be a density function for the BIW distribution if it is understood that the first two terms are densities with respect to two-dimensional Lebesgue measure and the third term is a density function with respect to one dimensional Lebesgue measure, see for example Bemis et al. [17]. It is well known that although in one dimension the practical use of a distribution with this property is usually pathological, but they do arise quite naturally in higher dimension. In case of BIW distribution, the presence of a singular part means that if X_1 and X_2 are BIW distribution, then $X_1 = X_2$ has a positive probability. In many practical situations it may happen that X_1 and X_2 both are continuous random variables, but $X_1 = X_2$ has a positive probability, see Marshall and Olkin [11] in this connection. The following Proposition will provide explicitly the absolute continuous part and the singular part of the BIW distribution function.

Proposition 2.3. If $(X_1, X_2) \sim BIW(\lambda_1, \lambda_2, \lambda_3, \alpha)$. Then,

$$F_{X_1,X_2}(x_1, x_2) = \frac{\lambda_1 + \lambda_2}{\lambda_1 + \lambda_2 + \lambda_3} F_a(x_1, x_2) + \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3} F_s(x_3), \tag{2.3}$$

where $x_3 = \min\{x_1, x_2\}$, $F_s(x_3) = e^{-(\lambda_1+\lambda_2+\lambda_3)x_3^{-\alpha}}$, and $F_a(x_1,x_2) = \frac{\lambda_1+\lambda_2+\lambda_3}{\lambda_2+\lambda_3} e^{-\lambda_1 x_1^{-\alpha}} \quad e^{-\lambda_2 x_2^{-\alpha}} \quad e^{-\lambda_3 x_3^{-\alpha}} - \frac{\lambda_3}{\lambda_1+\lambda_2+\lambda_3} e^{-(\lambda_1+\lambda_2+\lambda_3)x_3^{-\alpha}}$.

Here $F_s(.,.)$ and $F_a(.,.)$ are the singular and the absolutely continuous part respectively.

Proof.

Suppose A is the following event

$A = \{U_1 < U_3\} \cap \{U_2 < U_3\}$,

Then $P(A) = \frac{\lambda_3}{\lambda_1+\lambda_2+\lambda_3}$ and $P(A') = \frac{\lambda_1+\lambda_2}{\lambda_1+\lambda_2+\lambda_3}$.

Therefore,

$F_{X_1,X_2}(x_1,x_2) = P(X_1 \le x_1, X_2 \le x_2 \mid A)P(A) + P(X_1 \le x_1, X_2 \le x_2 \mid A')P(A')$.

Moreover for x_3 as defined before,

$$P(X_1 \le x_1, X_2 \le x_2 \mid A) = [P(A)]^{-1} P(U_1 \le U_3, U_2 \le U_3, U_3 \le x_3)$$
$$= e^{-(\lambda_1+\lambda_2+\lambda_3)x_3^{-\alpha}}$$

and $P(X_1 \le x_1, X_2 \le x_2 \mid A')$ can be obtained by subtraction as

$$P(X_1 \le x_1, X_2 \le x_2 \mid A') = \frac{\lambda_1+\lambda_2+\lambda_3}{\lambda_2+\lambda_3} e^{-\lambda_1 x_1^{-\alpha}} \quad e^{-\lambda_2 x_2^{-\alpha}} \quad e^{-\lambda_3 x_3^{-\alpha}} - \frac{\lambda_3}{\lambda_1+\lambda_2+\lambda_3} e^{-(\lambda_1+\lambda_2+\lambda_3)x_3^{-\alpha}}.$$

Clearly, $e^{-(\lambda_1+\lambda_2+\lambda_3)x_3^{-\alpha}}$ is the singular part as its mixed second partial derivative is zero when $x_1 \ne x_2$, and $P(X_1 \le x_1, X_2 \le x_2 \mid A')$ is the absolute continuous part as its mixed partial derivative is a density function.

Using Proposition 2.3, we immediately obtain the joint pdf of X_1 and X_2 also in the following form for $x_3 = \min\{x_1, x_2\}$;

$$f_{X_1,X_2}(x_1,x_2) = \frac{\lambda_2+\lambda_3}{\lambda_1+\lambda_2+\lambda_3} f_a(x_1,x_2) + \frac{\lambda_3}{\lambda_1+\lambda_2+\lambda_3} f_s(x_3) \qquad (2.4)$$

where $$f_a(x_1,x_2) = \frac{\lambda_1+\lambda_2+\lambda_3}{\lambda_2+\lambda_3} \times \begin{cases} f_{IW}(x_1;\lambda_1+\lambda_3)\cdot f_{IW}(x_2;\lambda_2) & if \quad x_1 < x_2 \\ f_{IW}(x_1;\lambda_1)\cdot f_{IW}(x_2;\lambda_2+\lambda_3) & if \quad x_1 > x_2 \end{cases}$$

and

$$f_s(x_3) = f_{IW}(x_3; \lambda_1 + \lambda_2 + \lambda_3).$$

Clearly, here $f_a(x_1, x_2)$ and $f_s(x_3)$ are the absolutely continuous and singular part respectively.

Note that the absolutely continuous bivariate Inverse Weibull (ACBIW) distribution can be obtained from Marshal and Olkin BIW distribution by removing the singular part and keeping only the continuous part. The joint pdf of ACBIW can be written as

$$f_{ACBIW}(y_1, y_2) = \begin{cases} c\ f_1(y_1, y_2) & if \quad y_1 < y_2 \\ c\ f_2(y_1, y_2) & if \quad y_1 > y_2 \end{cases}$$

$$= c. \begin{cases} f_{IW}(y_1; \lambda_1 + \lambda_3) \cdot f_{IW}(y_2; \lambda_2) & if \quad y_1 < y_2 \\ f_{IW}(y_1; \lambda_1) \cdot f_{IW}(y_2; \lambda_2 + \lambda_3) & if \quad y_1 > y_2 \end{cases},$$

Here c is the normalizing constant and $c = \frac{\lambda_1 + \lambda_2 + \lambda_3}{\lambda_1 + \lambda_2}$.

The absolutely continuous part of the BIW density may be unimodal depending on the values of $\lambda_1, \lambda_2, \lambda_3$ and α, that is $f_a(x_1, x_2)$ is unimodal and the respective modes are

$$\left\{[(\lambda_1 + \lambda_3)\frac{\alpha}{\alpha+1}]^{\frac{1}{\alpha}}, [\lambda_2 \frac{\alpha}{\alpha+1}]^{\frac{1}{\alpha}}\right\} \text{ and } \left\{[(\lambda_2 + \lambda_3)\frac{\alpha}{\alpha+1}]^{\frac{1}{\alpha}}, [\lambda_1 \frac{\alpha}{\alpha+1}]^{\frac{1}{\alpha}}\right\}.$$

The median for the BIW distribution is obtained as

$$\left[\frac{\lambda_1 + \lambda_2 + \lambda_3}{\ln(2)}\right]^{\frac{1}{\alpha}}.$$

3 Different Properties

In this Section we provide different basic properties of the BIW model. First we provide the marginal and conditional distributions of BIW model.

Proposition 3.1. If $(X_1, X_2) \sim BIW(\lambda_1, \lambda_2, \lambda_3, \alpha)$. Then,

1. $X_1 \sim IW(x_1; \lambda_1 + \lambda_3, \alpha)$ and $X_2 \sim IW(x_2; \lambda_2 + \lambda_3, \alpha)$
2. The conditional pdf of X_i, given $X_j = x_j$, denoted by $f_{i/j}(x_i/ x_j)$ $(i \neq j = 1,2)$, is given by

$$f_{i/j}(x_1,x_2)=\begin{cases} f_{i/j}^{(1)}(x_i/x_j) & if \quad x_i<x_j \\ f_{i/j}^{(2)}(x_i/x_j) & if \quad x_j<x_i \\ f_{i/j}^{(3)}(x_i/x_j) & if \quad x_i=x_j, \end{cases} \tag{3.1}$$

where

$$f_{i/j}^{(1)}(x_i/x_j)=(\lambda_1+\lambda_3)x_i^{-\alpha-1}\ \mathrm{e}^{-(\lambda_1+\lambda_3)x_i^{-\alpha}}\mathrm{e}^{-\lambda_3 x_j^{-\alpha}},$$

$$f_{i/j}^{(2)}(x_i/x_j)=\lambda_1\alpha\ x_i^{-\alpha-1}\ \mathrm{e}^{-\lambda_1 x_i^{-\alpha}},$$

$$f_{i/j}^{(3)}(x_i/x_j)=\frac{\lambda_3\ x_i^{-\alpha-1}}{(\lambda_2+\lambda_3)x_j^{-\alpha-1}}\ \mathrm{e}^{-(\lambda_1+\lambda_2+\lambda_3)x_i^{-\alpha}}\mathrm{e}^{(\lambda_2+\lambda_3)x_j^{-\alpha}}.$$

Proof: They can be obtained by routine calculation.

Algorithm to generate from BIW

Step 1. Generate U_1, U_2 and U_3 from $U(0,1)$,

Step 2. Compute

$$Z_1=\left[\frac{-\lambda_1}{\ln(U_2)}\right]^{\frac{1}{\alpha}},\ Z_2=\left[\frac{-\lambda_2}{\ln(U_2)}\right]^{\frac{1}{\alpha}}\ \text{and}\ \ Z_3=\left[\frac{-\lambda_3}{\ln(U_3)}\right]^{\frac{1}{\alpha}},$$

Step3. Obtain

$$X_1=\max(Z_1,Z_3)\ \text{ and }\ X_2=\max(Z_2,Z_3)\,.$$

Since the marginal distributions of the bivariate vector (X_1,X_2) are Inverse Weibull distributions, the moments of X_1 and X_2 can be obtained directly from their marginals

$$E({X_1}^r)=(\lambda_1+\lambda_3)^{\frac{r}{\alpha}}\ \Gamma(1-\frac{r}{\alpha})\ \text{and}\ E({X_2}^r)=(\lambda_2+\lambda_3)^{\frac{r}{\alpha}}\ \Gamma(1-\frac{r}{\alpha}),$$

where $\Gamma(.)$ is the Gamma function. Now, we present the product moments

Proposition 3.2. If $(X_1,X_2)\sim\ BIW(\lambda_1,\lambda_2,\lambda_3,\alpha)$. Then, The r^{th} and s^{th} joint moments of the X_1 and X_2, denoted by $\mu'_{r,s}$ is given by

$$\mu'_{r,s} = E(X_1^r, X_2^s) = \lambda_2\ (\lambda_1+\lambda_3)^{\alpha r} \quad \frac{\Gamma[2-\alpha(s+r)]\,F(1,\ [2-\alpha(s+r)];\ 2-\alpha\ r;\ \frac{1}{1+\lambda_2})}{(1-\alpha\ r)\ (1+\lambda_2)^{[2-\alpha(s+r)]}}$$

$$+\lambda_1\ (\lambda_1+\lambda_3)^{\alpha s} \quad \frac{\Gamma[2-\alpha(s+r)]\,F(1,\ [2-\alpha(s+r)];\ 2-\alpha\ s;\ \frac{1}{1+\lambda_1})}{(1-\alpha\ s)\ (1+\lambda_1)^{[2-\alpha(s+r)]}}$$

$$-\lambda_3\ (\lambda_1+\lambda_2+\lambda_3)^{\alpha(r+s)-1}\ \ \Gamma[2-\alpha(s+r)]. \tag{3.2}$$

where

$\Gamma(a) = \int_0^{\infty} t^{a-1}\ e^{-t} dt,$ is the gamma function,

$\mathrm{F}(\mathrm{a},\mathrm{b};\mathrm{c};\mathrm{z}) = \sum_{s=0}^{\infty} \frac{(a)_s (b)_s}{\Gamma(c+s) s!}\ z^s$ is a hypergeometric function,

and $(b)_i = b(b+1)\ldots(b+i-1) = \frac{\Gamma(b+i)}{\Gamma(b)}(b \neq 0, i = 1,2,\ldots).$

Proof

Starting with $E(X_1^r, X_2^s) = \int_0^{\infty}\int_0^{\infty} x_1^r x_2^s f(x_1, x_2) dx_1 dx_2 \qquad r, s = 1,2,3,\ldots$

and substituting for $f(x_1, x_2)$ .Then, using the fact that

$$\int_0^{\infty} x^{a-1} e^{-sx}\ \Gamma^*(b,x) dx = \frac{\Gamma(a+b)}{b(1+s)^{a+b}}\ \mathrm{F}\left(1, \mathrm{a}+\mathrm{b}; 1+\mathrm{b}; \frac{1}{1+s}\right).$$

We can derive the expression for $E(X_1^r, X_2^s)$.

4 Maximum likelihood Estimation

In this Section, we address the problem of computing the maximum likelihood estimators (MLEs) of the unknown parameters of the BIW distribution. Suppose $\{(x_{11}, x_{21}), \ldots, (x_{1n}, x_{2n})\}$ is a random sample from $BIW(\lambda_1, \lambda_2, \lambda_3, \alpha)$ distribution. Consider the following notation

$$I_1 = \{i; x_{1i} < x_{2i}\}, \quad I_2 = \{i; x_{1i} > x_{2i}\}, \quad I_3 = \{x_{1i} = x_{2i} = x_i\}, \quad I = I_1 \cup I_2 \cup I_3,$$

$$|I_1| = n_1, \quad |I_2| = n_2, \quad |I_3| = n_3, \quad \text{and} \quad n_1 + n_2 + n_3 = n.$$

The log-likelihood function of the sample of size n is given by

$$\ln L(\underline{\theta}) = \sum_{i \in I_1} \ln f_1(x_{1i}, x_{2i}) + \sum_{i \in I_2} \ln f_2(x_{1i}, x_{2i}) + \sum_{i \in I_3} \ln f_3(x_i) \tag{4.1}$$

$$\begin{aligned}\ln L(\underline{\theta}) = {} & n_1 \ln(\lambda_1 + \lambda_3) + n_1 \ln \lambda_2 + n_2 \ln(\lambda_2 + \lambda_3) + n_2 \ln \lambda_2 + n_3 \ln \lambda_3 + (2n_1 + 2n_2 + n_3) \ln(\alpha) \\ & - (\alpha + 1)[\sum_{i=1}^{n_1} \ln x_{1i} + \sum_{i=1}^{n_1} \ln x_{2i} + \sum_{i=1}^{n_2} (\ln x_{1i} + \ln x_{2i}) + \sum_{i=1}^{n_3} \ln x_i] \\ & - (\lambda_1 + \lambda_3) \sum_{i=1}^{n_1} x_{1i}^{-\alpha} - \lambda_2 \sum_{i=1}^{n_1} x_{2i}^{-\alpha} - (\lambda_2 + \lambda_3) \sum_{i=1}^{n_2} x_{2i}^{-\alpha} - \lambda_1 \sum_{i=1}^{n_2} x_{1i}^{-\alpha} - (\lambda_1 + \lambda_2 + \lambda_3) \sum_{i=1}^{n_3} x_i^{-\alpha}\end{aligned} \tag{4.2}$$

where $\underline{\theta} = (\lambda_1, \lambda_2, \lambda_3, \alpha)$.

On differentiating (4.2) with respect to $\lambda_1, \lambda_2, \lambda_3$ and α in turn and equating to zero, we obtain the following likelihood equations

$$\frac{n_1}{\hat{\lambda}_1 + \hat{\lambda}_3} + \frac{n_2}{\hat{\lambda}_1} - \sum_{i=1}^{n_1} x_{1i}^{-\hat{\alpha}} - \sum_{i=1}^{n_2} x_{1i}^{-\hat{\alpha}} - \sum_{i=1}^{n_3} x_i^{-\hat{\alpha}} = 0,$$

$$\frac{n_2}{\hat{\lambda}_2 + \hat{\lambda}_3} + \frac{n_1}{\hat{\lambda}_2} - \sum_{i=1}^{n_1} x_{2i}^{-\hat{\alpha}} - \sum_{i=1}^{n_2} x_{2i}^{-\hat{\alpha}} - \sum_{i=1}^{n_3} x_i^{-\hat{\alpha}} = 0,$$

$$\frac{n_1}{\hat{\lambda}_1 + \hat{\lambda}_3} + \frac{n_2}{\hat{\lambda}_2 + \hat{\lambda}_3} + \frac{n_3}{\hat{\lambda}_3} - \sum_{i=1}^{n_1} x_{1i}^{-\hat{\alpha}} - \sum_{i=1}^{n_2} x_{2i}^{-\hat{\alpha}} - \sum_{i=1}^{n_3} x_i^{-\hat{\alpha}} = 0,$$

and

$$\begin{aligned}& \frac{2n_1 + 2n_2 + n_3}{\hat{\alpha}} - [\sum_{i=1}^{n_1} \ln x_{1i} + \sum_{i=1}^{n_1} \ln x_{2i} + \sum_{i=1}^{n_2} (\ln x_{1i} + \ln x_{2i}) + \sum_{i=1}^{n_3} \ln x_i] + (\hat{\lambda}_1 + \hat{\lambda}_3) \sum_{i=1}^{n_1} x_{1i}^{-\hat{\alpha}} \ln x_{1i} \\ & + \hat{\lambda}_2 \sum_{i=1}^{n_1} x_{2i}^{-\hat{\alpha}} \ln x_{2i} + (\hat{\lambda}_2 + \hat{\lambda}_3) \sum_{i=1}^{n_2} x_{2i}^{-\hat{\alpha}} \ln x_{2i} + \hat{\lambda}_1 \sum_{i=1}^{n_2} x_{1i}^{-\hat{\alpha}} \ln x_{1i} + (\hat{\lambda}_1 + \hat{\lambda}_2 + \hat{\lambda}_3) \sum_{i=1}^{n_3} x_i^{-\hat{\alpha}} \ln x_i = 0.\end{aligned} \tag{4.3}$$

These four equations have not explicit form; therefore, their solutions are numerically obtained using Newton-Raphson method as will be seen in Section 5. They are solved simultaneously to obtain $\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3$ and $\hat{\alpha}$.

The asymptotic variance-covariance matrix of $(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\alpha})$ is obtained by inverting the Fisher information matrix with elements that are negatives of expected values of the second order derivatives of logarithms of the likelihood function. In the present situation, it seems appropriate to approximate the expected values by their maximum likelihood estimates (see Cohen [18]). Accordingly; we have the following approximate variance-covariance matrix

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}^{-1}$$

where

$$a_{11} = -\frac{\partial^2 \ln L}{\partial \lambda_1^2}\bigg|_{\hat{\lambda}_1,\hat{\lambda}_2,\hat{\lambda}_3,\hat{\alpha}} = \frac{n_1}{(\hat{\lambda}_1+\hat{\lambda}_3)^2} + \frac{n_2}{\hat{\lambda}_1^2},$$

$$a_{12} = -\frac{\partial^2 \ln L}{\partial \lambda_1 \partial \lambda_2}\bigg|_{\hat{\lambda}_1,\hat{\lambda}_2,\hat{\lambda}_3,\hat{\alpha}} = 0,$$

$$a_{13} = -\frac{\partial^2 \ln L}{\partial \lambda_1 \partial \lambda_3}\bigg|_{\hat{\lambda}_1,\hat{\lambda}_2,\hat{\lambda}_3,\hat{\alpha}} = \frac{n_1}{(\hat{\lambda}_1+\hat{\lambda}_3)^2}$$

$$a_{14} = -\frac{\partial^2 \ln L}{\partial \lambda_1 \partial \alpha}\bigg|_{\hat{\lambda}_1,\hat{\lambda}_2,\hat{\lambda}_3,\hat{\alpha}} = -\sum_{i=1}^{n_2} x_{1i}^{-\hat{\alpha}} \ln x_{1i} - \sum_{i=1}^{n_2} x_{1i}^{-\hat{\alpha}} \ln x_{1i} - \sum_{i=1}^{n_3} x_i^{-\hat{\alpha}} \ln x_i,$$

$$a_{22} = -\frac{\partial^2 \ln L}{\partial \lambda_2^2}\bigg|_{\hat{\lambda}_1,\hat{\lambda}_2,\hat{\lambda}_3,\hat{\alpha}} = \frac{n_2}{(\hat{\lambda}_2+\hat{\lambda}_3)^2} + \frac{n_1}{\hat{\lambda}_2^2},$$

$$a_{23} = -\frac{\partial^2 \ln L}{\partial \lambda_2 \partial \lambda_3}\bigg|_{\hat{\lambda}_1,\hat{\lambda}_2,\hat{\lambda}_3,\hat{\alpha}} = \frac{n_2}{(\hat{\lambda}_2+\hat{\lambda}_3)^2},$$

$$a_{24} = -\frac{\partial^2 \ln L}{\partial \lambda_2 \partial \alpha}\bigg|_{\hat{\lambda}_1,\hat{\lambda}_2,\hat{\lambda}_3,\hat{\alpha}} = -\sum_{i=1}^{n_1} x_{2i}^{-\hat{\alpha}} \ln x_{2i} - \sum_{i=1}^{n_2} x_{2i}^{-\hat{\alpha}} \ln x_{2i} - \sum_{i=1}^{n_3} x_i^{-\hat{\alpha}} \ln x_i,$$

$$a_{33} = -\frac{\partial^2 \ln L}{\partial \lambda_3^2}\bigg|_{\hat{\lambda}_1,\hat{\lambda}_2,\hat{\lambda}_3,\hat{\alpha}} = \frac{n_1}{(\hat{\lambda}_1+\hat{\lambda}_3)^2} + \frac{n_2}{(\hat{\lambda}_2+\hat{\lambda}_3)^2} + \frac{n_3}{\hat{\lambda}_3^2},$$

$$a_{34} = -\frac{\partial^2 \ln L}{\partial \lambda_3 \partial \alpha}\bigg|_{\hat{\lambda}_1,\hat{\lambda}_2,\hat{\lambda}_3,\hat{\alpha}} = -\sum_{i=1}^{n_1} x_{1i}^{-\hat{\alpha}} \ln x_{1i} - \sum_{i=1}^{n_2} x_{2i}^{-\hat{\alpha}} \ln x_{2i} - \sum_{i=1}^{n_3} x_i^{-\hat{\alpha}} \ln x_i,$$

$$a_{44} = -\frac{\partial^2 \ln L}{\partial \alpha^2}\bigg|_{\hat{\lambda}_1,\hat{\lambda}_2,\hat{\lambda}_3,\hat{\alpha}} = \frac{2n_1 + 2n_2 + n_3}{\hat{\alpha}^2} + (\hat{\lambda}_1 + \hat{\lambda}_3)\sum_{i=1}^{n_1} x_{1i}^{-\hat{\alpha}} (\ln x_{1i})^2 + \hat{\lambda}_2 \sum_{i=1}^{n_1} x_{2i}^{-\hat{\alpha}} (\ln x_{2i})^2$$

$$+ (\hat{\lambda}_2 + \hat{\lambda}_3)\sum_{i=1}^{n_2} x_{2i}^{-\hat{\alpha}} (\ln x_{2i})^2 + \hat{\lambda}_1 \sum_{i=1}^{n_2} x_{1i}^{-\hat{\alpha}} (\ln x_{1i})^2 + (\hat{\lambda}_1 + \hat{\lambda}_2 + \hat{\lambda}_3)\sum_{i=1}^{n_3} x_i^{-\hat{\alpha}} (\ln x_i)^2.$$

Now we state the asymptotic normality results to obtain the asymptotic confidence intervals of $\alpha, \lambda_1, \lambda_2$ and λ_3. It can be stated as follows

$$\sqrt{n}\,[(\underline{\hat{\alpha}} - \underline{\alpha}), (\hat{\lambda}_1 - \lambda_1), (\hat{\lambda}_2 - \lambda_2), (\hat{\lambda}_3 - \lambda_3)] \to N_4(0\,,\, I(\underline{\theta})^{-1}) \quad \text{as} \quad n \to \infty \qquad (4.4)$$

Where $I^{-1}(\underline{\theta})$ is the variance-covariance matrix, $\underline{\theta} = (\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\alpha})$ and

$\underline{\theta} = (\lambda_1, \lambda_2, \lambda_3, \alpha)$ .Since $\underline{\theta}$ is unknown in (4.4), $I^{-1}(\underline{\theta})$ is estimated by $I^{-1}(\underline{\hat{\theta}})$; the asymptotic variance-covariance matrix that defined above and this can be used to obtain the asymptotic confidence intervals of $\alpha, \lambda_1, \lambda_2$ and λ_3.

5 Simulation Results

In this Section we present, a Monte Carlo simulation experiment in which we evaluated the estimation of the model parameters of the BIW distribution by considering the direct maximization of the log-likelihood. The simulations were performed using the Mathcad program, the number of Monte Carlo replications $R = 1000$ and the tolerance level was 0.001.

The evaluation of the point estimation was performed based on the following quantities for each sample size: the Average Estimates (AE) and the Mean Squared Error, (MSE) are estimated from R Monte Carlo replications and the coverage rate of the 95% confidence interval for $\lambda_1, \lambda_2, \lambda_3$ and α. We set the sample size at $n = 20{,}40{,}60{,}80$ and 100, and considered some values for the parameters $\lambda_1, \lambda_2, \lambda_3$ and α.

It can be seen from Table 1 that the estimates are slightly positively biased and that the MSE decreases as the sample size increases, as expected. Note that the estimates are quite stable and, more importantly, are close to the true values for the sample sizes considered. We also notice that the coverage probabilities of the asymptotic confidence interval are close to the nominal level. These results indicate that the proposed model and the asymptotic approximation work well under the situation where no censoring occurs.

6 Applications to Real Data Set

For illustrative purposes we have analyzed one data set to see how the proposed model works in practice. The data set has been obtained from Meintanis [19]. The data represent the football (soccer) data where at least one goal scored by the home team and at least one goal scored directly from a penalty kick, foul kick or any other direct kick (all of them together will be called as kick goal) by any team have been considered. Here X_1 represents the time in minutes of the first kick goal scored by any team and X_2 represents the first goal of any type scored by the home team. In this case all possibilities are open, for example $X_1 < X_2$ or $X_1 > X_2$ or $X_1 = X_2 = X$. All the data points have been divided by 100 so that the shape and scale parameters are of the same order.

This is not going to make any difference in any statistical inference. Meintanis [19] used the Marshal-Olkin distribution to analyze the data. We would like to analyze the data using BIW model. Before going to analyze the data using BIW model, we fit the IW distribution to X_1 and X_2 separately. The MLEs of the parameters α and λ of the respective IW distribution for X_1 , X_2 and $\max(X_1, X_2)$ are (1.96, 1.056), (2.658, 1.036) and (2.86, 1.122) respectively. The Kolmogorov-Smirnov distances between the fitted distribution and the empirical distribution function and the corresponding p values (in brackets) for X_1 , X_2 and $\max(X_1, X_2)$ are 0.122 (0.556) , 0.131(0.543) and 0.165(0.041) respectively. Based on the p values IW distribution cannot be rejected for the marginals and for the maximum also. Now from the Kolmogorov-Smirnov distances, it is clear that although Meintanis [19] suggested using bivariate exponential (BE) distribution, BIW model is preferable in this case.

Now we try to test whether BIW model or BE model provides better fit to the above data set. The Akaike information criterion (AIC), Bayesian information criterion (BIC), the consistent Akaike information criterion (CAIC) and Hannan-Quinn information criterion (HQIC) are used to compare the candidate distributions. In case of BIW, based on the above estimates the log-likelihood value is -30.25 and in case of BE model, using the estimates obtained by Meintanis [19], the log-likelihood value becomes -44.56. The corresponding AIC, BIC, CAIC and HQIC values are (53.12, 51.80, 53.01 and 58.15) and (96.12, 95.46, 96.11 and 97.62) respectively. Therefore, all of the criteria suggest that BIW provides a better fit than the BE model.

Table 1: The average estimates (AE), the mean squared errors (MSE), and the coverage percentages (CI) of $\lambda_1, \lambda_2, \lambda_3$ and α for BIW model

n	parameters	AE	MSE	95% CI Coverage
20	λ_1	1.0956	0.01240	0.95
	λ_2	1.056	0.01596	0.91
	λ_3	1.2	0.01554	0.92
	α	0.36	0.0019	0.97
40	λ_1	1.0613	0.00303	0.96
	λ_2	1.035	0.00648	0.93
	λ_3	1.156	0.00615	0.95
	α	0.315	0.0014	0.95
60	λ_1	1.072	0.00123	0.95
	λ_2	1.0113	0.00421	0.95
	λ_3	1.0265	0.00610	0.93
	α	0.2990	0.00012	0.97
80	λ_1	1.0036	0.00120	0.97
	λ_2	1.00406	0.00211	0.98
	λ_3	0.9995	0.00412	0.95
	α	0.2007	0.00011	0.96
100	λ_1	1.001	0.0002	0.94
	λ_2	1.003	0.0001	0.92
	λ_3	0.999	0.0009	0.96
	α	0.2000	0.00008	0.95

7 Conclusions

In this paper we have proposed bivariate inverse Weibull distribution whose marginals are inverse Weibull distributions. It is observed that the BIW distribution is a singular distribution and it has an absolute continuous part and a singular part. Since the joint distribution function and the joint density function are in closed forms, therefore this distribution can be used in practice for non-negative and positively correlated random variables. We obtained the maximum likelihood estimates for the unknown parameters of this distribution and their approximate variance- covariance matrix. We perform some simulations to see the performances of the MLEs. One data set has been re-analyzed and it is observed that the bivariate inverse Weibull distribution provides a better fit than the bivariate exponential distribution. Along the same line as Block and Basu [20] bivariate exponential model, an absolute continuous version of the BIW also obtained. Work is in progress in this direction and it will be reported elsewhere.

References

[1] A.O. Langlands, S.J. Pocock, G.R. Kerr and S.M. Gore, Long term survival of patients with breast cancer: A study of curability of the disease. British Medical Journal. (1979), pp.1247-1251.

[2] S. Bennette, Log-logistic regression models for survival data. Applied Statistics. 32, (1983), pp.165-171.

[3] R.L. Prentice, Exponential survivals with censoring and explanatory variables. Biometrika. 60, (1973), pp. 279-288

[4] A. Z. Keller, , M. T. Giblin and N. R. Farnworth, Reliability Analysis of Commercial Vehicle Engines, Reliability Engineering, 10, (1985),15-25.

[5] D.N.P. Murthy, M. Xie and R. Jiang, *Weibull Models*. Wiley, New York, (2004)

[6] P. Erto, M. Rapone, Non-informative and practical Bayesian confidence bounds for reliable life in the Weibull model. Reliability Engineering. 7, (1984), pp.181-191.

[7] W. Nelson, *Applied Lifetime Data Analysis.* Wiley, New York, (1982).

[8] R.Calabria and G.Pulcini, Bayesian 2-sample prediction for the inverse Weibull distribution. Communications in Statistics: Theory and Methods. 23, (1994), pp.1811-1824.

[9] M. Aleem, The Bivariate Inverse Weibull Distribution Its Characteristics and Properties , Proceedings of the World Congress on Engineering, Volume I, (2012), pp.348

[10] E. J. Gumbel, Bivariate exponential distributions. Journal of the American Statistical Association, 55, (1960), pp. 698-707.

[11] A.W. Marshall and I. Olkin, A multivariate exponential distribution. Journal of the American Statistical Association, 62, (1967), pp.30- 44.

[12] A. Sarhan and N. Balakrishnan, A new class of bivariate distribution and its mixture. J. Multivariate Anal. 98, (2007), 1508–1527.

[13] D. Kundu and R.D. Gupta, Modified Sarhan–Balakrishnan singular bivariate distribution, J. Statist. Plann. Inference.140, (2010), 526–538.

[14] D. Kundu and A.K. Dey, Estimating the parameters of the Marshall–Olkin bivariate Weibull distribution by EM algorithm, Comput. Statist. Data Anal. 53, (2009), pp. 956–965.

[15] D. Kundu, and R.D. Gupta, Bivariate generalized exponential distribution. J.Multivariate Anal. 100, (2009), pp. 581–593.

[16] A. Sarhan, D.C. Hamilton, B. Smith and D. Kundu, The bivariate generalized linear failure rate distribution and its multivariate extension, Comput. Statist. Data Anal. 55, (2011), pp.644–654.

[17] B. Bemis, L.J. Bain and J.J. Higgins, Estimation and hypothesis testing for the parameters of a bivariate exponential distribution", Journal of the American Statistical Association. 67, (1972), pp.927-929.

[18] A. C. Cohen, Maximum likelihood estimation in the Weibull distribution based on complete and censored samples. Technometrics, 7, (1965), pp. 579-588.

[19] S. G. Meintanis, Test of fit for Marshall-Olkin distributions with applications. Journal of Statistical Planning and inference, 137, (2007), pp. 3954–3963.

[20] H. Block and A. P. Basu, A continuous bivariate exponential extension. Journal of the American Statistical Association, 69, (1974), pp. 1031-1037.

[21] R. Johnson, S. Kotz and N. Balakrishnan, *Continuous Univariate Distribution*, 2nd ed. Wiley and Sons, New York, (1995).

[22] D. Kundu, Bayesian inference and reliability sampling plan for Weibull distribution. Technometrics 50 (2), (2008), 144-154.

[23] D. Nordman and W.Q.Meeker, Weibull prediction intervals for a future number of failures. Technometrics. 44, (2002), 15-23.