A General Method of Construction of a Bivariate Lifetime Distribution with a Singular Component

Debasis Kundu *

Abstract

Marshall-Olkin bivariate exponential distribution is the most popular bivariate distribution with a singular component. Since then several other bivariate distributions with a singular component have been introduced in the literature. It is observed that there are mainly two main approaches to construct a bivariate distribution with a singular component. In this paper we have proposed a general method to construct a bivariate distribution with a singular component. All the existing bivariate distributions with a singular component can be obtained using this method. Moreover, more flexible bivariate distributions with a singular component also can be constructed using this method. It is a very simple procedure based on mixing. Using this approach, we have considered one special case, namely bivariate Weibull distribution, in detail. We have derived several properties of the proposed bivariate Weibull distribution and it seems to be more flexible than the popular Marshall-Olkin bivariate Weibull distribution. Maximum likelihood estimators can be obtained quite conveniently in this case. It can be used to model dependent competing risks data and it can be generalized to the multivariate set up also.

Key Words and Phrases: Marshall-Olkin bivariate exponential distribution; Block and Basu bivariate distributions;maximum likelihood estimators; competing risks.

AMS Subject Classifications: 62F10, 62F03, 62H12.

*Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India. E-mail: kundu@iitk.ac.in, Phone no. 91-512-2597141, Fax no. 91-512-2597500.

1 Introduction

Analyzing bivariate or multivariate data, particularly when they are dependent is a very important problem. It arises in different applications. Several bivariate and multivariate distributions have been proposed in the statistical literature. An extensive amount of work has been done in constructing different bivariate distributions developing their properties and providing various inferential procedures. Some of the well known bivariate distributions are bivariate normal, bivariate log-normal, bivariate-t, bivariate extreme value, bivariate gamma, bivariate exponential, bivariate logistic, bivariate Cauchy, bivariate beta, bivariate skew normal etc. An excellent review of all these bivariate distributions; method of constructions, properties and their applications can be obtained in the book by Balakrishnan and Lai [5]. Some of the recently developed bivariate or multivariate distributions are bivariate Birnbaum-Saunders distribution, bivariate weighted exponential distribution, bivariate generalized exponential distribution etc., see for example Kundu, Balakrishnan and Jamalizadeh [12], Al-Mutairi, Ghitany and Kundu [1], Kundu and Gupta [17] and see the references cited therein.

In all the above cases, the distributions have absolutely continuous cumulative distribution function (CDF). It means that the bivariate distribution function has a proper probability density function (PDF) with respect to a two dimensional Lebesgue measure. Moreover, if X and Y denote the marginals of the bivariate random variable (X, Y), then $P(X = Y) = 0$. Hence if there are ties in the data set, and if it is known that $P(X = Y) > 0$, then none of these distributions can be used to analyze these data sets. In many practical examples it has been observed that there are ties in the data set. It may happen due to truncation, or it may happen due to the physical process by which the data has been obtained, and it is known from the process that $P(X = Y) > 0$. Hence, to analyze these data sets we need a bivariate model so that $P(X = Y) > 0$.

Marshall and Olkin [22] first introduced such a model, and popularly it is known as the Marshall-Olkin bivariate exponential (MOBE) model. It is a bivariate distribution where the marginals are exponentials and in this case $P(X = Y) > 0$. It has a very interesting physical interpretation and it has an interesting connection with the homogeneous Poisson process also. In the same paper they have introduced bivariate Weibull model also, where the marginals are Weibull and in this case also $P(X = Y) > 0$. From now on we call this as the Marshall-Olkin bivariate Weibull (MOBW) model. For several properties and inferential issues one is referred to Lu [20, 21], Kundu and Dey [13, 18].

Several other such bivariate distributions have been introduced in the literature. For example, Barreto-Souza and Lemonte [6] introduced bivariate Kumaraswamy (BVK), bivariate Pareto (BVP), bivariate double generalized exponential (BDGE), bivariate exponentiated Frechet (BEF), bivariate Gumbel (BVG) distributions etc. Kundu and Gupta [14] proposed the bivariate generalized exponential (BVGE) distribution. Along the same line Sarhan-Balakrishnan bivariate (SBBV) distribution was suggested by Sarhan and Balakrishnan [27] and modified Sarhan-Balakrishnan bivariate (MSBB) distribution by Kundu and Gupta [15]. Similarly, Sarhan et al. [28] proposed bivariate generalized linear failure rate distribution and Muhammed [24] provided the bivariate inverse Weibull distribution. In all these cases the authors provided the method of constructions, derived several properties and developed inference procedures. See for example the recent article by Franco, Vivo and Kundu [10] in this respect.

It is observed that there are mainly two methods of constructions and they are mainly (a) minimization approach proposed by Marshall and Olkin [22] and (b) maximization approach proposed by Kundu and Gupta [14]. We will briefly describe them in Section 2. The aim of this paper is two fold. First we introduce a general method of construction of a bivariate distribution with a singular component. The method is very simple, and it is based on the mixture representation. All the existing bivariate distributions with a singular component can be obtained by using this method. Moreover, other more general bivariate distributions with a singular component can be obtained using the proposed method, which cannot be obtained by using the above two methods. The second aim of this paper is to consider one specific case, namely bivariate Weibull distribution with a singular component, which can be obtained by using the proposed method and discuss its properties. We call it as the BWE distribution. The well known MOBE and MOBW can be obtained as special cases of the proposed BWE distribution. It is observed that the maximum likelihood estimators of the unknown parameters can be obtained quite conveniently. Moreover, the proposed BWE model can be used quite conveniently to model dependent competing risks data.

The rest of the paper is organized as follows. In Section 2, we provide a brief background of the two different constructions of a bivariate distribution with a singular component. The general method is proposed in Section 3. The specific case, namely the BWE distribution is discussed in detail in Section 4. The analyses of two data sets have been presented in Section 5, and finally we conclude the paper in Section 6.

2 Background

In this section we provide briefly both the methods for constructing a bivariate distribution with a singular component can be constructed. We will be using the following notation. In this paper it is assumed that all the univariate random variables have non-negative support, although most of the results are valid for random variables with support on the whole real line. It is further assumed that all the univariate random variables are absolutely continuous and hence they have proper probability density functions. For a random variable X with parameter θ, the probability density function (PDF), the cumulative distribution function (CDF) and survival function (SF) will be denoted by $f_X(x;\theta)$, $F_X(x;\theta)$ and $S_X(x;\theta)$, respectively. Here the parameter θ can be vector valued also. A random variable X is said to follow an exponential distribution with the parameter λ, if the PDF of X is as follows:

$$f_{EX}(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & \text{if} \quad x > 0 \\ 0 & \text{if} \quad x \leq 0. \end{cases} \tag{1}$$

A random variable X with the PDF (1) will be denoted by EX(λ). A random variable X is said to follow a Weibull distribution with the scale parameter λ and the shape parameter α, if the PDF of X is as follows:

$$f_{WE}(x;\alpha,\lambda) = \begin{cases} \alpha\lambda x^{\alpha-1} e^{-\lambda x^{\alpha}} & \text{if} \quad x > 0 \\ 0 & \text{if} \quad x \leq 0. \end{cases} \tag{2}$$

A random variable X with the PDF (2) will be denoted by WE(α, λ). In this paper we will consider another lifetime distribution and it is known as the generalized exponential (GE) distribution. A two-parameter generalized exponential distribution is an extension of the one parameter exponential distribution and it has many properties which are very close to two-parameter gamma distribution. Since it has a very convenient PDF and CDF it can be used as an alternative to the gamma distribution. For a detailed discussion on GE distribution one is referred to the review article by Nadarajah [25]. A random variable X is said to follow a generalized exponential (GE) distribution with the scale parameter λ and the shape parameter α, if the PDF of X is as follows:

$$f_{GE}(x;\alpha,\lambda) = \begin{cases} \alpha\lambda e^{-\lambda x}(1-e^{-\lambda x})^{\alpha-1} & \text{if} \quad x > 0 \\ 0 & \text{if} \quad x \leq 0. \end{cases} \tag{3}$$

A random variable X with the PDF (3) will be denoted by GE(α, λ). Now we are going to define the MOBE, MOBW and BVGE distributions.

2.1 Minimization Approach

Suppose U_1, U_2 and U_3 are three independent random variables, then a bivariate random variable (X, Y) can be constructed as follows:

$$X = \min\{U_1, U_3\} \quad \text{and} \quad Y = \min\{U_2, U_3\}.$$

Marshall and Olkin [22] first proposed this method and it is assumed that U_1 follows ($\sim$) EX(λ_1), $U_2 \sim$ EX(λ_2) and $U_3 \sim$ EX(λ_3). Hence, the joint survival function, $S_{MOBE}(x,y) = P(X > x, Y > y)$, of (X,Y) for $x > 0, y > 0$, becomes

$$S_{MOBE}(x,y) = P(U_1 > x, U_2 > y, U_3 > z) = \begin{cases} e^{-(\lambda_1+\lambda_3)x-\lambda_2 y} & \text{if} \quad 0 < y < x < \infty \\ e^{-\lambda_1 x-(\lambda_2+\lambda_3)y} & \text{if} \quad 0 < x < y < \infty. \end{cases}$$

Here $z = \max\{x,y\}$. From now on, we call this as the Marshall-Olkin bivariate exponential (MOBE) distribution. The MOBE is not an absolutely continuous distribution, and in this case $P(X = Y) = \dfrac{\lambda_3}{\lambda_1+\lambda_2+\lambda_3} > 0$. It does not have a joint PDF with respect to a two-dimensional Lebesgue measure. It has the following joint PDF with respect to two-dimensional Lebesgue measure on $x \neq y$ and with respect to one-dimensional Lebesgue measure on $x = y$, see for example Bemis, Bain and Higgins [7]. The joint PDF of the MOBE is as follows:

$$f_{MOBE}(x,y) = \frac{\lambda_1+\lambda_2}{\lambda_1+\lambda_2+\lambda_3} f^{(ac)}_{MOBE}(x,y) + \frac{\lambda_3}{\lambda_1+\lambda_2+\lambda_3} f^{(si)}_{MOBE}(x,y),$$

where

$$f^{(ac)}_{MOBE}(x,y) = \begin{cases} \frac{\lambda_1+\lambda_2+\lambda_3}{\lambda_1+\lambda_2} f_{EX}(x;\lambda_1+\lambda_3) f_{EX}(y;\lambda_2) & \text{if} \quad x > y \\ \frac{\lambda_1+\lambda_2+\lambda_3}{\lambda_1+\lambda_2} f_{EX}(x;\lambda_1) f_{EX}(y;\lambda_2+\lambda_3) & \text{if} \quad y > x, \end{cases}$$

and

$$f^{(si)}_{MOBE}(x,y) = \begin{cases} f_{EX}(x;\lambda_1+\lambda_2+\lambda_3) & \text{if} \quad x = y \\ 0 & \text{if} \quad x \neq y. \end{cases}$$

It simply means

$$S_{MOBE}(x,y) = \frac{\lambda_1+\lambda_2}{\lambda_1+\lambda_2+\lambda_3} \int_x^\infty \int_y^\infty f^{(ac)}_{MOBE}(u,v)dudv + \frac{\lambda_3}{\lambda_1+\lambda_2+\lambda_3} \int_z^\infty f^{(si)}_{MOBE}(u,u)du.$$

It can be easily seen that $f^{ac}_{MOBE}(x,y)$ is a proper bivariate PDF. Now MOBW distribution can be obtained by using Weibull distributions instead of exponential distributions. The MOBW distribution can be defined as follows. Suppose $U_1 \sim$ WE(α, λ_1), $U_2 \sim$ WE(α, λ_2), $U_3 \sim$ WE(α, λ_3), and they are independently distributed, then (X,Y), where

$$X = \min\{U_1, U_3\} \qquad \text{and} \qquad Y = \min\{U_2, U_3\},$$

is said to have a MOBW distribution. The joint PDF of MOBW can be obtained as given below:

$$f_{MOBW}(x,y) = \frac{\lambda_1+\lambda_2}{\lambda_1+\lambda_2+\lambda_3} f^{(ac)}_{MOBW}(x,y) + \frac{\lambda_3}{\lambda_1+\lambda_2+\lambda_3} f^{(si)}_{MOBW}(x,y),$$

where

$$f^{(ac)}_{MOBW}(x,y) = \begin{cases} \frac{\lambda_1+\lambda_2+\lambda_3}{\lambda_1+\lambda_2} f_{WE}(x;\alpha,\lambda_1+\lambda_3) f_{WE}(y;\alpha,\lambda_2) & \text{if} \quad x > y \\ \frac{\lambda_1+\lambda_2+\lambda_3}{\lambda_1+\lambda_2} f_{WE}(x;\alpha,\lambda_1) f_{WE}(y;\alpha,\lambda_2+\lambda_3) & \text{if} \quad y > x, \end{cases}$$

and

$$f^{(si)}_{MOBW}(x,y) = \begin{cases} f_{WE}(x;\alpha,\lambda_1+\lambda_2+\lambda_3) & \text{if} \quad x = y \\ 0 & \text{if} \quad x \neq y. \end{cases}$$

Several other bivariate distributions, for example Sarhan-Balakrishnan bivariate distribution by Sarhan and Balakrishnan [27] or modified Sarhan-Balakrishnan bivariate distribution by Kundu and Gupta [15] have been obtained using the same procedure. Now we will provide the maximization method.

2.2 Maximization Approach

Suppose U_1, U_2 and U_3 are three independent random variables, then based on these three random variables, the following bivariate random variable (X, Y) can be constructed, where

$$X = \max\{U_1, U_3\} \qquad \text{and} \qquad Y = \max\{U_2, U_3\}.$$

Kundu and Gupta [14] first proposed this method to construct BVGE distribution, and it can be obtained as follows. Suppose $U_1 \sim \text{GE}(\alpha_1, \lambda)$, $U_2 \sim \text{GE}(\alpha_2, \lambda)$, $U_3 \sim \text{GE}(\alpha_3, \lambda)$, and they are independently distributed. They (X, Y) as defined above is said to have a BVGE distribution. The joint CDF of (X, Y) for $x > 0$ and $y > 0$, is

$$F_{BVGE}(x,y) = P(U_1 \le x, U_2 \le y, U_3 \le z) = \begin{cases} (1-e^{-\lambda x})^{\alpha_1}(1-e^{-\lambda y})^{\alpha_2+\alpha_3} & \text{if} \quad 0 < y < x < \infty \\ (1-e^{-\lambda x})^{\alpha_1+\alpha_3}(1-e^{-\lambda y})^{\alpha_2} & \text{if} \quad 0 < x \le y < \infty, \end{cases}$$

and the joint PDF becomes:

$$f_{BVGE}(x,y) = \frac{\alpha_1+\alpha_2}{\alpha_1+\alpha_2+\alpha_3} f^{(ac)}_{BVGEW}(x,y) + \frac{\alpha_3}{\alpha_1+\alpha_2+\alpha_3} f^{(si)}_{BVGE}(x,y),$$

where

$$f^{(ac)}_{BVGE}(x,y) = \begin{cases} \frac{\alpha_1+\alpha_2+\alpha_3}{\alpha_1+\alpha_2} f_{GE}(x;\alpha_1,\lambda) f_{WE}(y;(\alpha_2+\alpha_3),\lambda) & \text{if} \quad y < x \\ \frac{\alpha_1+\alpha_2+\alpha_3}{\alpha_1+\alpha_2} f_{GE}(x;(\alpha_1+\alpha_3),\lambda) f_{WE}(y;\alpha_2,\lambda) & \text{if} \quad x < y, \end{cases}$$

and

$$f^{(si)}_{BVGE}(x,y) = \begin{cases} f_{GE}(x;\alpha_1+\alpha_2+\alpha_3,\lambda) & \text{if} \quad x = y \\ 0 & \text{if} \quad x \neq y. \end{cases}$$

Several other bivariate distributions with a singular component have been developed by using this method, for example the bivariate inverse Weibull distribution by Muhammed [24], the bivariate generalized linear failure rate distribution proposed by Sarhan et al. [28] or a very general proportional reversed hazard bivariate model as suggested by Kundu and Gupta [15] etc. In the next section we propose a very general method of constructing a bivariate distribution with a singular component using that all these distributions can be obtained as special cases.

3 Proposed Method

A bivariate random variable (X,Y) is said to have a bivariate distribution with a singular component, if the joint PDF of (X,Y) can be written as

$$f_{X,Y}(x,y) = p f^{(ac)}_{X,Y}(x,y) + (1-p) f^{(si)}_{X,Y}(x,y). \tag{4}$$

Here, $0 < p < 1$, $f^{(ac)}_{X,Y}(x,y)$ is a proper two dimensional PDF, and $f^{(si)}_{X,Y}(x,x)$ is a proper one dimensional PDF, and $f^{(si)}_{X,Y}(x,y) = 0$, if $x \neq y$. Now it is very clear that with proper choice of p, $f^{(ac)}_{X,Y}(x,y)$ and $f^{(si)}_{X,Y}(x,y)$ it is possible to obtain all the existing bivariate distributions with a singular component.

Some of the advantages of the proposed class of distributions can be described as follows. Note that the bivariate Weibull geometric distribution which has been proposed by Kundu and Gupta [19] cannot be obtained by the above minimization or maximization approach, but with the proper choice of p, $f_{X,Y}^{(ac)}(x,y)$ and $f_{X,Y}^{(si)}(x,y)$, it is possible to obtain based on the proposed method. It may be mentioned that all the existing bivariate distributions with a singular component have positive correlation coefficient, and it is due to construction. It is not possible to obtain negative correlation between the two variables based on the existing methods. But in our proposed method it is not a restriction. It is possible to obtain correlation on the entire range. For example, we can construct $f_{X,Y}^{(ac)}(x,y)$ based on a Gaussian copula with any marginal distribution function, and it is possible to obtain the correlation on the entire range, namely from (-1,1).

The following interpretation can be given for the proposed model. Suppose U, V and W are three random variables, and (U,V) has a joint PDF $f_{U,V}(u,v)$ and W has the PDF $f_W(w)$. We consider the following bivariate random variable (X,Y) as follows:

$$(X,Y) = \begin{cases} (U,V) & \text{with probability } p \\ (W,W) & \text{with probability } 1-p, \end{cases}$$

then for $0 < p < 1$ and if $f_{U,V}(u,v) = f_{X,Y}^{(ac)}(u,v)$, $f_W(w) = f_{X,Y}^{(si)}(w,w)$, then (X,Y) will have the same joint PDF as in (4).

In this section we mainly derive some of the general properties of this proposed bivariate distribution and in the subsequent sections we consider one specific case and discuss its properties. The joint CDF and the joint SF of (X,Y) will be denoted by $F_{X,Y}(x,y)$ and $S_{X,Y}(x,y)$, respectively. The joint CDF and the joint SF corresponding to the joint PDF $f_{X,Y}^{(ac)}(x,y)$ will be denoted by $F_{X,Y}^{(ac)}(x,y)$ and $S_{X,Y}^{(ac)}(x,y)$, respectively. Similarly, the CDF and SF of the PDF $f_{X,Y}^{(si)}(x,x)$ will be denoted by $F_{X,Y}^{(si)}(x)$ and $S_{X,Y}^{(si)}(x)$, respectively. Therefore, the joint CDF and the joint SF of (X,Y) can be written as

$$F_{X,Y}(x,y) = pF_{X,Y}^{(ac)}(x,y) + (1-p)F_{X,Y}^{(si)}(\min\{x,y\}) \quad \text{and} \tag{5}$$

$$S_{X,Y}(x,y) \quad = \quad pS^{(ac)}_{X,Y}(x,y) + (1-p)S^{(si)}_{X,Y}(\max\{x,y\}), \tag{6}$$

respectively. Hence, the CDF of X and Y can be obtained as

$$\begin{aligned} P(X \le x) &= F_X(x) = pF^{(ac)}_{X,Y}(x,\infty) + (1-p)F^{(si)}_{X,Y}(x,x), \qquad \text{and} \\ P(Y \le y) &= F_Y(y) = pF^{(ac)}_{X,Y}(\infty,y) + (1-p)F^{(si)}_{X,Y}(y,y). \end{aligned}$$

Therefore, it is clear that in general it will be a mixture distribution, but with the proper choice of p, $F^{(ac)}_{X,Y}(x,y)$ and $F^{(si)}_{X,Y}(x)$, it may not be a mixture distribution. Some of the properties can be easily obtained for (X,Y), such that $P(X=Y)=p$, $P(X \ne Y) = 1-p$, and

$$\begin{aligned} P(X < Y) &= p\int_0^\infty \int_x^\infty f^{(ac)}_{X,Y}(u,v)dvdu \\ P(X > Y) &= p\int_0^\infty \int_0^x f^{(ac)}_{X,Y}(u,v)dvdu. \end{aligned}$$

We can further obtain quite convenoently the distribution of $\max\{X,Y\}$ and $\min\{X,Y\}$, i.e.

$$\begin{aligned} P(\max\{X,Y\} \le x) &= pF^{(ac)}_{X,Y}(x,x) + (1-p)F^{(si)}_{X,Y}(x,x) \quad \text{and} \\ P(\min\{X,Y\} \ge x) &= pS^{(ac)}_{X,Y}(x,x) + (1-p)S^{(si)}_{X,Y}(x,x). \end{aligned}$$

This model can be used quite conveniently for modeling data from a dependent series system, dependent parallel system, analyzing dependent competing risks data and also dependent complementory risks data. We will explain those in details in the subsequent sections.

4 Bivariate Weibull Distribution

4.1 Joint, Marginal and Conditional PDFs

The main aim of this section is to define a bivariate Weibull (BWE) distribution based on the proposed method, which has close similarity with the popular MOWE distribution, but it is more flexible than the MOBW distribution. We discuss different properties of the BWE distribution and provide the estimation method. We will also explore how this model can be used to analyze dependent competing risks data.

Consider the bivariate distribution which has the following absolute continuous part and the singular part.

$$f_{X,Y}^{(ac)}(x,y) = c\begin{cases} f_{WE}(x;\alpha,\delta_1)f_{WE}(y;\alpha,\delta_2) & \text{if} \quad x < y \\ f_{WE}(x;\alpha,\delta_3)f_{WE}(y;\alpha,\delta_4) & \text{if} \quad y < x, \end{cases} \tag{7}$$

where $c^{-1} = \dfrac{\delta_1}{\delta_1+\delta_2} + \dfrac{\delta_4}{\delta_3+\delta_4}$, and

$$f_{X,Y}^{(si)}(x,y) = \begin{cases} f_{WE}(x;\alpha,\delta_5) & \text{if} \quad x = y \\ 0 & \text{if} \quad x \neq y. \end{cases} \tag{8}$$

Then the random variable (X,Y) has the following joint PDF of (X,Y):

$$f_{X,Y}(x,y) = pf_{X,Y}^{(ac)}(x,y) + (1-p)f_{X,Y}^{(si)}(x,y). \tag{9}$$

We make the following restrictions on the parameter.

$$\delta_1+\delta_2 = \delta_3+\delta_4 = \theta \quad \text{(say)} \qquad \text{and} \qquad \delta_5 = \delta_1+\delta_2 = \theta. \tag{10}$$

These restrictions have been made so that the proposed bivariate distribution has a similar structure as the MOBW distribution and at the same time it is more flexible than the later. From now on a bivariate distribution with the joint PDF (9) and with the restriction (10) will be called BWE distribution. It may be mentioned that $f_{X,Y}^{(ac)}(x,y)$ is continuous for all $0 < x, y < \infty$, when $\delta_1 + \delta_2 = \delta_3 + \delta_4$ and $|\delta_1 - \delta_2| = |\delta_3 - \delta_4|$, similar to the MOBW distribuion, otherwise it is not continuous on the line $x = y$. When $\alpha = 1$, we call it as the bivariate exponential (BEX) distribution.

The corresponding survival function, $S_{X,Y}(x,y)$ for $x \leq y$ becomes:

$$S_{X,Y}(x,y) = pc\left\{(e^{-\delta_1 x^\alpha} - e^{-\delta_1 y^\alpha})e^{-\delta_2 y^\alpha} + \frac{\delta_1+\delta_4}{\theta}e^{-\theta y^\alpha}\right\} + (1-p)e^{-\theta y^\alpha} \tag{11}$$

and for $x > y$,

$$S_{X,Y}(x,y) = pc\left\{(e^{-\delta_4 y^\alpha} - e^{-\delta_4 x^\alpha})e^{-\delta_3 x^\alpha} + \frac{\delta_1+\delta_4}{\theta}e^{-\theta x^\alpha}\right\} + (1-p)e^{-\theta x^\alpha}. \qquad (12)$$

It can be easily seen that if we take:

$$\delta_1 = \lambda_1, \quad \delta_2 = \lambda_2 + \lambda_3, \quad \delta_3 = \lambda_1 + \lambda_3, \quad \delta_4 = \lambda_2, \quad p = \frac{\lambda_1+\lambda_2}{\lambda_1+\lambda_2+\lambda_3},$$

it satisfies the constraint (10), and $S_{X,Y}(x,y)$ becomes

$$S_{X,Y}(x,y) = \begin{cases} e^{-\lambda_1 x^\alpha}e^{-(\lambda_2+\lambda_3)y^\alpha} & \text{if} \quad x \le y \\ e^{-(\lambda_1+\lambda_3)x^\alpha}e^{-\lambda_2 y^\alpha} & \text{if} \quad x > y. \end{cases}$$

It shows that the proposed BWE distribution becomes the MOBW distribution, and when $\alpha = 1$, it matches with the joint PDF of the MOBE distribution.

It may be noted that the proposed BWE distribution is more flexible than the MOBW distribution, as the former has one extra parameter. Moreover, it is well known that the estimation of the unknown parameters in case of MOBW distribution is not a trivial issue, see for example Kundu and Dey [13], where as it is observed that in case of BWE distribution, the maximum likelihood estimators can be obtained in a routine manner. Hence, the proposed BWE distribution can be used quite effectively for data analysis purposes.

Now we consider the marginal distribution functions of BWE.

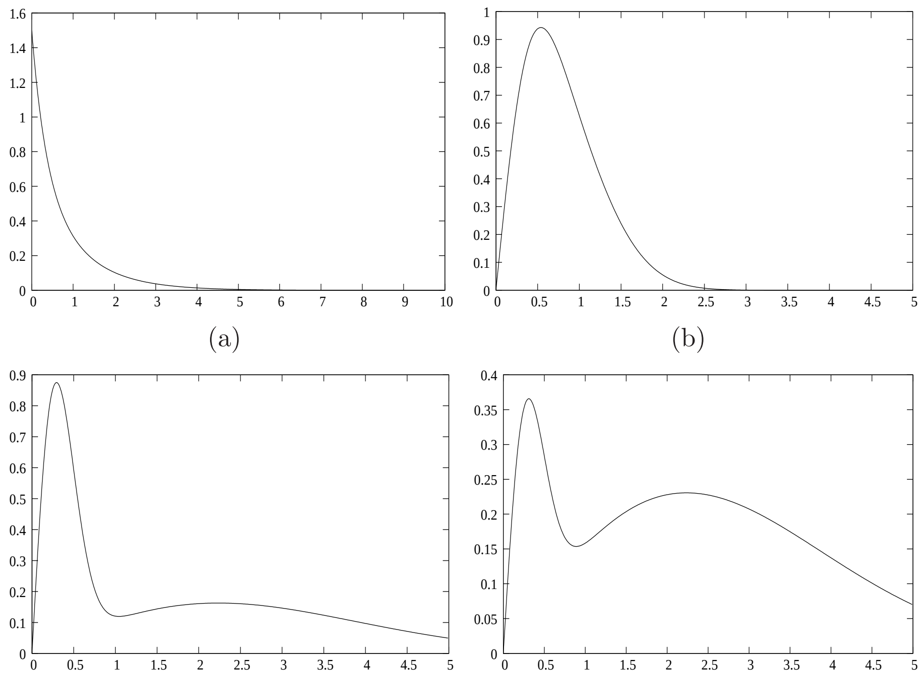
$$\begin{aligned} S_X(x) &= P(X > x) = p\frac{\delta_1+\delta_4}{\theta}e^{-\delta_3 x^\alpha} + \left(1 - p\frac{\delta_1+\delta_4}{\theta}\right)e^{-\theta x^\alpha} \\ S_Y(y) &= P(Y > y) = p\frac{\delta_1+\delta_4}{\theta}e^{-\delta_2 y^\alpha} + \left(1 - p\frac{\delta_1+\delta_4}{\theta}\right)e^{-\theta y^\alpha}. \end{aligned}$$

It is interesting to see that if $p\dfrac{\delta_1+\delta_4}{\delta_1+\delta_2} < 1$, then the marginal distribution functions can be written as the mixture of two Weibull distribution functions, and if $p\dfrac{\delta_1+\delta_4}{\delta_1+\delta_2} > 1$, then it can be written as the generalized mixture of Weibull distributions, see for example Franco et al. [9]. The PDFs of X and Y can be written as

$$f_X(x) = p\frac{\delta_1+\delta_4}{\theta}f_{WE}(x;\alpha,\delta_3) + \left(1 - p\frac{\delta_1+\delta_4}{\theta}\right)f_{WE}(x;\alpha,\theta)$$

$$f_Y(y) \;=\; p\frac{\delta_1+\delta_4}{\theta} f_{WE}(y;\alpha,\delta_2) + \left(1 - p\frac{\delta_1+\delta_4}{\theta}\right) f_{WE}(y;\alpha,\theta),$$

respectively. Since the marginals are mixtures of Weibull distributions the PDFs can take variety of shapes. It can be increasing, decreasing and even bimodal also. Moreover, the hazard functions of the marginals also can be of different types. The PDFs and hazard functions of the marginals for different parameter values have been plotted in Figure 1. It may be mentioned that Figures 1(a)-1(c) correspond to mixture of Weibull distributions where as Figure 1(d) corresonds to generalized mixture of Weibull distributions.



(a) (b) (c) (d)

Figure 1: The PDF plot of the marginal distribution of X, when (a) α = 1.0, δ_1 = 2.0, δ_2 = 1.0, δ_3 = 1.0, δ_4 = 2.0, p = 0.75, (b) α = 2.0, δ_1 = 2.0, δ_2 = 1.0, δ_3 = 1.0, δ_4 = 2.0, p = 0.75, (c) α = 2, δ_1 = 3.0, δ_2 = 3.0, δ_3 = 0.1, δ_4 = 5.9, p = 0.6, (d) α = 2, δ_1 = 3.0, δ_2 = 3.0, δ_3 = 0.1, δ_4 = 5.9, p = 0.85.

In Figure 2 we provide the hazard functions of the marginal distribution X for different parameter values. It is clear that it can take variety of shapes. In this case also Figures 2(a) - 2(c) corresond to the mixture of Weibull distributions, and Figure 2(d) corresonds to the generalized mixture of Weibull distributions.

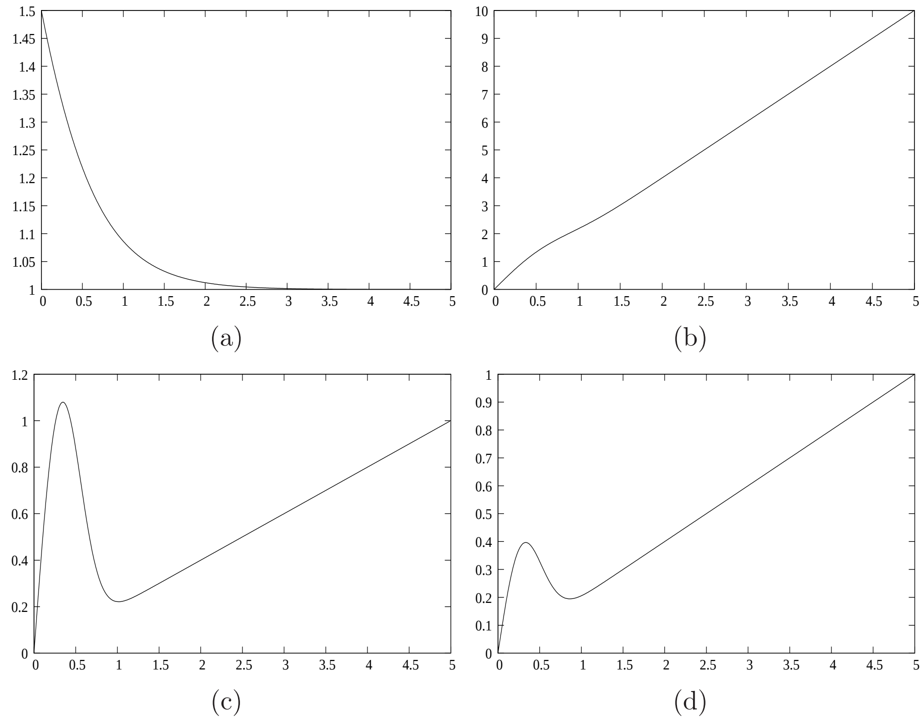


Figure 2: The hazard function plot of the marginal distribution of X, when (a) $\alpha = 1.0$, $\delta_1 = 2.0$, $\delta_2 = 1.0$, $\delta_3 = 1.0$, $\delta_4 = 2.0$, $p = 0.75$, (b) $\alpha = 2.0$, $\delta_1 = 2.0$, $\delta_2 = 1.0$, $\delta_3 = 1.0$, $\delta_4 = 2.0$, $p = 0.75$, (c) $\alpha = 2$, $\delta_1 = 3.0$, $\delta_2 = 3.0$, $\delta_3 = 0.1$, $\delta_4 = 5.9$, $p = 0.6$, (d) $\alpha = 2$, $\delta_1 = 3.0$, $\delta_2 = 3.0$, $\delta_3 = 0.1$, $\delta_4 = 5.9$, $p = 0.85$.

Now we discuss some conditional distributions which will be of interest in data analysis, and it may have some independent interests also. For example, if (X, Y) has a BWE as described in (9) then it can be easily seen that

$$X|\{X < Y\} \sim \text{WE}(\alpha, \theta) \quad \text{and} \quad Y|\{Y < X\} \sim \text{WE}(\alpha, \theta). \tag{13}$$

Moreover, the conditional PDF of $Y|\{X < Y\}$ and $X|\{Y < X\}$ can be written as

$$f_{Y|\{X<Y\}}(y) = \frac{\theta}{\delta_1}\alpha\delta_2 y^{\alpha-1} e^{-\delta_2 y^{\alpha}}(1 - e^{-\delta_1 y^{\alpha}}) \tag{14}$$

$$f_{X|\{Y<X\}}(y) = \frac{\theta}{\delta_4}\alpha\delta_3 y^{\alpha-1} e^{-\delta_3 y^{\alpha}}(1 - e^{-\delta_4 y^{\alpha}}). \tag{15}$$

It can be seen that (14) and (15) are the PDFs of the weighted Weibull (WWE) distribution, as introduced by Gupta and Kundu [11], see also Shahbaz, Shahbaz and Butt [29] in this respect. It may be mentioned that the WWE distribution has an interesting interpretation similar to the skew normal distribution as introduced by the Azzalini [3]. For several interesting properties on WWE distribution, one may refer to Al-Mutairi, Ghitany and Kundu [2].

4.2 Absolute Continuous Part of a BWE and its Generation

In this section we study some basic feature of the absolute continuous part of the proosed BWE distribution. Suppose (X, Y) has a BWE distribution with the absolute continuous part as given in (7). Let us assume that an absolute continuous bivariate random variable (U, V) has the joint PDF

$$f_{U,V}(u, v) = \frac{\theta}{\delta_1 + \delta_4} \begin{cases} f_{WE}(u; \alpha, \delta_1) f_{WE}(v; \alpha, \delta_2) & \text{if} \quad u < v \\ f_{WE}(u; \alpha, \delta_3) f_{WE}(v; \alpha, \delta_4) & \text{if} \quad v < u. \end{cases} \tag{16}$$

It may be mentioned that the joint PDF of (U, V) may be compared with the joint PDF of the Block and Basu bivariate Weibull (BBBW) distribution. The BBBW distribution can be obtanied from a MOBW distribution by removing the singular component, see for example the original article by Block and Basu [8], see also Pradhan and Kundu [26] in this respect. It may be recalled that joint PDF of a BBBW distribution with parameters $\alpha, \lambda_0, \lambda_1, \lambda_2$, can be written as follows:

$$f_{BBBW}(u, v) = \frac{\lambda_0 + \lambda_1 + \lambda_2}{\lambda_1 + \lambda_2} \begin{cases} f_{WE}(u; \alpha, \lambda_1) f_{WE}(v; \alpha, \lambda_0 + \lambda_2) & \text{if} \quad u < v \\ f_{WE}(v; \alpha, \lambda_0 + \lambda_1) f_{WE}(v; \alpha, \lambda_2) & \text{if} \quad v < u. \end{cases} \tag{17}$$

It is clear that the joint PDF (17) can be obtained as a special case of (16). Now we provide the shape of the joint PDF of (16). It may be noted that if $\alpha \le 1$, then the joint PDF of (U, V) is a decreasing function both in u and v directions for all values of δ_1, δ_2, δ_3 and δ_4. Therefore, we are mainly interested when $\alpha > 1$ and then we have the following result.

Theorem 1: For $\alpha > 1$, if (a) $\delta_2 < \delta_1$ and $\delta_4 > \delta_3$, then it is a bimodal function and the two modes are at (i) $\left(\left[\frac{\alpha - 1}{\alpha \delta_1} \right]^{1/\alpha}, \left[\frac{\alpha - 1}{\alpha \delta_2} \right]^{1/\alpha} \right)$ and (ii) $\left(\left[\frac{\alpha - 1}{\alpha \delta_3} \right]^{1/\alpha}, \left[\frac{\alpha - 1}{\alpha \delta_4} \right]^{1/\alpha} \right)$, (b)

$\delta_1 < \delta_2$, $\delta_4 < \delta_3$, then it is unimodal, and the mode is at $\left(\left[\frac{2(\alpha-1)}{\alpha(\delta_1+\delta_2)} \right]^{1/\alpha}, \left[\frac{2(\alpha-1)}{\alpha(\delta_1+\delta_2)} \right]^{1/\alpha} \right)$, (c) $\delta_1 < \delta_2$, $\delta_4 > \delta_3$ and $\delta_1\delta_2 > \delta_3\delta_4$, then it is a bimodal function, and and the two modes are at (i) $\left(\left[\frac{2(\alpha-1)}{\alpha(\delta_1+\delta_2)} \right]^{1/\alpha}, \left[\frac{2(\alpha-1)}{\alpha(\delta_1+\delta_2)} \right]^{1/\alpha} \right)$ and (ii) $\left(\left[\frac{\alpha-1}{\alpha\delta_3} \right]^{1/\alpha}, \left[\frac{\alpha-1}{\alpha\delta_4} \right]^{1/\alpha} \right)$, (d) $\delta_1 < \delta_2$, $\delta_4 > \delta_3$ and $\delta_1\delta_2 < \delta_3\delta_4$, then it is unimodal and the mode is at $\left(\left[\frac{\alpha-1}{\alpha\delta_3} \right]^{1/\alpha}, \left[\frac{\alpha-1}{\alpha\delta_4} \right]^{1/\alpha} \right)$.

PROOF: The proof is not very difficult to obtain. It can be obtained by studying the derivatives of the log of the joint PDF of the BWE distribution. The details are avoided. ■

The following Figure 3 provides the contour plots of the absolute continuous part of the BWE distribution for different parameter values.

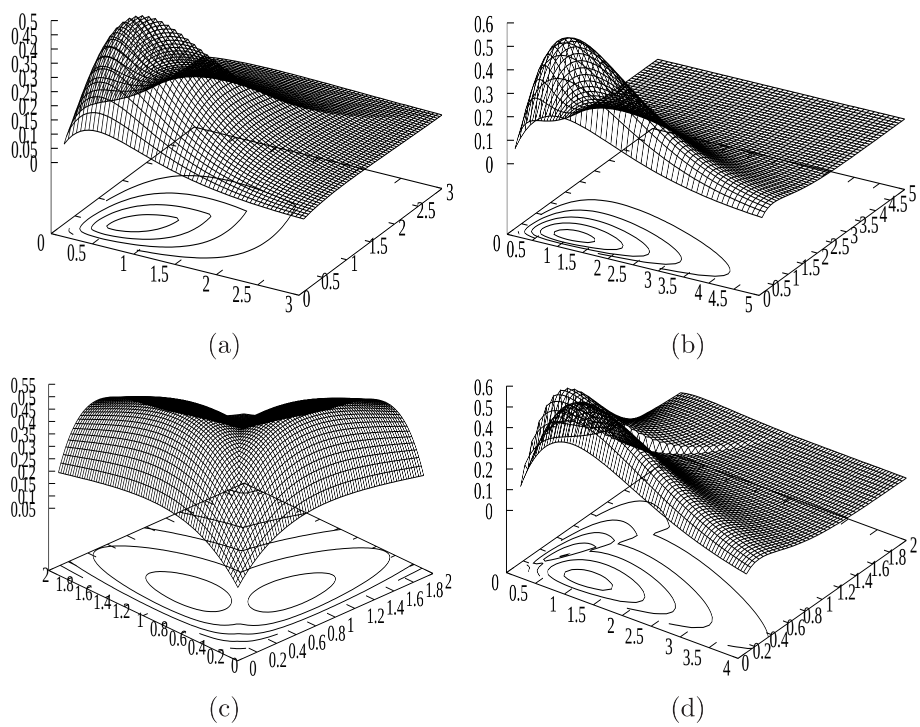


Figure 3: The contour plot of the joint PDF of the absolute continuous part of the BWE distribution when (a)$\alpha = 2$, $\delta_1 = 1.0$, $\delta_2 = 2.0$, $\delta_3 = 2.0$, $\delta_4 = 1.0$, (b) $\alpha = 2$, $\delta_1 = 1.0$, $\delta_2 = 2.0$, $\delta_3 = 1.0$, $\delta_4 = 2.0$, (c) $\alpha = 2$, $\delta_1 = 2.0$, $\delta_2 = 1.0$, $\delta_3 = 1.0$, $\delta_4 = 2.0$, (d) $\alpha = 2$, $\delta_1 = 2.0$, $\delta_2 = 3.0$, $\delta_3 = 1.0$, $\delta_4 = 4.0$.

It can be easily checked that if (U, V) has the joint PDF (16), then $P(U < V) = \frac{\delta_1}{\delta_1 + \delta_4}$

and $P(U > V) = \dfrac{\delta_4}{\delta_1 + \delta_4}$. The following decomposition is useful to generate samples from (U, V), they may have some independent interests also. The joint PDF of (U, V) can be written as follows:

$$f_{U,V}(u, v) = \frac{\delta_1}{\delta_1 + \delta_4} f_{U_1,V_1}(u, v) + \frac{\delta_4}{\delta_1 + \delta_4} f_{U_2,V_2}(u, v). \tag{18}$$

Here

$$f_{U_1,V_1}(u, v) = \begin{cases} \frac{\delta_1+\delta_2}{\delta_1} f_{WE}(u; \alpha, \delta_1) f_{WE}(v; \alpha, \delta_2) & \text{if} \quad u < v \\ 0 & \text{if} \quad u \geq v \end{cases} \tag{19}$$

and

$$f_{U_2,V_2}(u, v) = \begin{cases} \frac{\delta_3+\delta_4}{\delta_4} f_{WE}(u; \alpha, \delta_3) f_{WE}(v; \alpha, \delta_4) & \text{if} \quad u > v \\ 0 & \text{if} \quad u \leq v. \end{cases} \tag{20}$$

It follows that if (U_1, V_1) has a joint PDF (19), then $U_1 \sim \text{WE}(\alpha, \delta_1 + \delta_2)$ and $P(V_1 > v | U_1 = u) = e^{-\delta_2(v^\alpha - u^\alpha)}$, for $v > u$. Similarly, if (U_2, V_2) has a joint PDF (20), then $V_2 \sim \text{WE}(\alpha, \delta_3 + \delta_4)$ and $P(U_2 > u | V_2 = v) = e^{-\delta_3(u^\alpha - v^\alpha)}$, for $u > v$. It is quite simple to generate random samples from (U_1, V_1) and (U_2, V_2), and hence, generation random samples from (U, V) is straight forward.

Note that if (X, Y) has BWE distribution, then it can be written as follows:

$$(X, Y) = \begin{cases} (U_1, V_1) & \text{with probability} \quad \frac{p\delta_1}{\delta_1+\delta_4} \\ (U_2, V_2) & \text{with probability} \quad \frac{p\delta_4}{\delta_1+\delta_4} \\ (W, W) & \text{with probability} \quad 1 - p, \end{cases} \tag{21}$$

here (U_1, V_1) and (U_2, V_2) are same as defined above, and $W \sim \text{WE}(\alpha, \delta_1 + \delta_2)$. The above decomposition (21) can be used quite effectively to generate random samples from a BWE distribution.

We have provided the scatter plots of (X, Y) generated from BWE distribution for different parameter values in Figure 4. In each case we have reported the corresponding sampling correlation (r) also based 100 data points. It may be observed that the sample correlation coefficient can be negative also in this case for certain set of parameters.

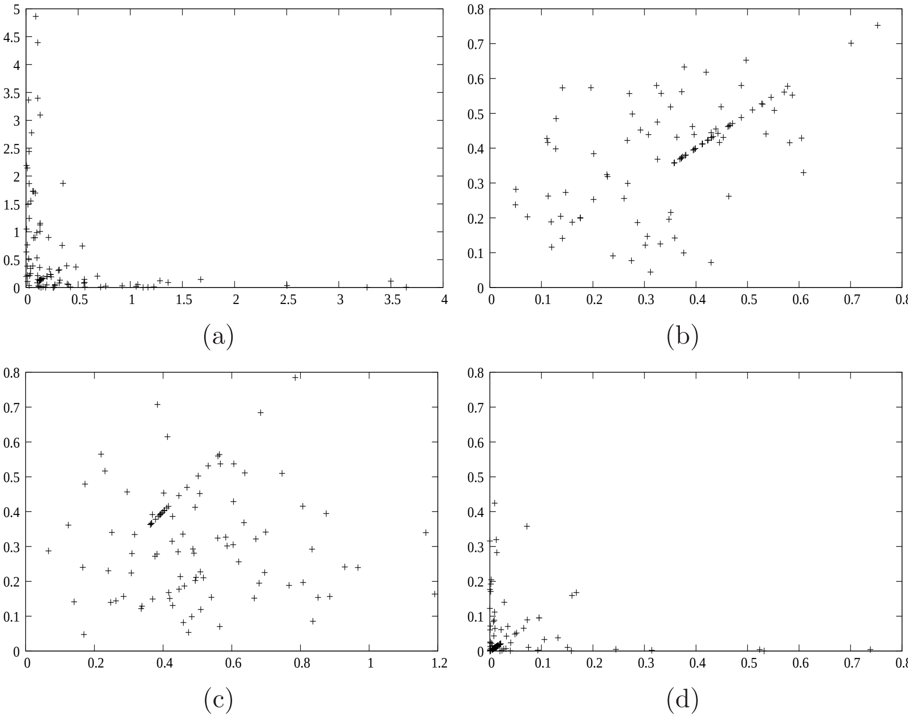


Figure 4: The scatter of (X, Y) generated from BWE distribution when (a) α = 1.0, δ_1 = 10.0, δ_2 = 1.0, δ_3 = 1.0, δ_4 = 10.0, p = 0.75, r = -0.32 (b) α = 2.0, δ_1 = 1.0, δ_2 = 10.0, δ_3 = 10.0, δ_4 = 1.0, p = 0.75, r = 0.71 (c) α = 2, δ_1 = 1.0, δ_2 = 10.0, δ_3 = 5.0, δ_4 = 6.0, p = 0.75, r = -0.28 (d) α = 0.5, δ_1 = 8.0, δ_2 = 3.0, δ_3 = 5.0, δ_4 = 6.0, p = 0.5, r = 0.009.

4.3 Maximum Likelihood Estimators

In this section we disucss about the maximum likelihood estimators of the unknown parameters. It is assumed we have a random sample of size n from a BWE distribution with the constraint on the parameters (10). Therefore, the proposed BWE distribution has five independent parameters. Let $\mathcal{D} = \{(x_1, y_1), \ldots, (x_n, y_n)\}$ be a random sample of size n from a BWE distribution, and we use the following notation

$$I_1 = \{i : x_i < y_i\}, \quad I_2 = \{i : x_i > y_i\} \quad I_0 = \{i : x_i = y_i = u_i\}.$$

and $|I_1| = n_1$, $|I_2| = n_2$, $|I_0| = n_0$. Moreover, we use $\theta = \delta_1 + \delta_2 = \delta_3 + \delta_4$, as before. Therefore, it is assumed that p, α, θ, δ_2, δ_3 are the independent parameters of the proposed model, and $\boldsymbol{\Theta} = (p, \alpha, \boldsymbol{\Gamma})^\top$, where $\boldsymbol{\Gamma} = (\theta, \delta_2, \delta_3)^\top$. Now based on the observed sample $\mathcal{D}$, the log-likelihood function can be written as

$$\begin{aligned} l(p, \alpha, \boldsymbol{\Gamma}|\mathcal{D}) \;=\; & (n_1 + n_2)\{\ln p - \ln(2\theta - \delta_2 - \delta_3) + \ln\theta\} + 2n_1 \ln\alpha + n_1 \ln\delta_2 + n_1 \ln(\theta - \delta_2) + \\ & (\alpha - 1)\sum_{i \in I_1 \cup I_2} \ln x_i - (\theta - \delta_2)\sum_{i \in I_1} x_i^\alpha + (\alpha - 1)\sum_{i \in I_1 \cup I_2} \ln y_i - \delta_2 \sum_{i \in I_1} y_i^\alpha + \\ & 2n_2 \ln\alpha + n_2 \ln\delta_3 + n_2 \ln(\theta - \delta_3) - \delta_3 \sum_{i \in I_2} x_i^\alpha - (\theta - \delta_3)\sum_{i \in I_2} y_i^\alpha + n_0 \ln(1 - p) + \\ & n_0 \ln\alpha + n_0 \ln\theta + (\alpha - 1)\sum_{i \in I_0} \ln u_i - \theta \sum_{i \in I_0} u_i^\alpha. \end{aligned} \tag{22}$$

The maximum likelihood estimators (MLEs) of $p, \alpha, \boldsymbol{\Gamma}$ can be obtained by maximizing $l(p, \alpha, \boldsymbol{\Gamma}|\mathcal{D})$ with respect to the unknown parameters. It can be easily seen that the MLE of p becomes

$$\widehat{p} = \frac{n_1 + n_2}{n},$$

and the MLEs of $(\alpha, \boldsymbol{\Gamma})^\top$ can be obtained by maximizing

$$\begin{aligned} l_0(\alpha, \boldsymbol{\Gamma}) \;=\; & (n_1 + n_2)\{-\ln(2\theta - \delta_2 - \delta_3) + \ln\theta\} + 2n_1 \ln\alpha + n_1 \ln\delta_2 + n_1 \ln(\theta - \delta_2) + \\ & \alpha \sum_{i \in I_1 \cup I_2} \ln x_i - (\theta - \delta_2)\sum_{i \in I_1} x_i^\alpha + \alpha \sum_{i \in I_1 \cup I_2} \ln y_i - \delta_2 \sum_{i \in I_1} y_i^\alpha + \\ & 2n_2 \ln\alpha + n_2 \ln\delta_3 + n_2 \ln(\theta - \delta_3) - \delta_3 \sum_{i \in I_2} x_i^\alpha - (\theta - \delta_3)\sum_{i \in I_2} y_i^\alpha + \\ & n_0 \ln\alpha + n_0 \ln\theta + \alpha \sum_{i \in I_0} \ln u_i - \theta \sum_{i \in I_0} u_i^\alpha. \end{aligned} \tag{23}$$

with respect to the unknown parameters. It is a four dimensional optimization problem. If we try to solve directly, then we can obtain the normal equations and we need to solve four non-linear equations simultaneously. To avoid that we have used the profile likelihood method, and the method provided by Song, Fan and Kalbfleisch [30]. It is observed that the the MLEs of the unknown parameters can be obtained by solving only one non-linear equation. The details are provided in Appendix A. Once we obtain the MLEs of the unknown parameters, the observed Fisher information matrix can be easily constructed, and the asymptotic confidence intervals of the unknown parameters also can be computed.

It may be noted that the algorithm which has been described in the previous section, is an iterative process, hence good initial estimates are needed for α, θ, δ_2 and δ_3. Now we describe how to obtain the initial guesses.

Now to compute α and θ, we use the results (13). The data points which are obtained as below

$$\{x_i : i \in I_0 \cup I_1\}, \quad \text{and} \quad \{y_i : i \in I_2\},$$

are i.i.d. WE(α, θ) random variables. Hence, the estimates of α and θ can be obtained quite conveniently as several efficient methods are available to estimate the shape and scale parameters of a Weibull distribution. Further, the initial estimates of δ_1 and δ_4 can be obtained from

$$\{x_i : i \in I_2\} \quad \text{and} \quad \{y_i : i \in I_1\},$$

and using the results (14) and (15). Note that $\{x_i : i \in I_2\}$ and $\{y_i : i \in I_1\}$, are WWE distributions. Hence, assuming the shape parameter α to be known, the method proposed by Gupta and Kundu [11] can be used quite conveniently to estimate δ_1 and δ_4.

5 Data Analysis

In this section we perform the analyses of two data sets, one simulated and one real data set.

5.1 Simulated Data Set:

We have generated a data set with $n = 50$, and $p = 0.75$, $\alpha = 1.5$, $\delta_1 = \delta_2 = \delta_3 = \delta_4 = 1.0$. The data set has been plotted in Figure 5 In this case the intitial estimates of α, θ, δ_1 and δ_4 can be obtained as

$$\widetilde{\alpha} = 1.142, \quad \widetilde{\theta} = 1.842, \quad \widetilde{\delta}_1 = 0.597, \quad \widetilde{\delta}_4 = 0.911.$$

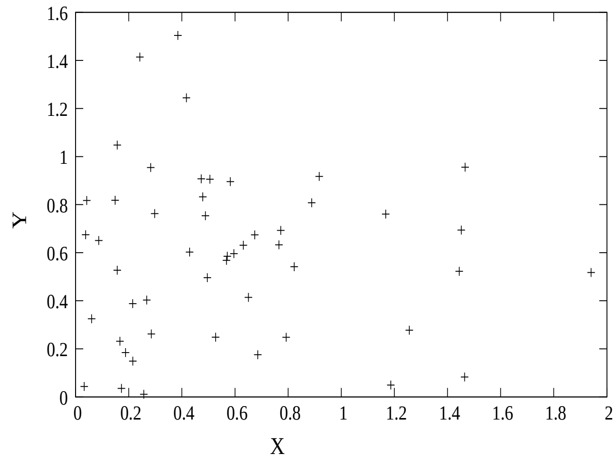


Figure 5: The scatter of (X, Y) for simulated sample.

We have used the above initial estimates to compute the MLEs of the unknown parameters. We have reported the MLEs and the associated 95% confidence intervals (in brackets) based on the observed Fisher information matrix.

$$\widehat{\alpha} = 1.368(\mp 0.352), \quad \widehat{\theta} = 1.833(\mp 0.451), \quad \widehat{\delta}_1 = 0.864(0.235), \quad \widehat{\delta}_2 = 0.969(\mp 0.268),$$

$$\widehat{\delta}_3 = 0.877(\mp 0.235), \quad \widehat{\delta}_4 = 0.956(0.254), \quad \widehat{p} = 0.88(0.17).$$

5.2 Soccer Data

In this section we have analyzed on soccer data set based on the proposed BWE model. The data set has been obtained from Meintanis [23] and it represents the soccer data where at least one goal has been scored by the home team and at least one goal has been scored directly either from a penalty kick, foul kick or any other direct kick. All of them together we call tham as the *kick goal*. Here X reresents the time in minutes of the first *kick goal* by any team and Y represents the time in minutes of the first goal of any type scored by the home team. It may be noted that in this case all possibilities are open, namely $X < Y$, $X > Y$ or $X = Y$. The data set has been presented in Table 1.

Table 1: UEFA Champion's League data

2005-2006	X	Y	2004-2005	X	Y
Lyon-Real Madrid	26	20	Internazionale-Bremen	34	34
Milan-Fenerbahce	63	18	Real Madrid-Roma	53	39
Chelsea-Anderlecht	19	19	Man. United-Fenerbahce	54	7
Club Brugge-Juventus	66	85	Bayern-Ajax	51	28
Fenerbahce-PSV	40	40	Moscow-PSG	76	64
Internazionale-Rangers	49	49	Barcelona-Shakhtar	64	15
Panathinaikos-Bremen	8	8	Leverkusen-Roma	26	48
Ajax-Arsenal	69	71	Arsenal-Panathinaikos	16	16
Man. United-Benfica	39	39	Dynamo Kyiv-Real Madrid	44	13
Real Madrid-Rosenborg	82	48	Man. United-Sparta	25	14
Villarreal-Benfica	72	72	Bayern-M. TelAviv	55	11
Juventus-Bayern	66	62	Bremen-Internazionale	49	49
Club Brugge-Rapid	25	9	Anderlecht-Valencia	24	24
Olympiacos-Lyon	41	3	Panathinaikos-PSV	44	30
Internazionale-Porto	16	75	Arsenal-Rosenborg	42	3
Schalke-PSV	18	18	Liverpool-Olympiacos	27	47
Barcelona-Bremen	22	14	M. Tel-Aviv-Juventus	28	28
Milan-Schalke	42	42	Bremen-Panathinaikos	2	2
Rapid-Juventus	36	52			

Before procedding further, we have divided all the data points by 100 mainly for numerical purposes. It is not going to make any difference in the inference procedure. It is observed that in this case $N_1 = 6$, $N_2 = 17$ and $N_0 = 14$. We obtain the initial estimates of α, θ, δ_1 and δ_4 as

$$\widetilde{\alpha} = 1.477, \quad \widetilde{\theta} = 5.341, \quad \widetilde{\delta}_1 = 4.199, \quad \widetilde{\delta}_4 = 4.699.$$

Based on these initial estimates we compute the MLEs of the unknown parameters. The MLEs and the associated 95% confidence intervals (in brackets) based on the observed Fisher information matrix are reported below.

$$\widehat{\alpha} = 1.649(\mp 0.473), \quad \widehat{\theta} = 5.981(\mp 1.114), \quad \widehat{\delta}_1 = 0.483(0.097), \quad \widehat{\delta}_2 = 5.499(\mp 1.654),$$

$$\widehat{\delta}_3 = 4.653(\mp 0.998), \quad \widehat{\delta}_4 = 1.323(0.189), \quad \widehat{p} = 0.622(0.201).$$

6 Conclusions

In this paper we have proposed a general construncation of a bivariate distribution with a singular component. It is well known that most of the existing methods which can be used to generate a bivariate distribution with a singular component can be classifed into two classes, namely the minimization approach proposed by Marshall and Olkin [22] and the maximization approach proposed by Kundu and Gupta [14]. The present method actually unifies both the approaches. All the existing bivariate distributions with a singular component can be obtained as special cases of the proposed model. We have considered one specific model based on the Weibull distributions and discussed some of its properties, and showed how it can be used in practice. It may be mentioned that this model can be used quite effectively as a dependent competing risks model and it can be extended to the multivariate case also. More work is needed along that direction.

Acknowledgements:

The author would like to thank the reviewers for their helpful comments which have helped to imporve the manuscript.

Appendix A

In this section we show how we can maximize $l_0(\alpha, \mathbf{\Gamma})$ with respect to the unknown parameters. This is a four dimensional optimization problem. To avoid that we use the profile likelihood method. In that case for a fixed α first we maximize with respect to $\mathbf{\Gamma}$ and then we maximize with respect to α. Now maximizing $l_0(\alpha, \mathbf{\Gamma})$ with respect to $\mathbf{\Gamma}$ for a fixed α is a three dimensional optimization problem, and we use the method of Song, Fan, and

Kalbfleisch [30] to optimize it. It can be described as follows. Let us write

$$l_0(\alpha, \mathbf{\Gamma}) = l_1(\mathbf{\Gamma}) + l_2(\mathbf{\Gamma}) + l_3(\alpha),$$

here

$$\begin{aligned}
l_3(\alpha) &= \ln\alpha(2n_1 + 2n_2 + n_0) + \alpha\left(\sum_{i\in I_1\cup I_2}\ln x_i + \sum_{i\in I_1\cup I_2}\ln y_i + \sum_{i\in I_0}\ln u_i\right), \\
l_{1\alpha}(\mathbf{\Gamma}) &= n\ln\theta - \theta\left(\sum_{i\in I_1} x_i^\alpha + \sum_{i\in I_2} y_i^\alpha + \sum_{i\in I_0} u_i^\alpha\right) + n_1\ln\delta_2 - \delta_2\sum_{i\in I_1}(y_i^\alpha - x_i^\alpha) \\
&\quad + n_2\ln\delta_3 - \delta_3\sum_{i\in I_2}(x_i^\alpha - y_i^\alpha) \\
l_{2\alpha}(\mathbf{\Gamma}) &= -(n_1+n_2)\ln(2\theta - \delta_2 - \delta_3) + n_1\ln(\theta - \delta_2) + n_2\ln(\theta - \delta_3).
\end{aligned}$$

Therefore, for a given α, if $\widehat{\mathbf{\Gamma}}_\alpha$ maximizes $l_1(\mathbf{\Gamma}) + l_2(\mathbf{\Gamma})$, then the MLE of α, say $\widehat{\alpha}$ can be obtained by the argument maximum of $l_0(\alpha, \widehat{\mathbf{\Gamma}}_\alpha)$, and the MLE of $\mathbf{\Gamma}$, say $\widehat{\mathbf{\Gamma}}$ can be obtained as $\widehat{\mathbf{\Gamma}} = \widehat{\mathbf{\Gamma}}_{\widehat{\alpha}}$

For a given α, we want to maximize $l_{1\alpha}(\mathbf{\Gamma}) + l_{2\alpha}(\mathbf{\Gamma})$, with respect to $\mathbf{\Gamma}$. It means we need to find the solution of the vector equation:

$$\dot{l}_{1\alpha}(\mathbf{\Gamma}) + \dot{l}_{2\alpha}(\mathbf{\Gamma}) = \mathbf{0},$$

equivalently

$$\dot{l}_{1\alpha}(\mathbf{\Gamma}) = -\dot{l}_{2\alpha}(\mathbf{\Gamma}),$$

here $\mathbf{0} = (0,0,0)^\top$,

$$\dot{l}_{1\alpha}(\mathbf{\Gamma}) = \left(\frac{\partial}{\partial\theta}l_{1\alpha}(\mathbf{\Gamma}), \frac{\partial}{\partial\delta_2}l_{1\alpha}(\mathbf{\Gamma}), \frac{\partial}{\partial\delta_3}l_{1\alpha}(\mathbf{\Gamma})\right)^\top, \dot{l}_{2\alpha}(\mathbf{\Gamma}) = \left(\frac{\partial}{\partial\theta}l_{2\alpha}(\mathbf{\Gamma}), \frac{\partial}{\partial\delta_2}l_{2\alpha}(\mathbf{\Gamma}), \frac{\partial}{\partial\delta_3}l_{2\alpha}(\mathbf{\Gamma})\right)^\top.$$

Song, Fan, and Kalbfleisch [30] suggested the following, first solve

$$\dot{l}_{1\alpha}(\mathbf{\Gamma}) = 0, \tag{24}$$

if $\mathbf{\Gamma}^{(0)}$ is the solution of (24), then find $\mathbf{\Gamma}^{(1)}$ such that

$$\dot{l}_{1\alpha}(\mathbf{\Gamma}) = -\dot{l}_{2\alpha}(\mathbf{\Gamma}^{(0)}).$$

Continue the process, until the convergence takes place. We will provide the explicit expressions of If $\mathbf{\Gamma}^{(j)} = (\theta^{(j)}, \delta_2^{(j)}, \delta_3^{(j)})^\top$, the value of $\mathbf{\Gamma}^{(j)}$ at the j-th iteration. Let us use the following notations:

$$A(\alpha) = \sum_{i \in I_1} x_i^\alpha + \sum_{i \in I_2} y_i^\alpha + \sum_{i \in I_0} u_i^\alpha, \quad B(\alpha) = \sum_{i \in I_1} (y_i^\alpha - x_i^\alpha), \quad C(\alpha) = \sum_{i \in I_2} (x_i^\alpha - y_i^\alpha),$$

$a_0 = b_0 = c_0 = 0$, and for $j = 1, 2, \ldots,$

$$\begin{aligned} a_j &= \frac{n_1 + n_2}{2\theta^{(j)} - \delta_2^{(j)} - \delta_3^{(j)}} - \frac{n_1}{\theta^{(j)} - \delta_2^{(j)}} - \frac{n_2}{\theta^{(j)} - \delta_3^{(j)}} \\ b_j &= -\frac{n_1 + n_2}{2\theta^{(j)} - \delta_2^{(j)} - \delta_3^{(j)}} + \frac{n_1}{\theta^{(j)} - \delta_2^{(j)}} \\ c_j &= -\frac{n_1 + n_2}{2\theta^{(j)} - \delta_2^{(j)} - \delta_3^{(j)}} + \frac{n_2}{\theta^{(j)} - \delta_3^{(j)}}. \end{aligned}$$

Then

$$\theta^{(j+1)} = \frac{n}{A(\alpha) + a_j}, \quad \delta_2^{(j+1)} = \frac{n_1}{B(\alpha) + b_j}, \delta_3^{(j+1)} = \frac{n_2}{C(\alpha) + c_j}.$$

References

[1] Al-Mutairi, D.K., Ghitany, M.E. and Kundu, D. (2011), "A new bivariate distribution with weightes exponential marginals and its multivariate generalization", *Statistical Papers* vol. 52, 921 – 936.

[2] Al-Mutairi, D.K., Ghitany, M.E. and Kundu, D. (2018), "Weighted Weibull distribution: bivariate and multivariate cases", *Brazilian Journal of Probability and Statistics*, vol. 32, no. 1, 20 – 43, 2018

[3] Azzalini, A. (1985), "A class of distribution which includes the normal ones", *Scandinavian Journal of Statistics*, vol. 12, 171 - 178.

[4] Azzalini, A.A. and Dalla Valle, A. (1996), "The multivariate skew normal distribution", *Biometrika*, vol. 83, 715 - 726.

[5] Balakrishnan, N. and Lai, C. (2009),Continuous bivariate distributions, 2nd edition, Springer, New York.

[6] Barreto-Souza, W. and Lemonte, A.J. (2013), "Bivariate Kumaraswamy distribution: properties and a new method to generate bivariate classes", *Statistics*, vol.47, 1321 – 1342

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

[8] Block, H.W. and Basu, A.P. (1974), "A continuous bivariate exponential extension", *Journal of the American Statistical Association*, vol. 69, 1031¡V1037.

[9] Franco, M., Balakrishnan, N., Kundu, D. and Vivo, J-M (2014), "Generalized mixture of Weibull distributions", *TEST*, vol. 23, 515–535.

[10] Franco, M., Vivo, J-M and Kundu, D. (2020), "A generator of bivariate distributions: Properties, estimation, and applications", *Mathematics*, vol. 8, 1776; https://doi.org/10.3390/math8101776.

[11] Gupta, R.D. and Kundu, D. (2009), "A new class of weighted exponential distribution", *Statistics*, vol. 43, 621 - 643.

[12] Kundu, D., Balakrishnan, N. and Jamalizadeh, A. (2010), "Bivariate Birnbaum-Saunders distribution and associated inference", *Journal of Multivariate Analysis*, vol. 101, 113 – 125.

[13] Kundu, D. and Dey, A. K. (2009), "Estimating the parameters of the Marshall-Olkin bivariate Weibull distribution by EM Algorithm", *Computational Statistics and Data Analysis*, vol. 53, 956 – 965.

[14] Kundu, D. and Gupta, R.D. (2009), "Bivariate generalized exponential distribution", *Journal of Multivariate Analysis*, vol. 100, 581 – 593.

[15] Kundu, D. and Gupta, R.D. (2010), "Modified Sarhan-Balakrishnan singular bivariate distribution", *Journal of Statistical Planning and Inference*, vol. 140, 526 – 538.

[16] Kundu, D. and Gupta, R.D. (2010), "A class of bivariate models with proportional reversed hazard marginals", *Sankhya, Ser. B*, vol. 72, 236 – 253.

[17] Kundu, D. and Gupta, R.D. (2011), "Absolute continuous bivariate generalized exponential distribution", *Advances in Statistical Analysis*, vol. 95, 169 – 185.

[18] Kundu, D. and Gupta, A.K. (2013), "Bayes estimation for the Marshall-Olkin bivariate Weibull distribution", *Computational Statistics and Data Analysis*, vol.57, 271 – 281.

[19] Kundu, D. and Gupta, A.K. (2014), "On bivariate Weibull Geometric distribution", *Journal of Multivariate Analysis*, vol. 123, no. 1, 19 - 29, 2014.

[20] Lu, Jye-Chyi (1989), "Weibull extension of the Freund and Marshall-Olkin bivariate exponential model", *IEEE Transactions on Reliability*, vol. 38, 615 – 619.

[21] Lu, Jye-Chyi (1992), "Bayes parameter estimation for the bivariate Weibull model of Marshall-Olkin for censored data", *IEEE Transactions on Reliability*, vol. 41, 608 – 615.

[22] Marshall, A.W. and Olkin, I. (1967), "A multivariate exponential distribution", *Journal of the American Statistical Association*, vol. 62, 30 – 44.

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

[24] Muhammed, H. (2016), "Bivariate inverse Weibull distribution", *Journal of Statistical Computation and Simulation*, vol. 86, 2335 – 2345.

[25] Nadarajah, S. (2011), "The exponentiated exponential distribution; a survey", *Advances in Statistical Analysis*, vol. 95, 219 - 251.

[26] Pradhan, B and Kundu, D. (2016), "Bayes estimation for the Block and Basu bivariate and multivariate Weibull distributions", *Journal of Statistical Computation and Simulation*, vol. 86, 170 – 182.

[27] Sarhan, A.M. and Balakrishnan, N. (2007), "A new class of bivariate distribution-and its mixture", *Journal of Multivariate Analysis*, vol. 98, 1508 – 1527.

[28] Sarhan, A.M., Hamilton, D.C., Smith, B. and Kundu, D. (2011), "The bivariate generalized linear failure rate distribution and its multivariate extension", *Computational Statistics and Data Analysis*, vol. 55, 644 – 654.

[29] Shahbaz, S., Shahbaz, M.Q. and Butt, N.S. (2010), "A class of weighted Weibull distri- bution", *Pakistan Journal of Statistics and Operation Research*, vol. VI, 53 - 59.

[30] Song, P.X.,Fan, Y. and Kalbfleisch, J.D.(2005), Maximization by parts in likelihood inference (with discussions)", *Journal of the American Statistical Association*, vol. 100, 1145 – 1167.