On Bivariate Weibull Frailty Model

Debasis Kundu *

Abstract

In this paper we have considered bivariate Weibull frailty (BWF) model. It has a singularity along the line $x = y$. There is a positive probability that the two marginals can be equal similar to the Marshall-Olkin bivariate exponential or Marshall-Olkin bivariate Weibull models. The Marshall-Olkin bivariate exponential or Weibull distribution can be obtained as a limiting case of the proposed model. It is a very flexible model. The joint probability density function can take variety of shapes. Different properties of the BWF model have been studied. Different dependency measures have been investigated. The model has five parameters. Computing the maximum likelihood estimators of the unknown parameters involves solving a five dimensional optimization problem. An effective EM algorithm has been proposed and it can be implemented quite conveniently in practice. Extensive simulations have been performed to show the effectiveness of the proposed method. One diabetic retinopathy data set has been analyzed. We have further proposed to analyze dependent competing risks data, and one competing risks data set has been analyzed. The results are quite satisfactory.

Key Words and Phrases: Marshall-Olkin bivariate exponential distribution; Marshall-Olkin bivariate Weibull distribution; bivariate singular distribution; bivariate copula; positive dependence, maximum likelihood estimators; EM algorithm.

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

The motivation of this work came when we were trying to analyze a diabetic retinopathy data set. The diabetic retinopathy is a medical condition of the eyes. This particular eye conditions may depend on various factors of the individual namely age, sex, blood sugar level, cholesterol level etc. One major issue of this disease is that unless somebody goes for a regular eye check-up this may not be detected at the early stage. The final outcome of this disease is blindness. Till today we do not have any treatment available so that this disease can be cured. The available treatment can only delay the onset of blindness. Due to this reason an extensive amount of work is going on to develop new treatment so that the onset of the blindness can be delayed. One such treatment which has been recently introduce is the laser treatment. Different experiments have been conducted to test whether the laser treatment has any significant effect in delaying the onset of blindness or not, compared to the traditional treatment.

We will be discussing two such data sets. One data set is of the form (X, Y). Here X and Y denote the time to blindness of the laser treated eye and the other eye, respectively, of the same individual. Clearly, here both $X > 0$ and $Y > 0$. The other data set is of the form (T, Δ), here $T = \min\{X, Y\}$, where X and Y are same as defined above, and $\Delta = 1$, $\Delta = 2$ or $\Delta = 3$, if $T = X$, $T = Y$ or $T = X = Y$, respectively. One important feature of these data sets is that there is a significant portion of the data points where $X = Y$, hence it cannot be ignored. Due to this reason, several authors have analyzed these data sets based on the assumptions that (X, Y) follows Marshall-Olkin bivariate Weibull (MOBW) distribution, see for example Feizjavadian and Hashemi [4], Cai, Shi and Liu [1], Shen and Xu [16], Kundu [8] and Samanta and Kundu [15].

In the first data set few covariates (age, sex and blood sugar level) are available, where

as in the second data set no covariates is available. It is quite possible that many other variables also influence these survival times. Such factors are usually unknown and thus cannot be explicitly included in the analysis. Vaupel et al. [17] suggested a mathematical model for this. They have introduced a random variable, which is not observed, for each individual to the associated survival function. Since it is not observed it is integrated out. There is some identifiability issue but it can be sorted out.

The main aim of this paper is to introduce the frailty for the MOBW model in a very natural way. We call this new model as the Bivariate Weibull Frailty (BWF) model. It may be mentioned that the MOBW distribution has received a considerable amount of attention in analyzing bivariate data with ties. Extensive work has been done in establishing different properties and developing both the classical and Bayesian inference procedures for MOBW model, see for example the review article by Kundu [9] and the references cited there in. The MOBW model can be obtained as a limiting case of the BWF mode. The MOBW distribution has four parameters, where as the proposed BWF model has five parameters. The BWF model is a very flexible model and it also has a singularity along the line $x = y$ similar to the MOBW model. Hence, this model also can be used quite effectively if there are ties in the data set. The joint probability density function (JPDF) can be take variety of shapes, and we have derived different properties of the BWF model. Different dependency properties and dependency measures also have been established. The maximum likelihood estimators of the unknown parameters cannot be obtained in explicit forms. It involves solving five dimensional optimization process. Moreover, finding the five dimensional initial guesses also is not a trivial issue. To avoid that we have proposed to use EM algorithm, which can be implemented quite conveniently in practice. Extensive simulations have been performed to show the effectiveness of the proposed method and one bivariate diabetic retinopathy data set has been analyzed based on the proposed model. We have further proposed a competing risks model based on the BWF model, and one diabetic retinopathy competing risks data

has been analyzed based on this model.

The rest of the paper is organized as follows. In Section 2 we have defined the BWF model and provided several of its properties. The maximum likelihood estimators based on the EM algorithm has been proposed in Section 3. The analysis of a bivariate diabetic retinopathy data set has been provided in Section 4. In Section 5 we have indicated how the proposed BWF model can be used effectively to analyze competing risks data and the analysis of a data set has also been presented. Finally we have concluded the paper in Section 6.

2 Bivariate Weibull Frailty Model

2.1 Notations and Preliminaries

We will use the following notations for the rest of the paper. The two-parameter Weibull distribution with the shape parameter $\alpha > 0$ and $\lambda > 0$ has the following probability density function (PDF);

$$f_{WE}(x;\alpha,\lambda) = \alpha\lambda x^{\alpha-1}e^{-\lambda x^{\alpha}}; \quad \text{for} \quad x > 0, \tag{1}$$

and zero, otherwise. The corresponding cumulative distribution function (CDF), survival function (SF) and the hazard function (HF) will be denoted by $F_{WE}(x;\alpha,\lambda)$, $S_{WE}(x;\alpha,\lambda)$ and $h_{WE}(x;\alpha,\lambda)$, respectively, and for $x > 0$, they are as follows:

$$F_{WE}(x;\alpha,\lambda) = 1 - e^{-\lambda x^{\alpha}}, \quad S_{WE}(x;\alpha,\lambda) = e^{-\lambda x^{\alpha}}, \quad h_{WE}(x;\alpha,\lambda) = \alpha\lambda x^{\alpha-1}.$$

From now on a Weibull distribution with the shape parameter α and scale parameter λ will be denoted by $\text{WE}(\alpha,\lambda)$. The two-parameter gamma distribution with the shape parameter $\alpha > 0$ and $\lambda > 0$ has the following probability density function (PDF);

$$f_{GA}(x;\alpha,\lambda) = \frac{\lambda^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}; \quad \text{for} \quad x > 0, \tag{2}$$

and zero, otherwise. It will be denoted by $\text{GA}(\alpha,\lambda)$.

Let $U_1 \sim$ (follows) 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 $X = \min\{U_1, U_3\}$, $Y = \min\{U_2, U_3\}$ follows Marshall-Olkin bivariate Weibull (MOBW) distribution with parameters $\alpha, \lambda_1, \lambda_2, \lambda_3$. From now on it will be denoted by MOBW$(\alpha, \lambda_1, \lambda_2, \lambda_3)$. The joint PDF of (X, Y) can be written as

$$f_{MOBW}(x,y) = \begin{cases} f_{WE}(x;\alpha,\lambda_1) f_{WE}(y;\alpha,\lambda_2+\lambda_3) & \text{if} \quad x<y \\ f_{WE}(x;\alpha,\lambda_1+\lambda_3) f_{WE}(y;\alpha,\lambda_2) & \text{if} \quad x>y \\ \frac{\lambda_3}{\lambda_1+\lambda_2+\lambda_3} f_{WE}(z;\alpha,\lambda_1+\lambda_2+\lambda_3) & \text{if} \quad x=y=z. \end{cases} \tag{3}$$

2.2 Model Descriptions

Definition: Suppose $U_1 \sim$ WE(α, λ_1), $U_2 \sim$ WE(α, λ_2), $U_3 \sim$ WE(α, λ_3), $V \sim$ GA(β, β) and they are all independently distributed. Let us define

$$X = \min\left\{\frac{U_1}{V^{1/\alpha}}, \frac{U_3}{V^{1/\alpha}}\right\} \quad \text{and} \quad Y = \min\left\{\frac{U_2}{V^{1/\alpha}}, \frac{U_3}{V^{1/\alpha}}\right\}. \tag{4}$$

Then (X, Y) is said to have BWF distribution with parameters $\alpha, \lambda_1, \lambda_2, \lambda_3, \beta$. It will be denoted by BWF$(\alpha, \lambda_1, \lambda_2, \lambda_3, \beta)$.

The joint SF of (X, Y) for $x > 0$ and $y > 0$ becomes

$$\begin{aligned} S_{BWF}(x,y) &= P(X>x, Y>y) \\ &= \int_0^\infty P(U_1 > v^{1/\alpha}x, U_2 > v^{1/\alpha}y, U_3 > v^{1/\alpha}\max\{x,y\}) f_{GA}(v;\beta,\beta)dv \\ &= \int_0^\infty P(U_1 > v^{1/\alpha}x, U_2 > v^{1/\alpha}y, U_3 > v^{1/\alpha}\max\{x,y\}) f_{GA}(v;\beta,\beta)dv \\ &= \int_0^\infty e^{-v\lambda_1 x^\alpha} e^{-v\lambda_2 y^\alpha} e^{-v\lambda_3 \max\{x^\alpha, y^\alpha\}} f_{GA}(v;\beta,\beta)dv \\ &= \frac{\beta^\beta}{\Gamma(\beta)} \int_0^\infty v^{\beta-1} e^{-v(\beta+\lambda_1 x^\alpha+\lambda_2 y^\alpha+\lambda_3 \max\{x^\alpha,y^\alpha\})} dv \\ &= \left[1 + \frac{\lambda_1}{\beta}x^\alpha + \frac{\lambda_2}{\beta}y^\alpha + \frac{\lambda_3}{\beta}\max\{x^\alpha, y^\alpha\}\right]^{-\beta} \\ &= \begin{cases} [1+\theta_1 x^\alpha + (\theta_2+\theta_3)y^\alpha]^{-\beta} & \text{if} \quad x<y \\ [1+(\theta_1+\theta_3)x^\alpha + \theta_2 y^\alpha]^{-\beta} & \text{if} \quad y \le x. \end{cases} \end{aligned}$$

Here we have denoted $\theta_1 = \lambda_1/\beta$, $\theta_2 = \lambda_2/\beta$ and $\theta_3 = \lambda_3/\beta$. Hence, the marginal SFs of X and Y become

$$P(X > x) = [1 + (\theta_1 + \theta_3)x^{\alpha}]^{-\beta} \quad \text{and} \quad P(Y > y) = [1 + (\theta_2 + \theta_3)y^{\alpha}]^{-\beta}.$$

The PDFs of X and Y for $x > 0$ and $y > 0$ become

$$f_X(x) = \frac{\alpha\beta(\theta_1 + \theta_3)x^{\alpha-1}}{[1 + (\theta_1 + \theta_3)x^{\alpha}]^{\beta+1}} \quad \text{and} \quad f_Y(y) = \frac{\alpha\beta(\theta_2 + \theta_3)x^{\alpha-1}}{[1 + (\theta_2 + \theta_3)y^{\alpha}]^{\beta+1}},$$

respectively. We introduce the following notation. A random variable is said to have a univariate Weibull frailty (UWF) distribution with parameters α, β, θ, if it has the following PDF for $x > 0$

$$f_{WF}(x; \alpha, \beta, \theta) = \frac{\alpha\beta\theta x^{\alpha-1}}{[1 + \theta x^{\alpha}]^{\beta+1}}, \tag{5}$$

and zero, otherwise. It will be denoted by UWF(α, β, θ). The generation of random sample from a UWF model is quite simple by using the inverse transformation. Hence, the generation of random sample from a BWF model can be performed very easily.

Following exactly the same procedure as in Kundu and Gupta [12] the joint PDF of (X, Y) if (X, Y) can be obtained from the joint SF of the BWF. Alternatively, it can be obtained as follows. Observe that $\{(X, Y)|V = v\} \sim$ MOBW$(\alpha, \lambda_1 v, \lambda_2 v, \lambda_3 v)$. Hence, using (3) we can write the joint PDF of (X, Y) given $V = v$ as follows;

$$f_{(X,Y)|V=v}(x, y) = \begin{cases} f_{WE}(x; \alpha, \lambda_1 v) f_{WE}(y; \alpha, (\lambda_2 + \lambda_3)v) & \text{if} \quad x < y \\ f_{WE}(x; \alpha, (\lambda_1 + \lambda_3)v) f_{WE}(y; \alpha, \lambda_2 v) & \text{if} \quad y < x \\ \frac{\lambda_3}{\lambda_1+\lambda_2+\lambda_3} f_{WE}(z; \alpha, (\lambda_1 + \lambda_2 + \lambda_3)v) & \text{if} \quad x = y = z. \end{cases} \tag{6}$$

Hence, the joint PDF of (X, Y) becomes

$$f_{BWF}(x, y) = \begin{cases} \frac{\alpha^2\beta(\beta+1)\theta_1\theta_2 x^{\alpha-1}y^{\alpha-1}}{(1+\theta_1 x^{\alpha}+(\theta_2+\theta_3)y^{\alpha})^{\beta+2}} & \text{if} \quad x < y \\ \frac{\alpha^2\beta(\beta+1)\theta_1\theta_2 x^{\alpha-1}y^{\alpha-1}}{(1+(\theta_1+\theta_3)x^{\alpha}+\theta_2 y^{\alpha})^{\beta+2}} & \text{if} \quad y < x \\ \frac{\alpha\theta_3\beta z^{\alpha-1}}{(1+(\theta_1+\theta_2+\theta_3)z^{\alpha})^{\beta+1}} & \text{if} \quad x = y = z. \end{cases} \tag{7}$$

The BWF distribution has an absolutely continuous part and a singular part. It is absolutely continuous on $0 < x \neq y < \infty$ and it has a singular part on $0 < x = y < \infty$. The joint PDF

(7) can be written as

$$f_{BWF}(x,y) = \frac{\theta_1+\theta_2}{\theta_1+\theta_2+\theta_3} f_{ac}(x,y) + \frac{\theta_3}{\theta_1+\theta_2+\theta_3} f_{si}(x,y), \tag{8}$$

where

$$f_{ac}(x,y) = \frac{\theta_1+\theta_2+\theta_3}{\theta_1+\theta_2} \begin{cases} \frac{\alpha^2\beta(\beta+1)\theta_1\theta_2 x^{\alpha-1}y^{\alpha-1}}{(1+\theta_1 x^\alpha+(\theta_2+\theta_3)y^\alpha)^{\beta+2}} & \text{if} \quad x<y \\ \frac{\alpha^2\beta(\beta+1)\theta_1\theta_2 x^{\alpha-1}y^{\alpha-1}}{(1+(\theta_1+\theta_3)x^\alpha+\theta_2 y^\alpha)^{\beta+2}} & \text{if} \quad y<x \end{cases}$$

and

$$f_{si}(x,y) = \begin{cases} \frac{\alpha\beta(\theta_1+\theta_2+\theta_3)z^{\alpha-1}}{(1+(\theta_1+\theta_2+\theta_3)z^\alpha)^{\beta+1}} & \text{if} \quad x=y=z \\ 0 & \text{if} \quad x\neq y. \end{cases}$$

It can be easily seen that

$$P(X<Y) = \frac{\theta_1}{\theta_1+\theta_2+\theta_3}, \quad P(Y<X) = \frac{\theta_2}{\theta_1+\theta_2+\theta_3}, \quad P(X=Y) = \frac{\theta_3}{\theta_1+\theta_2+\theta_3},$$

The correlation coefficient of X and Y varies from zero to one. X and Y are independent if $\theta_3 = 0$ and the correlation tends to one as $\theta_3 \to \infty$.

The following results will be used in the implementation of the EM algorithm and they can be obtained after some calculations.

$$V|\{(X,Y)=(x,y)\} \sim \text{Gamma}(\beta+2, (\beta+\lambda_1 x^\alpha+(\lambda_2+\lambda_3)y^\alpha) \quad \text{if } x<y, \tag{9}$$

$$V|\{(X,Y)=(x,y)\} \sim \text{Gamma}(\beta+2, (\beta+(\lambda_1+\lambda_3)x^\alpha+\lambda_2 y^\alpha) \quad \text{if } x>y, \tag{10}$$

$$V|\{(X,Y)=(x,y)\} \sim \text{Gamma}(\beta+1, (\beta+(\lambda_1+\lambda_2+\lambda_3)z^\alpha) \quad \text{if } x=y=z. \tag{11}$$

Hence,

$$E(V|\{(X,Y)=(x,y)\}) = \frac{\beta+2}{(\beta+\lambda_1 x^\alpha+(\lambda_2+\lambda_3)y^\alpha)} \quad \text{if } x<y, \tag{12}$$

$$E(V|\{(X,Y)=(x,y)\}) = \frac{\beta+2}{(\beta+(\lambda_1+\lambda_3)x^\alpha+\lambda_2 y^\alpha)} \quad \text{if } x>y, \tag{13}$$

$$E(V|\{(X,Y)=(x,y)\}) = \frac{\beta+1}{(\beta+(\lambda_1+\lambda_2+\lambda_3)z^\alpha)} \quad \text{if } x=y=z \tag{14}$$

and

$$E(\ln V|\{(X,Y)=(x,y)\}) = \psi(\beta+2) - \psi(\beta+\lambda_1 x^\alpha+(\lambda_2+\lambda_3)y^\alpha)) \quad \text{if } x<y, \tag{15}$$

$$E(\ln V|\{(X,Y)=(x,y)\}) = \psi(\beta+2)-\psi(\beta+(\lambda_1+\lambda_3)x^\alpha+\lambda_2 y^\alpha)) \quad \text{if } x>y, \quad (16)$$

$$E(\ln V|\{(X,Y)=(x,y)\}) = \psi(\beta+1)-\psi(\beta+(\lambda_1+\lambda_2+\lambda_3)z^\alpha)) \quad \text{if } x=y=z. \quad (17)$$

Here $\psi(\cdot)$ is the digamma function, i.e. $\psi(x)=\dfrac{d}{dx}\Gamma(x)=\dfrac{\Gamma'(x)}{\Gamma(x)}$.

2.3 Properties

In this section we provide some properties of the UWF and BWF models. If $X \sim \text{UWF}(\alpha,\beta,\theta)$, then it can be easily seen that the PDF of X is a decreasing function for $0<\alpha\leq 1$ and it is an unimodal function for all values of $\beta>0$ and $\theta>0$. The hazard function of X becomes

$$h_{WF}(x)=\frac{\alpha\beta x^{\alpha-1}}{1+\theta x^\alpha}; \quad x>0. \quad (18)$$

It can be easily shown that if $0<\alpha\leq 1$, the hazard function is a decreasing function and for $\alpha>1$, the hazard function is an unimodal function. It is clear that the shape (whether it will be decreasing or unimodal) of the PDF or HF does not depend on β and θ, it depends only on α. In Figure 1 we provide the plot of the PDF and HF for different values of α.

The following results are useful for data analysis purposes or it may have some independent interests also.

Theorem 1: Suppose $(X,Y)\sim \text{BWF}(\alpha,\beta,\theta_1,\theta_2,\theta_3)$. Then we have the following results:

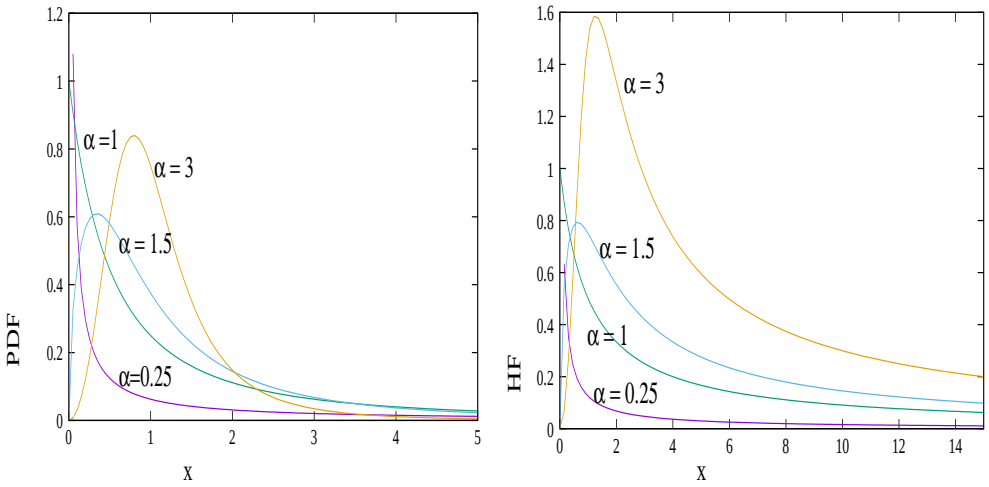
(a) $X\sim \text{UWF}(\alpha,\beta,(\theta_1+\theta_3))$.

(b) $Y\sim \text{UWF}(\alpha,\beta,(\theta_2+\theta_3))$.

(c) $\min\{X,Y\}\sim \text{UWF}(\alpha,\beta,(\theta_1+\theta_2+\theta_3))$

Proof: The proof can be easily obtained from the joint survival function. ∎

The following results provide the shapes of the joint PDF of the absolute continuous part of the BWF.



(a) (b)

Figure 1: The PDF and HF of WF(α, β, θ) for different values of α, when $\beta = \theta = 1$. (a) PDF and (b) HF.

Theorem 2: Suppose $(X, Y) \sim$ BWF$(\alpha, \beta, \theta_1, \theta_2, \theta_3)$. Then we have the following results:

(a) The joint PDF of the absolute continuous part of (X, Y) is continuous everywhere if $\theta_1 = \theta_2$.

(b) If $\theta_1 \neq \theta_2$, the joint PDF of the absolute continuous part of (X, Y) is continuous everywhere except on the line $x = y$.

(c) If $0 < \alpha \leq 1$, the joint PDF of the absolute continuous part of (X, Y) is a decreasing function for all values of $\beta > 0$, $\theta_1 > 0$, $\theta_2 > 0$ and $\theta_3 > 0$.

(d) If $\theta_1 = \theta_2 = \theta$ and $\alpha > 1$ the joint PDF of the absolute continuous part of (X, Y) has a unique mode at $\left(\left(\frac{\alpha - 1}{(2\theta + \theta_3)(\alpha\beta + 1)} \right)^{1/\alpha}, \left(\frac{\alpha - 1}{(2\theta + \theta_3)(\alpha\beta + 1)} \right)^{1/\alpha} \right)$.

(e) If $\theta_1 > \theta_2 + \theta_3$ and $\alpha > 1$ the joint PDF of the absolute continuous part of (X, Y) has a unique mode at $\left(\left(\frac{\alpha - 1}{\theta_1(\alpha\beta + 2)} \right)^{1/\alpha}, \left(\frac{\alpha - 1}{(\theta_2 + \theta_3)(\alpha\beta + 2)} \right)^{1/\alpha} \right)$.

(f) $\theta_2 > \theta_1 + \theta_3$ and $\alpha > 1$ the joint PDF of the absolute continuous part of (X, Y) has a unique mode at $\left(\left(\frac{\alpha-1}{(\theta_1+\theta_3)(\alpha\beta+2)}\right)^{1/\alpha}, \left(\frac{\alpha-1}{\theta_2(\alpha\beta+2)}\right)^{1/\alpha}\right)$.

(g) If $\theta_2 < \theta_1 < \theta_2 + \theta_3$ or $\theta_1 < \theta_2 < \theta_1 + \theta_3$, the joint PDF of the absolute continuous part of (X, Y) does not have any mode on $0 < x \neq y < \infty$.

Proof: The Proof of Theorem 2 is not very difficult, hence it is avoided. ∎

In Figure 2 we provide the contour plot of the joint PDF of the absolute continuous part of BWF for different parameter values. It shows that when $\alpha > 1$, the joint PDF is an unimodal function Figures 2 (a),(c),(d), if $\alpha \leq 1$, the joint PDF is an decreasing function Figure 2 (b), if $\theta_1 = \theta_2$, the mode of the joint PDF is on $x = y$, Figure 2 (a), if $\theta_1 > \theta_2 + \theta_3$, the mode of the joint PDF is on $x < y$, Figure 2 (c) and if $\theta_2 > \theta_1 + \theta_3$, the mode of the joint PDF is on $x > y$, Figure 2 (d).

The hazard gradient of BWF model according to Johnson and Kotz [6] can be defined as follows;

$$h_1(x, y) = -\frac{\partial}{\partial x} \ln S_{X,Y}(x, y) \qquad \text{and} \qquad h_2(x, y) = -\frac{\partial}{\partial y} \ln S_{X,Y}(x, y).$$

Hence, the hazard gradients of BWF are as follows':

$$h_1(x, y) = -\frac{\partial}{\partial x} \ln S_{X,Y}(x, y) = \begin{cases} \frac{\beta\alpha\theta_1 x^{\alpha-1}}{1+\theta_1 x^\alpha + (\theta_2+\theta_3) y^\alpha} & \text{if} \quad x < y \\ \frac{\beta\alpha(\theta_1+\theta_3) x^{\alpha-1}}{1+(\theta_1+\theta_3) x^\alpha + \theta_2 y^\alpha} & \text{if} \quad x > y \end{cases}$$

$$h_2(x, y) = -\frac{\partial}{\partial y} \ln S_{X,Y}(x, y) = \begin{cases} \frac{\beta\alpha(\theta_2+\theta_3) y^{\alpha-1}}{1+\theta_1 x^\alpha + (\theta_2+\theta_3) y^\alpha} & \text{if} \quad x < y \\ \frac{\beta\alpha\theta_2 y^{\alpha-1}}{1+(\theta_1+\theta_3) x^\alpha + \theta_2 y^\alpha} & \text{if} \quad x > y \end{cases}$$

Now we will show that the survival function of BWF satisfies the total positivity of order two (TP_2) property, and also it satisfies some hazard rate ordering properties. Note that a function $g : \mathbb{R} \longrightarrow \mathbb{R}$ is said to have TP_2 property, if for all $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$, $\mathbf{y} = (y_1, y_2) \in \mathbb{R}^2$, $g(\mathbf{x})g(\mathbf{x}) \leq g(\mathbf{x} \wedge \mathbf{y})g(\mathbf{x} \vee \mathbf{y})$. Here $\mathbf{x} \wedge \mathbf{y} = (\min\{x_1, y_1\}, \min\{x_2, y_2\})$ and $\mathbf{x} \vee \mathbf{y} = (\max\{x_1, y_1\}, \max\{x_2, y_2\})$. Let $\boldsymbol{U}$ and $\boldsymbol{V}$ be two bivariate random vectors with

survival function S_U and S_V, respectively. We say that $\boldsymbol{U}$ is smaller that $\boldsymbol{V}$ in the bivariate hazard rate order (denoted by $\boldsymbol{U} \leq_{hr} \boldsymbol{V}$) if

$$S_U(\boldsymbol{x})S_V(\boldsymbol{y}) \leq S_U(\boldsymbol{x} \wedge \boldsymbol{y})S_V(\boldsymbol{x} \vee \boldsymbol{y}); \quad \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2.$$

We say that $\boldsymbol{U}$ is smaller that $\boldsymbol{V}$ in the bivariate weak hazard rate order (denoted by $\boldsymbol{U} \leq_{whr} \boldsymbol{V}$) if

$$S_U(\boldsymbol{x})S_V(\boldsymbol{y}) \leq S_U(\boldsymbol{x})S_V(\boldsymbol{y}); \quad \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2. \tag{19}$$

We have the following results.

Theorem 3:

(a) If $(X, Y) \sim \text{BWF}(\alpha, \beta, \theta_1, \theta_2, \theta_3)$, then the survival function of (X, Y) has the TP_2 property.

(b) If $\boldsymbol{X} = (X_1, X_2) \sim \text{BWF}(\alpha, \beta, \theta_1, \theta_2, \theta_3)$ and $\boldsymbol{Y} = (Y_1, Y_2) \sim \text{BWF}(\alpha, \beta, \theta_1, \theta_2, \widetilde{\theta_3})$. If $\theta_3 > \widetilde{\theta_3}$, then $\boldsymbol{X} \leq_{whr} \boldsymbol{Y}$.

(c) If $\boldsymbol{X} = (X_1, X_2) \sim \text{BWF}(\alpha, \beta, \theta_1, \theta_2, \theta_3)$ and $\boldsymbol{Y} = (Y_1, Y_2) \sim \text{BWF}(\alpha, \beta, \theta_1, \theta_2, \widetilde{\theta_3})$. If $\theta_3 > \widetilde{\theta_3}$, then $\boldsymbol{X} \leq_{hr} \boldsymbol{Y}$.

Proof:

(a) To prove this, we need to show that for all possible values of $\boldsymbol{x} = (x_1, x_2) \in \mathbb{R}^2$ and $\boldsymbol{y} = (y_1, y_2) \in \mathbb{R}^2$

$$S_{X,Y}(x_1, x_2)S_{X,Y}(y_1, y_2) \leq S_{X,Y}(x_1 \wedge y_1, x_2 \wedge y_2)S_{X,Y}(x_1 \vee y_1, x_2 \vee y_2). \tag{20}$$

Now the above inequality (20) can be shown by considering all possible twenty four cases namely $x_1 < x_2 < y_1 < y_2$, $x_1 < x_2 < y_2 < y_1$, and so on.

(b) To prove this, we need to show (19). It can be shown again by considering all possible twenty four cases as above.

(c) Using (a), (b), and Theorem 2.1 of Hu et al. [5], the result follows.

The BWF has the following survival copula

$$\overline{C}(u,v) = \begin{cases} \left(\frac{\theta_1}{\theta_1+\theta_3}(u^{-1/\beta}-1)+v^{-1/\beta}\right)^{-\beta} & \text{if} \quad (\theta_2+\theta_3)(u^{-1/\beta}-1) < (\theta_1+\theta_3)(v^{-1/\beta}-1) \\ \left(u^{-1/\beta}+\frac{\theta_2}{\theta_2+\theta_3}(v^{-1/\beta}-1)\right)^{-\beta} & \text{if} \quad (\theta_2+\theta_3)(u^{-1/\beta}-1) > (\theta_1+\theta_3)(v^{-1/\beta}-1) \end{cases} \tag{21}$$

If $\theta_1 = \theta_2$ and if we denote $\delta = \dfrac{\theta_1}{\theta_1+\theta_3} = \dfrac{\theta_2}{\theta_2+\theta_3}$, then (21) becomes

$$\overline{C}(u,v) = \begin{cases} \left(\delta(u^{-1/\beta}-1)+v^{-1/\beta}\right)^{-\beta} & \text{if} \quad u > v \\ \left(u^{-1/\beta}+\delta(v^{-1/\beta}-1)\right)^{-\beta} & \text{if} \quad u < v \end{cases} \tag{22}$$

Hence, different dependency properties of BWF may be explored through copula function.

3 Inference

3.1 Modified EM Algorithm

In this section we provide the maximum likelihood estimators of the unknown parameters based on the observations $D = \{(x_i, y_i); i = 1, \ldots, n\}$. Let us denote $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2 \cup \mathcal{D}_3$, where $\mathcal{D}_1 = \{(x_i, y_i) : x_i < y_i\}$,$\mathcal{D}_2 = \{(x_i, y_i) : x_i > y_i\}$, $\mathcal{D}_3 = \{(x_i, y_i) : x_i = y_i = z_i\}$ and the number of elements in the set $\mathcal{D}_j$ is n_j for $j = 1, 2, 3$. The log-likelihood function of the observed data D becomes

$$\begin{aligned} l(\Theta|D) \;=\; & (2n_1+2n_2+n_3)\ln\alpha + n\ln\beta + (n_1+n_2)(\ln(\beta+1)+\ln\theta_1+\ln\theta_2) + n_3\ln\theta_3 + \\ & (\alpha-1)\left(\sum_{i\in D_1\cup D_2}(\ln x_i+\ln y_i)+\sum_{i\in D_3}\ln z_i\right) - (\beta+1)\ln(1+(\theta_1+\theta_2+\theta_3)z_i^{\alpha}) - \\ & (\beta+2)\left[\sum_{i\in D_1}\ln(1+\theta_1 x_i^{\alpha}+(\theta_2+\theta_3)y_i^{\alpha})+\sum_{i\in D_1}\ln(1+(\theta_1+\theta_3)x_i^{\alpha}+\theta_2 y_i^{\alpha})\right], \end{aligned} \tag{23}$$

where $\Theta = (\alpha, \beta, \theta_1, \theta_2, \theta_3)$. It may be mentioned that to avoid introducing another notation, we sometime denote $\Theta = (\alpha, \beta, \lambda_1, \lambda_2, \lambda_3)$ and it should be clear from the context. Hence, the MLEs of the unknown parameters can be obtained by maximizing (23) with respect to unknown parameters. It involves a 5-dimensional optimization problem. To avoid that we treat this as a missing value problem. First observe that $\{(X, Y)|V\}$ has MOBW distribution, and there is a very effective EM algorithm has been proposed by Kundu and Dey [10]. It is assumed that the complete observations are coming from the following random vector $(X, Y, \Delta_1, \Delta_2, V)$. Here V is the frailty random variable, and (Δ_1, Δ_2) are defined as follows:

$$\Delta_1 = \begin{cases} 0 & \text{if} \quad X = \frac{U_3}{V^{1/\alpha}} \\ 1 & \text{if} \quad X = \frac{U_1}{V^{1/\alpha}} \end{cases} \qquad \Delta_2 = \begin{cases} 0 & \text{if} \quad Y = \frac{U_3}{V^{1/\alpha}} \\ 2 & \text{if} \quad Y = \frac{U_2}{V^{1/\alpha}} \end{cases}$$

Here U_1, U_2, U_3 are same as defined before. Note that

$$\begin{aligned} P(\Delta = 1, \Delta_2 = 0|X < Y) &= \frac{\theta_3}{\theta_2 + \theta_3}, \quad P(\Delta = 1, \Delta_2 = 2|X < Y) = \frac{\theta_2}{\theta_2 + \theta_3}, \\ P(\Delta = 0, \Delta_2 = 2|X > Y) &= \frac{\theta_3}{\theta_1 + \theta_3}, \quad P(\Delta = 1, \Delta_2 = 2|X > Y) = \frac{\theta_1}{\theta_1 + \theta_3}. \end{aligned}$$

It can be easily seen that if we have a complete observations $D_c = \{(x_i, y_, \delta_{1i}, \delta_{2i}, v_i); i = 1, \ldots, n\}$, the MLEs of the unknown parameters can be obtained by solving two one dimensional optimization problems. We use the following notations for further development. At the $k - th$ stage of the EM algorithm, the parameter vector the estimate of the parameter vector Θ will be denoted by $\Theta^{(k)} = (\alpha^{(k)}, \beta^{(k)}, \theta_1^{(k)}, \theta_2^{(k)}, \theta_3^{(k)})$ and

$$a_1^{(k)} = \frac{\theta_3^{(k)}}{\theta_2^{(k)} + \theta_3^{(k)}}, \quad a_2 = \frac{\theta_2^{(k)}}{\theta_2^{(k)} + \theta_3^{(k)}}, \quad b_1 = \frac{\theta_3^{(k)}}{\theta_1^{(k)} + \theta_3^{(k)}}, \quad b_2 = \frac{\theta_1^{(k)}}{\theta_1^{(k)} + \theta_3^{(k)}},$$

$$c_{1i}^{(k)} = E(V_i|x_i = X < Y = y_i), \quad c_{2i}^{(k)} = E(V_i|x_i = X > Y = y_i), \quad c_{3i}^{(k)} = E(V_i|X = Y = z_i, \Theta^{(k)})$$

$$d_{1i}^{(k)} = E(\ln V_i|x_i = X < Y = y_i), \quad d_{2i}^{(k)} = E(\ln V_i|x_i = X > Y = y_i), \quad d_{3i}^{(k)} = E(\ln V_i|X = Y = z_i).$$

At the k-th stage of the EM algorithm, the pseudo log-likelihood contribution of (x_i, y_i) with out the constant for different cases are as follow;

Case 1: $x_i < y_1$

$$2\ln\alpha + \ln\lambda_1 + a_1^{(k)}\ln\lambda_3 + a_2^{(k)}\ln\lambda_2 + \alpha(\ln x_i + \ln y_i) - \lambda_1 c_{1i}^{(k)} x_i^{\alpha} - (\lambda_2+\lambda_3)c_{1i}^{(k)} y_i^{\alpha} + \beta\ln\beta - \ln(\Gamma(\beta)) + \beta(d_{1i}^{(k)} - c_{1i}^{(k)})$$

Case 2: $x_i < y_1$

$$2\ln\alpha + \ln\lambda_2 + b_1^{(k)}\ln\lambda_3 + b_2^{(k)}\ln\lambda_1 + \alpha(\ln x_i + \ln y_i) - (\lambda_1+\lambda_3)c_{2i}^{(k)} x_i^{\alpha} - \lambda_2 c_{2i}^{(k)} y_i^{\alpha} + \beta\ln\beta - \ln(\Gamma(\beta)) + \beta(d_{2i}^{(k)} - c_{2i}^{(k)})$$

Case 3: $x_i = y_1 = z_i$

$$\ln\alpha + \ln\lambda_3 + \alpha\ln z_i - (\lambda_1+\lambda_2+\lambda_3)c_{3i}^{(k)} z_i^{\alpha} + \beta\ln\beta - \ln(\Gamma(\beta)) + \beta(d_{3i}^{(k)} - c_{3i}^{(k)}).$$

Hence, the pseudo log-likelihood function at the $(k+1)-th$ stage can be written as

$$\begin{aligned} l_{pseudo}(\Theta|\Theta^{(k)}) &= (n_1 + n_2 b_2^{(k)})\ln\lambda_1 - \lambda_1\left(\sum_{i\in D_1} c_{1i}^{(k)} x_i^{\alpha} + \sum_{i\in D_2} c_{2i}^{(k)} x_i^{\alpha} + \sum_{i\in D_3} c_{3i}^{(k)} z_i^{\alpha}\right) + \\ &\quad (n_2 + n_1 a_2^{(k)})\ln\lambda_2 - \lambda_2\left(\sum_{i\in D_1} c_{1i}^{(k)} y_i^{\alpha} + \sum_{i\in D_2} c_{2i}^{(k)} y_i^{\alpha} + \sum_{i\in D_3} c_{3i}^{(k)} z_i^{\alpha}\right) + \\ &\quad (n_1 a_1^{(k)} + n_2 b_1^{(k)} + n_3)\ln\lambda_3 - \lambda_3\left(\sum_{i\in D_1} c_{1i}^{(k)} y_i^{\alpha} + \sum_{i\in D_2} c_{2i}^{(k)} x_i^{\alpha} + \sum_{i\in D_3} c_{3i}^{(k)} z_i^{\alpha}\right) + \\ &\quad (2n_1 + 2n_2 + n_3)\ln\alpha + \alpha\left(\sum_{i\in D_1\cup D_2}(\ln x_i + \ln y_i) + \sum_{i\in D_3}\ln z_i\right) + \\ &\quad n\beta\ln\beta - n\ln(\Gamma(\beta)) + \beta\left(\sum_{i\in D_1}(d_{1i}^{(k)} - c_{1i}^{(k)}) + \sum_{i\in D_2}(d_{2i}^{(k)} - c_{2i}^{(k)}) + \sum_{i\in D_3}(d_{3i}^{(k)} - c_{3i}^{(k)})\right). \end{aligned} \tag{24}$$

Now $\Theta^{(k+1)}$ can be obtained by maximizing (24) with respect Θ, and they are as follows. If we denote

$$\lambda_1^{(k+1)}(\alpha) = \frac{(n_1 + n_2 b_2^{(k)})}{\sum_{i\in D_1} c_{1i}^{(k)} x_i^{\alpha} + \sum_{i\in D_2} c_{2i}^{(k)} x_i^{\alpha} + \sum_{i\in D_3} c_{3i}^{(k)} z_i^{\alpha}}$$

$$\lambda_2^{(k+1)}(\alpha) = \frac{(n_2 + n_1 a_2^{(k)})}{\sum_{i\in D_1} c_{1i}^{(k)} y_i^{\alpha} + \sum_{i\in D_2} c_{2i}^{(k)} y_i^{\alpha} + \sum_{i\in D_3} c_{3i}^{(k)} z_i^{\alpha}}$$

$$\lambda_3^{(k+1)}(\alpha) = \frac{(n_1 a_1^{(k)} + n_2 b_1^{(k)} + n_3)}{\sum_{i\in D_1} c_{1i}^{(k)} y_i^{\alpha} + \sum_{i\in D_2} c_{2i}^{(k)} x_i^{\alpha} + \sum_{i\in D_3} c_{3i}^{(k)} z_i^{\alpha}},$$

then first obtain $\alpha^{(k+1)}$ by maximizing $g(\alpha)$ with respect to α, where

$$g(\alpha) = (2n_1 + 2n_2 + n_3)\ln\alpha + \alpha\left(\sum_{i\in D_1\cup D_2}(\ln x_i + \ln y_i) + \sum_{i\in D_3}\ln z_i\right) + (n_1 + n_2 b_2^{(k)})\ln\lambda_1^{(k+1)}(\alpha) + (n_2 + n_1 a_2^{(k)})\ln\lambda_2^{(k+1)}(\alpha) + (n_1 a_1^{(k)} + n_2 b_1^{(k)} + n_3)\ln\lambda_3^{(k+1)}(\alpha). \quad (25)$$

Once $\alpha^{(k+1)}$ is obtained, then obtain $\lambda_1^{(k+1)}$, $\lambda_2^{(k+1)}$, $\lambda_3^{(k+1)}$ as $\lambda_1^{(k+1)}(\alpha^{(k+1)})$, $\lambda_2^{(k+1)}(\alpha^{(k+1)})$ and $\lambda_3^{(k+1)}(\alpha^{(k+1)})$, respectively. Obtain $\beta^{(k+1)}$ by maximizing $h(\beta)$ with respect to β, where

$$h(\beta) = n\beta\ln\beta - n\ln(\Gamma(\beta)) + \beta\left(\sum_{i\in D_1}(d_{1i}^{(k)} - c_{1i}^{(k)}) + \sum_{i\in D_2}(d_{2i}^{(k)} - c_{2i}^{(k)}) + \sum_{i\in D_3}(d_{3i}^{(k)} - c_{3i}^{(k)})\right). \quad (26)$$

The following algorithm can be used for that purpose.

Algorithm:

Step 1: Take some initial values of $\Theta = (\alpha, \beta, \lambda_1, \lambda_2, \lambda_3)$, say $\Theta^{(0)} = (\alpha^{(0)}, \beta^{(0)}, \lambda_1^{(0)}, \lambda_2^{(0)}, \lambda_3^{(0)})$.

Step 2: Based on $\Theta^{(0)}$, compute $a_1^{(0)}, a_2^{(0)}, b_1^{(0)}, b_2^{(0)}$, $\{(c_{1i}^{(0)}, c_{2i}^{(0)}, c_{3i}^{(0)}, d_{1i}^{(0)}, d_{2i}^{(0)}, d_{3i}^{(0)}); i = 1, \ldots, n\}$.

Step 3: Obtain $\alpha^{(1)}$ by maximizing (25), obtain $\lambda_1^{(1)}, \lambda_2^{(1)}, \lambda_3^{(1)}$ and obtain $\beta^{(1)}$ by maximizing (26).

Step 4: Go back to Step 1 and replace '0' by '1' and continue the process until convergence takes place.

3.2 Testing of Hypothesis

In this section we want to discuss the following testing of hypothesis problem which has some practical importance. We want to test the following hypothesis

$$H_0 : \lambda_1 = \lambda_2 \quad vs. \quad H_1 : \lambda_1 \neq \lambda_2. \tag{27}$$

It manly tests the equality of the two marginals. To test the hypothesis (27), we propose to use the likelihood ratio test. Hence, we need to maximize the log-likelihood function (23) under H_0. In this case also the modified EM algorithm can be used quite effectively with the necessary changes. If we denote the estimate without restriction of the unknown parameter vector Θ as $\widehat{\Theta}$ and by $\widehat{\Theta}_0$ with restriction, then we reject the null hypothesis if $-2(l_0(\widehat{\Theta}_0|D) - l(\widehat{\Theta}|D))$ is greater than the appropriate upper percentage of χ^2_1 value.

4 Data Analysis

In this section we analyze one bivariate data set to show how the proposed model and the method can be used in practice. The proposed model has been used on a diabetic retinopathy data set. Diabetic retinopathy is a physical disorder of the eye, and it is observed mainly among the diabetic patients. This particular disease leads to blindness. An extensive amount of work has been done in developing different treatment for this disease. Among different methods one recent treatment is the laser treatment. The main aim of this experiment is to test whether the laser treatment has significant different effect compared to the traditional treatment in delaying the onset of blindness. The experiment has been conducted as follows. For each patient one eye has been chosen at random and the laser treatment has been given where in the other eye the traditional treatment has been administered. For each patient the time to onset of blindness of both the eyes have been recorded. Here X and Y denote the times for the laser treated eye and the traditionally treated eye, respectively. The data

set has been presented in Table 1.

Table 1: Bivariate Diabetic Retinopathy Data Set. Here Y_1 denotes the time to blindness of the laser treated eye and Y_2 denotes the same for the other eye.

Sl. No.	X	Y	Sl. No.	X	Y	Sl. No.	X	Y	Sl. No.	X	Y
1.	20.17	6.90	2.	10.27	1.63	3.	5.67	13.83	4.	5.77	1.33
5.	5.90	35.53	6.	25.63	21.90	7.	33.90	14.80	8.	1.73	6.20
9.	30.20	22.00	10.	25.80	13.87	11.	5.73	48.30	12.	9.90	9.90
13.	1.73	1.73	14.	1.77	43.03	15.	8.30	8.30	16.	18.70	6.53
17.	42.17	42.17	18.	14.30	48.43	19.	13.33	9.60	20.	14.27	7.60
21.	34.57	1.80	22.	4.10	12.20	23.	21.57	9.90	24.	13.77	13.77
25.	33.63	33.63	26.	63.33	27.60	27.	38.47	1.63	28.	10.33	0.83
29.	13.83	1.57	30.	11.07	1.97	31.	2.10	11.30	32.	12.93	4.97
33.	24.43	9.87	34.	13.97	30.40	35.	6.30	56.97	36.	13.80	19.00
37.	13.57	5.43	38.	42.77	42.77	39.	42.43	46.63	40.	2.70	2.70

For this data set $n = 38$, and 12, 20 and 6 observations for which $X < Y$, $X > Y$ and $X = Y$, respectively. Based on this data set we want to test whether laser treatment has any significant different effect compared to the traditional treatment or not. Before progressing further we provide some of the basic statistics of the data set. We have provided the median, Q1 (first quartile) and Q3 (third quartile) of X, Y and $\min\{X, Y\}$. We have also fitted UWF models to X, Y and $\min\{X, Y\}$, their fitted parameter values, the Kolmogorv-Smirnov distances (KSD) and the associated p-values have been presented in Table 2. Based on the KSD and the associated p-values, it is clear that UWF model fits X, Y and $\min\{X, Y\}$. We have plotted the empirical survival curves of the time to blindness of both the eyes in Figure 3. We have also plotted the best fitted UWF models to the marginals in Figure 4. They fit loots quite quite well in both the cases. Hence, it is reasonable to try to fit BWF model to the bivariate data set (X, Y).

We would like to compute the MLEs of the fitted BWF model using the EM algorithm as proposed in Section 3. Based on the fitted UWF models to the marginals and the minimum,

Table 2: Some basic statistics of the Bivariate Retinopathy Data Set. The median, Q1 and Q3 of the marginals and their minimum. The estimated parameters of the UWF models, their respective K-S distances and the associated p-values have been presented.

Variable	Median	Q1	Q3	α	β	θ	KSD	p-Value
X	13.82	13.35	25.03	1.4698	5.3205	0.0026	0.1278	0.5640
Y	10.60	5.20	21.95	1.1516	4.5085	0.0099	0.1186	0.6581
$\min\{X, Y\}$	7.25	1.88	13.83	1.1577	4.6788	0.0115	0.1365	0.4787

we have obtained the following initial guesses; $\alpha^{(0)} = 1.2597$, $\beta^{(0)} = 4.8359$, $\lambda_1^{(0)} = 0.0016$, $\lambda_2^{(0)} = 0.0089$ and $\lambda_3^{(0)} = 0.0110$. We have stopped the EM algorithm when the relative difference between to consecutive log-likelihood functions is less than $\epsilon = 10^{-6}$. The iteration stops after 12 steps. The MLEs of the unknown parameters and the associated 95% confidence intervals are also obtained from the las t step of the EM algorithm based on the method of Louis [13]. They are as follows: $\widehat{\alpha} = 1.2077(\mp 0.3765)$, $\widehat{\beta} = 11.4772(\mp 2.6754)$, $\widehat{\lambda}_1 = 0.00220(\mp 00034)$, $\widehat{\lambda}_2 = 0.00245(\mp 0.00037)$, $\widehat{\lambda}_3 = 0.00031(\mp 0.00001)$. The corresponding log-likelihood value is -292.9115.

One natural question is whether the proposed BWF distribution provides a good fit to the above bivariate data. For that purpose we consider the following statistic

$$D = \sup_{x,y} |S_n(x, y) - \widehat{S}_{BWF}(x, y)|.$$

Here $S_n(x, y)$ denotes the empirical survival function, i.e. $S_n(x, y) = \dfrac{\#\{i; x_i \geq x, y_i \geq y\}}{n}$ and $\widehat{S}_{BWF}(x, y)$ denotes the estimated value of $S_{BWF}(x, y)$ based on MLEs. We obtain the $D = 0.1134$ and based on simulation we obtain the associated p value as 0.681. Hence, it shows that the proposed BWF provides a good fit to the bivariate Retinopathy data set.

Now we want to test whether the laser treatment has any significant effect in delaying the blindness or not. It is equivalent in testing (27). We have obtained the MLEs and

the associated 95% confidence intervals of the unknown parameters under the null hypothesis as follows: $\widehat{\alpha}_0 = 1.2041(\mp 0.3668)$, $\widehat{\beta}_0 = 13.2188(\mp 2.5753)$, $\widehat{\lambda}_{10} = 0.00202(\mp 00032)$, $\widehat{\lambda}_{20} = 0.00202(\mp 0.00032)$, $\widehat{\lambda}_{30} = 0.00027(\mp 0.00001)$. The corresponding log-likelihood value is -292.9893. Hence, based on $-2(l_0(\widehat{\Theta}_0|D) - l(\widehat{\Theta}|D))$, we cannot reject the null hypothesis. Therefore, the present data indicates that there is no significant difference between the laser treatment and the traditional treatment.

5 Application: Competing Risks

In many life testing experiment the failure might occur due to different causes. In this type of experiment one observes the failure time as well as the cause of failure. It is important to study the effect of one cause in presence of other causes. In the statistical literature it is known as the competing risks problem. There are mainly two approaches to analyze competing risks data; one is latent failure type approach of Cox [2] and the other one is cause specific hazard function approach by Prentice et al. [14]. In case of exponential or Weibull failure time distributions it has been shown by Kundu [7] that both the methods provide the same likelihood function, although their interpretations are different. An extensive amount of work has been done in the area of competing risks both in the parametric and non-parametric set up. One is referred to the book by Crowder [3] for a comprehensive treatment on this topic.

A typical competing risk data is of the form (T, Δ), here T denotes the observed failure time and Δ denotes the cause of failure. The failure time T is usually assumed to be continuous, where as Δ is a discrete random variable. In this paper we have assumed Cox's latent failure approach to analyze the competing risks data. In this case it is assumed that there are total K causes of failures, i.e. Δ can take values $1, \ldots, K$. Further, there are K lifetimes say $T_1, \ldots, T_K$ due to K different causes, where $T = \min\{T_1, \ldots, T_K\}$ and $\Delta = j$ if

$T_j < \{T_1, \ldots, T_{j-1}, T_{j+1}, \ldots, T_K\}$.

We have another data set from a diabetic retinopathy study has been obtained as before. The experiment was same as before, but here the minimum time to blindness (T) and the indicator specifying whether the laser treated eye ($\Delta = 1$), the traditionally treated eye ($\Delta = 2$) or both the eyes ($\Delta = 3$) have failed simultaneously have been recorded. The data set has been presented in Table 3.

Table 3: Competing Risks Diabetic Retinopathy Data Set. Here T denotes the minimum time to blindness in days and Δ denotes its causes.

T	Δ	T	Δ	T	Δ	T	Δ	T	Δ	T	Δ
266	1	272	3	203	3	91	2	1137	3	84	1
154	2	1484	1	392	1	285	3	315	1	1140	2
583	1	287	2	901	1	547	2	1252	1	1247	3
79	1	717	2	448	2	622	3	642	1	904	2
707	2	141	2	276	1	469	2	407	1	520	1
93	1	356	1	485	2	1313	2	1653	3	248	2
805	1	427	2	503	1	344	1	699	1	423	2
790	2	36	2	285	2	125	2	667	1	315	2
777	2	588	2	727	2	306	1	471	3	210	2
415	1	126	1	409	2	307	2	350	2	584	1
637	2	350	1	355	1	577	2	663	3	1302	1
178	1	567	2	227	2	517	2	966	3		

The problem is same as before, i.e. we want to test whether there is any significant between the laser treatment and the traditional treatment in delaying the onset of blindness of the affected eyes. We treat this data as a competing risks data where the two treatments can be considered as the two different causes of failures. Here $T = \min\{T_1, T_2\}$, where T_1 (T_2) denotes the lifetime of the laser (traditionally) treated eye, and $\Delta = 1$, if $T_1 < T_2$, $\Delta = 2$, if $T_1 > T_2$ and $\Delta = 3$ if $T_1 = T_2$. Here both T_1 and T_2 are continuous random variables, but there is a positive probability that $T_1 = T_2$. We have assumed that $(T_1, T_2) \sim$ BWF$(\alpha, \beta, \lambda_1, \lambda_2, \lambda_3)$. Now based on the observations $D = \{(t_i, \delta_i); i = 1, \ldots, n\}$, the log-

likelihood function can be written as

$$l(\Theta|D) = n\log\beta + n\ln\alpha + n_1\ln\theta_1 + n_2\ln\theta_2 + n_3\ln\theta_3 + (\alpha-1)\sum_{i=1}^{n}\ln t_i - (\beta+1)\sum_{i=1}^{n}\ln(1+(\theta_1+\theta_2+\theta_3)t_i^{\alpha}). \quad (28)$$

Here n_1, n_2 and n_3 denote the number of $\delta_i = 1$, $\delta_i = 2$ and $\delta_i = 3$, respectively. The MLEs of the unknown parameters can be obtained by maximizing (28) with respect to the unknown parameters. The MLEs of the unknown parameters and the associated 95% confidence intervals are: $\widehat{\alpha} = 6.6043(\mp 1.1967)$, $\widehat{\beta} = 0.0242(\mp 0.0061)$, $\widehat{\theta}_1 = 0.5984(\mp 0.1534)$, $\widehat{\theta}_2 = 0.8245(\mp 0.2278)$ and $\widehat{\theta}_3 = 3.0644(\mp 1.0014)$. The corresponding log-likelihood value is -5191.931. We want to test whether there is any significant difference between the laser treatment and the traditional treatment and it is equivalent to test (27). Under the null hypothesis the estimates and the associated 95% confidence intervals are: $\widehat{\alpha}_0 = 6.6113(\mp 1.1913)$, $\widehat{\beta}_0 = 0.0239(\mp 0.0058)$, $\widehat{\theta}_{10} = 0.8154(\mp 0.2256)$, $\widehat{\theta}_{20} = 0.8154(\mp 0.2256)$ and $\widehat{\theta}_{30} = 3.0657(\mp 0.9976)$. The corresponding log-likelihood value is -5203.363. Based on the test statistic $-2(l_0(\widehat{\Theta}_0|D) - l(\widehat{\Theta}|D))$ we reject the null hypothesis. Hence, in this case the laser treatment has a significant different effect than the traditional treatment.

6 Conclusions

In this paper we have proposed a bivariate Weibull frailty model which is a singular distribution. The proposed model has five parameters and the joint PDF can take variety of shapes. We have derived different properties and developed the classical inference of the unknown parameters. We have used this model to analyze a dependent competing risks data. Although we have developed the classical inference, we have not developed any Bayesian inference of the unknown parameters. It will be interesting to develop the Bayesian inference of the unknown parameters for this model. More work is needed in that direction.

References

[1] Cai, J., Shi, Y. and Liu, B. (2017), "Analysis of incomplete data in the presence of dependent competing risks from Marshall-Olkin bivariate Weibull distribution under progressive hybrid censoring", *Communications in Statistics - Theory and Methods*, vol. 46, 6497– 6511.

[2] Cox D. R. (1959), "The analysis of exponentially distributed lifetimes with two types of failures", *Journal of the Royal Statistical Society Series B*, vol. 21, 411 – 421.

[3] Crowder, M. J. (2001), *Classical Competing Risks*, Chapman & Hall, Boca Raton, Florida.

[4] Feizjavadian, S.H. and Hashemi, R. (2015),"Analysis of dependent competing risks in the presence of progressive hybrid censoring using Marshall-Olkin bivariate Weibull distribution", *Computational Statistics and Data Analysis*, vol. 82, 19–34.

[5] Hu. T., Khaledi, B-E and Shaked, M. (2003), "Multivariate hazard rate", *Journal of Multivariate Analysis*, vol. 84, 173 –189.

[6] Johnson, N.L. and Kotz, S. (1975), "A vector multivariate hazard rate", *Journal of Multivariate Analysis*, vol. 5, 53–66.

[7] Kundu, D. (2004), "Parameter Estimation of the Partially Complete Time and Type of Failure Data", *Biometrical Journal*, vol. 46, no. 2, 165–179.

[8] Kundu, D. (2022), "Bivariate semi-parametric singular family of distributions and its applications", *Sankhya*, Ser B, vol. 84, Part 2, 846 – 872.

[9] Kundu, D. (2023), "Bivariate distributions with singular components", *Springer Handbook of Engineering Statistics*, Editor: Hoang Pham, Springer, pp 733 – 761.

[10] 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, no. 4, 956 – 965.

[11] 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.

[12] Kundu, D. and Gupta, R.D. (2009), "Bivariate generalized exponential distribution",*Journal of Multivariate Analysis*, vol. 100, no. 4, 581 –593.

[13] Louis, T. A. (1982), "Finding the observed information matrix when using the EM algorithm", *Journal of the Royal Statistical Society*, Series B, vol. 44, 226–233.

[14] Prentice, R. L., Kalbfleish, J. D., Peterson, Jr. A. V., Flurnoy, N., Farewell, V. T.,Breslow, N. E. (1978), "The analysis of failure times in presence of competing risks", *Biometrics*, vol. 34, 541 – 554.

[15] Samanta, D. and Kundu, D. (2023), "Bivariate semi-parametric model: Bayesian inference", *Methodology and Computing in Applied Probability*, vol. 25, Article no. 87.

[16] Shen, Y. and Xu, A. (2018), "On the dependent competing risks using Marshall-Olkin bivariate Weibull model: Parameter estimation with different methods", *Communications in Statistics - Theory and Methods*, vol. 47, 5558–5572.

[17] Vaupel, J.W., Manton, K.G. and Stallard, E. (1979), "The impact of heterogeneity in individual frailty on the dynamics of mortality", *Demography*, vol. 16, 439–454.

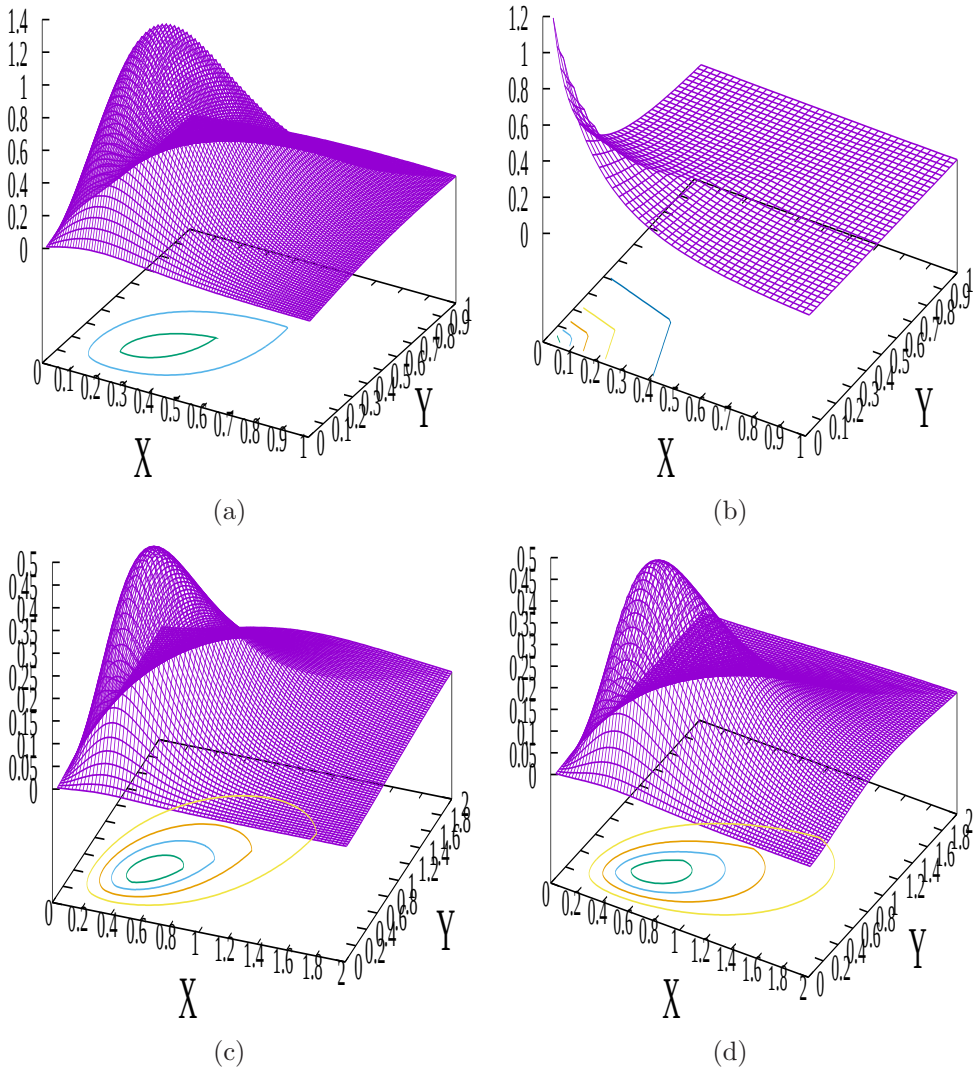


Figure 2: The contour plot of the joint PDF of the absolute continuous part of BWF$(\alpha, \beta, \theta_1, \theta_2, \theta_3)$ for different parameter values. (a) $\alpha = 2$, $\beta = 2$, $\theta_1 = \theta_2 = 1$, $\theta_3 = 2$, (b) $\alpha = 1$, $\beta = 0.5$, $\theta_1 = \theta_2 = 1$, $\theta_3 = 2$, (c) $\alpha = 2.5$, $\beta = 0.5$, $\theta_1 = 3$, $\theta_2 = \theta_3 = 1$, (d) $\alpha = 2.5$, $\beta = 0.5$, $\theta_1 = \theta_3 = 1$, $\theta_2 = 3$,

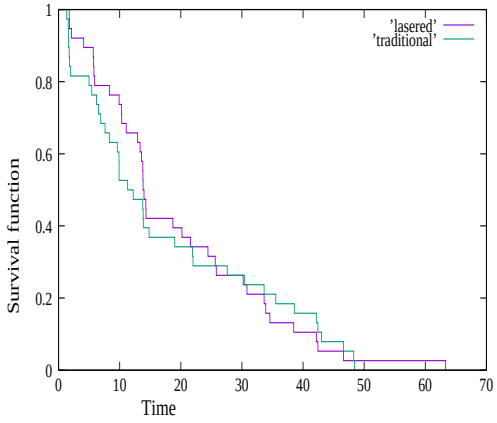
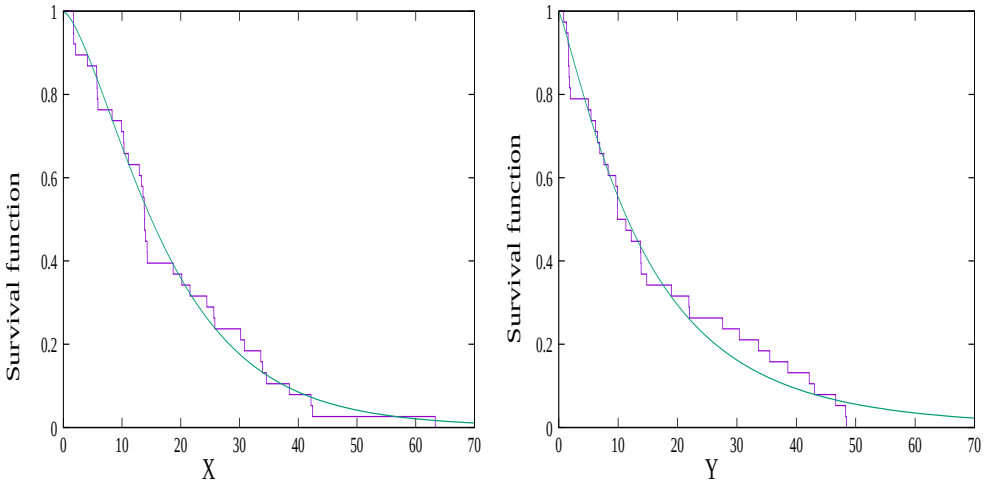


Figure 3: Empirical survival functions of the time to blindness of the two eyes.



(a) (b)

Figure 4: Empirical and fitted survival functions of the time to blindness of the two eyes (a) laser treated eye and (b) traditionally treated eye.