A new bivariate lifetime distribution: properties, estimations and its extension

Ammar M. Sarhan[(1)], Joeseph Apaloo[(2)] and Debasis Kundu[(3)]

[(1)]Department of Mathematics and Statistics, Dalhousie University, Canada. ammar.sarhan@dal.ca

[(2)]Department of Mathematics and Statistics, St. Francis Xavier University, Canada. japaloo@stfx.ca

[(3)]Department of Mathematics and Statistics, I.I.T. Kanpur, India. kundu@iitk.ac.in

Abstract

In this paper a new bivariate lifetime distribution is introduced. Its marginal distribution functions follow two-parameter Chen distribution, which has a bathtub shaped or increasing hazard rate functions. The proposed distribution, which we call a bivariate Chen distribution (BCD), is of Marshall-Olkin type and it is a singular distribution. Several properties of this proposed distribution are discussed. The BCD distribution has four unknown parameters. The maximum likelihood (ML) method and the Bayes techniques are used to estimate the unknown parameters. The maximum likelihood estimators or the Bayes estimators cannot be obtained in closed form. Numerical methods have been used in both cases. A real data set is analyzed using the proposed distribution for illustrative and comparison purposes. An application to dependent competing risks data is discussed, and finally we have extended the BCD to the multivariate case.

Key words: Marshal-Olkin, Maximum likelihood, Bayes, multivariate distribution, Competing risks models, Reliability.

1 Introduction

The bathtub shaped hazard rate distributions play an important role in modeling certain lifetime data that arise in fields such as medicine and engineering. These distributions exhibit a combination of monotone and constant hazard rate behavior. An extensive amount of work has been done related to bathtub shaped hazard

function, see for example Smith and Bain (1975), Leemis (1986), Gaver and Acar (1979), Hjorth (1980), Mudholkar and Srivastava (1993), Lemonte (2013) and see the references therein. Chen (2000) had introduced a two-parameter lifetime distribution with the following hazard function

$$h(x;\alpha,\beta) = \alpha\beta x^{\beta-1}e^{x^{\beta}}\,, x \geq 0;\ \alpha,\beta > 0. \tag{1.1}$$

The hazard function $h(.)$ can be an increasing or a bathtub shaped. It may be observed that for any $\alpha > 0$, if $\beta < 1$, the shape of the hazard function is a bathtub shaped, and for $\beta \geq 1$, it is an increasing function. This distribution is referred to as the univariate Chen distribution (UCD) and will be denoted by UCD(α,β). The associated survival function and the probability density function are respectively:

$$S_{UCD}(u;\alpha,\beta) = \exp\left\{-\alpha\left(e^{u^{\beta}}-1\right)\right\};\quad u>0, \tag{1.2}$$

$$f_{UCD}(u;\alpha,\beta) = \alpha\beta u^{\beta-1}e^{u^{\beta}}\exp\left\{-\alpha\left(e^{u^{\beta}}-1\right)\right\};\quad u>0. \tag{1.3}$$

It has been observed that the UCD is a very flexible distribution. The pdf and the hazard functions can take variety of shapes.

Sarhan and Balakrishnan (2007) proposed a bivariate distribution using the generalized exponential distribution and exponential distribution, and derived several interesting properties of that distribution. This distribution was named the Sarhan-Balakrishnan bivariate distribution (SBBD). The SBBD was obtained using the generalized exponential and exponential distributions, but the marginal distributions are not in analytic forms. Kundu and Gupta (2009) following the same idea of Sarhan and Balakrishnan (2007) proposed the bivariate generalized exponential distribution (BGED), whose marginals are generalized exponential distribution. Further, Sarhan, et al. (2011) introduced the bivariate generalized linear failure rate (BGLFR) distribution which generalizes both SBBD and BGED.

Although an extensive amount of work has been done for univariate bathtub shaped hazard function, not much work has been done in the bivariate setup. The main aim of this manuscript is to introduce bivariate Chen distribution, whose marginals are univariate Chen distribution. The proposed bivariate Chen distribution (BCD) has four parameters. Due to presence of the four parameters, the proposed BCD is a very flexible distribution. The joint pdf of the BCD can take variety of shapes. It has a very convenient copula structure. Hence, several dependency measures and dependency properties can be easily established. The estimation of the unknown parameters is an important problem for using any distribution to a real life application. We have proposed to use the maximum likelihood (ML) method and the Bayes procedure to estimate the unknown parameters. Due to complicated nature of the joint pdf, the ML estimates or the Bayes estimates cannot be obtained in closed form. Hence, we have proposed to use numerical method to compute the ML estimates, and we have used Markov Chain Monte Carlo (MCMC) method to compute the Bayes estimates and the associated credible intervals. We have performed some simulation experiments to see how the proposed methods work for different sample sizes and for different sets of parameters. Also, we have provided the analysis of a data set to see how the proposed model works in practice. Further, we have used this model to analyze dependent competing risk data, and finally we have extended it to the multivariate case also.

The rest of the paper is organized as follows. In Section 2, we introduce the model and provide some basic properties. The ML method and the Bayes estimation procedure are discussed in Section 3 and Section 4, respectively. Data generation and some simulation results have been presented in Section 5. In Section 6, we provide the analysis of a real data set. In Section 7, we have shown how this model can be used to analyze competing risks data. Finally, the multivariate generalization and

the conclusion appear in Section 8.

2 Bivariate Chen Distribution

The BCD can be defined as follows. Let U_1, U_2 and U_3 be three independent random variables, such that $U_1 \sim$ (follows) UCD(α_1, β), $U_2 \sim$ UCD(α_2, β) and $U_3 \sim$ UCD(α_3, β). Then the distribution of the random vector (X_1, X_2), where $X_1 = \min\{U_1, U_3\}$ and $X_2 = \min\{U_2, U_3\}$, is called the bivariate Chen distribution, and it will be denoted by BCD$(\alpha_1, \alpha_2, \alpha_3, \beta)$.

Note that the BCD has the same physical interpretation as the Marshall-Olkin bivariate exponential model or the shock model. Consider a system consisting of two units. For example, a desktop computer that consists of a central processing unit and a co-processor, or an aircraft with two engines. The system units are subject to three independent fatal shocks that affect one or both units. Suppose, the first shock affects component 1, the second shock affects component 2 and the third shock affects both the components. If the shocks follow UCD as defined above, then the lifetime of the two components follow the BCD.

The following Theorem 2.1 gives the joint survival function, the marginal distributions, and the joint pdf of the BCD.

Theorem 2.1 *Suppose* $(X_1, X_2) \sim BCD(\alpha_1, \alpha_2, \alpha_3, \beta)$*, then:*

1. *The joint survival function of* (X_1, X_2) *is*

$$S_{X_1,X_2}(x_1, x_2) = \exp\left\{-\sum_{i=1}^{3} \alpha_i \left(e^{x_i^{\beta}} - 1\right)\right\},\ x_3 = \max(x_1, x_2). \tag{2.1}$$

2. *The marginal distributions of* X_i *is* $UCD(\alpha_i + \alpha_3, \beta)$*,* $i = 1, 2$.

92 *3. The joint pdf of* (X_1, X_2) *is*

$$f_{X_1,X_2}(x_1,x_2) = \begin{cases} f_1(x_1,x_2) & \text{if } x_1 > x_2 > 0, \\ f_2(x_1,x_2) & \text{if } x_2 > x_1 > 0, \\ f_0(x) & \text{if } x_1 = x_2 = x > 0, \end{cases} \tag{2.2}$$

93 *where*

$$\begin{aligned} f_1(x_1,x_2) &= f_{UCD}(x_1;\alpha_1+\alpha_3,\beta) f_{UCD}(x_2;\alpha_2,\beta) \\ f_2(x_1,x_2) &= f_{UCD}(x_1;\alpha_1,\beta) f_{UCD}(x_2;\alpha_2+\alpha_3,\beta) \\ f_0(x) &= \frac{\alpha_3}{\alpha_1+\alpha_2+\alpha_3} f_{UCD}(x;\alpha_1+\alpha_2+\alpha_3,\beta). \end{aligned}$$

94 **Proof.**

95 1. The joint survival function of (X_1, X_2) is

$$\begin{aligned} S_{X_1,X_2}(x_1,x_2) &= P(X_1 > x_1, X_2 > x_2) \\ &= P(\min(U_1,U_3) > x_1, \min(U_2,U_3) > x_2) \\ &= P(U_1 > x_1, U_3 > x_1, U_2 > x_2, U_3 > x_2) \\ &= P(U_1 > x_1) P(U_2 > x_2) P(U_3 > \max(x_1,x_2)) \\ &= \prod_{j=1}^{3} S_{UCD}(x_j;\alpha_j,\beta), \text{ where } x_3 = \max(x_1,x_2). \end{aligned} \tag{2.3}$$

96 Substituting $S_{UCD}(x_j;\alpha_j,\beta)$ into (2.3), we get (2.1) that completes the proof of
97 part 1.

2. The marginal survival function of X_j, $(j = 1, 2)$, can be obtained from the joint survival function of (X_1, X_2), by setting $x_i \to 0, i \neq j$, in this case $x_3 = x_j$, and therefore,

$$S_{X_j}(x_j) = \exp\left\{-(\alpha_j+\alpha_3)\left(e^{x_j^{\beta}} - 1\right)\right\}.$$

98 That is, X_j follows UCD$(\alpha_j + \alpha_3, \beta)$ that completes the proof of part 2.

3. Using (2.1), we can be re-express $S_{X_1,X_2}(x_1,x_2)$ as

$$S_{X_1,X_2}(x_1,x_2)=\begin{cases}\exp\left\{-(\alpha_1+\alpha_3)\left(e^{x_1^\beta}-1\right)-\alpha_2\left(e^{x_2^\beta}-1\right)\right\} & \text{if } x_1>x_2>0,\\ \exp\left\{-\alpha_1\left(e^{x_1^\beta}-1\right)-(\alpha_2+\alpha_3)\left(e^{x_2^\beta}-1\right)\right\} & \text{if } x_2>x_1>0,\\ \exp\left\{-(\alpha_1+\alpha_2+\alpha_3)\left(e^{x^\beta}-1\right)\right\} & \text{if } x_1=x_2=x>0.\end{cases}$$

For $x_1\neq x_2$, the expressions for $f_1(x_1,x_2)$ and $f_2(x_1,x_2)$ can be obtained from $S_{X_1,X_2}(x_1,x_2)$ using $\frac{\partial^2}{\partial x_1\partial x_2}S_{X_1,X_2}(x_1,x_2)$. While, when $x_1=x_2=x$, we can get the expression of $f_0(x_1,x_2)$ using the following fact

$$\int_0^\infty\int_0^{x_2} f_2(x_1,x_2)\,\mathrm{d}x_1\mathrm{d}x_2+\int_0^\infty\int_0^{x_1} f_1(x_1,x_2)\,\mathrm{d}x_2\mathrm{d}x_1+\int_0^\infty f_0(x)\mathrm{d}x=1$$

That is,

$$\begin{aligned}\int_0^\infty f_0(x)\mathrm{d}x &= 1-\int_0^\infty\int_0^{x_2} f_2(x_1,x_2)\,\mathrm{d}x_1\mathrm{d}x_2+\int_0^\infty\int_0^{x_1} f_1(x_1,x_2)\,\mathrm{d}x_2\mathrm{d}x_1\\ &= 1-\int_0^\infty(\alpha_1+\alpha_2)\beta t^{\beta-1}e^{t^\beta}e^{-(\alpha_1+\alpha_2+\alpha_3)\left(e^{t^\beta}-1\right)}\mathrm{d}t\\ &= \int_0^\infty\alpha_3\beta t^{\beta-1}e^{t^\beta}e^{-(\alpha_1+\alpha_2+\alpha_3)\left(e^{t^\beta}-1\right)}\mathrm{d}t.\end{aligned}$$

Thus, we get $f_0(x)$ as expressed above, which completes the proof of part 3.

Therefore, it has been observed that the BCD is a singular distribution. It has an absolute continuous part on $x_1\neq x_2$ and it has a positive support on $x_1=x_2$. The joint pdf of the absolute continuous part of the BCD can be written as

$$f_{ac}(x_1,x_2)=\frac{\alpha_1+\alpha_2+\alpha_3}{\alpha_1+\alpha_2}\begin{cases}f_{UCD}(x_1;\alpha_1+\alpha_3,\beta)f_{UCD}(x_2;\alpha_2,\beta) & \text{if } x_1\geq x_2>0\\ f_{UCD}(x_1;\alpha_1,\beta)f_{UCD}(x_2;\alpha_2+\alpha_3,\beta) & \text{if } x_2>x_1>0.\end{cases}\tag{2.4}$$

It may be seen that $f_{ac}(x_1,x_2)$ has been defined as (2.4) with $x_1\geq x_2$, but it can be defined the other way also. The function $f_{ac}(x_1,x_2)$ is continuous everywhere on $\{(x_1,x_2);x_1\neq x_2\}$. It is discontinuous on $x_1=x_2$, unless $\alpha_1=\alpha_2$. When $\alpha_1=\alpha_2$, $f_{ac}(x_1,x_2)$ is continuous everywhere. The following result provides the shape of $f_{ac}(x_1,x_2)$.

Theorem 2.2 *The function $f_{ac}(x_1, x_2)$ for $\beta \geq 1$ is an unimodal function, and the mode will be as follows.*

1. *If $\alpha_1 = \alpha_2$, then the mode will be on $x_1 = x_2$.*
2. *If $\alpha_1 + \alpha_3 < \alpha_2$, then the mode will be in the region $\{(x_1, x_2); x_1 > x_2\}$.*
3. *If $\alpha_2 + \alpha_3 < \alpha_1$, then the mode will be in the region $\{(x_1, x_2); x_1 < x_2\}$.*
4. *If $\alpha_1 + \alpha_3 > \alpha_2$ then $\alpha_2 + \alpha_3 > \alpha_1$ and the mode will be on $x_1 = x_2$.*

Proof: The proof can be obtained in a routine manner by studying the stationary points at different regions namely $\{(x_1, x_2); x_1 < x_2\}$, $\{(x_1, x_2); x_1 > x_2\}$ and $\{(x_1, x_2); x_1 = x_2\}$ along the same line as the proof of Theorem 2.3 of Kundu and Gupta (2017). Hence, the details are avoided.

It can be easily seen that for $0 < u, v < 1$, the BCD has the following survival copula,

$$\overline{C}(u,v) = \begin{cases} uv^{\frac{\alpha_2}{\alpha_2+\alpha_3}} & \text{if} \quad u^{\frac{1}{\alpha_1+\alpha_3}} \geq v^{\frac{1}{\alpha_2+\alpha_3}} \\ u^{\frac{\alpha_1}{\alpha_1+\alpha_3}}v & \text{if} \quad u^{\frac{1}{\alpha_1+\alpha_3}} < v^{\frac{1}{\alpha_2+\alpha_3}}. \end{cases}$$

If we write $\gamma = \dfrac{\alpha_3}{\alpha_1+\alpha_3}$ and $\delta = \dfrac{\alpha_3}{\alpha_2+\alpha_3}$, then

$$\overline{C}(u,v) = \begin{cases} uv^{1-\delta} & \text{if} \quad u^{\gamma} \geq v^{\delta} \\ u^{1-\gamma}v & \text{if} \quad u^{\gamma} < v^{\delta}. \end{cases}$$

If $\alpha_1 = \alpha_2$, then $\gamma = \delta$, hence

$$\overline{C}(u,v) = \begin{cases} uv^{1-\delta} & \text{if} \quad u \geq v \\ u^{1-\delta}v & \text{if} \quad u < v. \end{cases}$$

This is the famous Marshall-Olkin copula (see [15]). It may be mentioned that using the copula properties several dependency measures and dependency properties can be explored. It has not been attempted here.

3 Maximum Likelihood Estimators

3.1 Direct Maximization

In this section we discuss the problem of computing the maximum likelihood estimates (MLEs) of the four unknown parameters of the BCD. It is assumed that we have a random sample of size n, namely $\{(x_{11}, x_{21}), (x_{12}, x_{22}), \cdots, (x_{1n}, x_{2n})\}$ from a BCD$(\alpha_1, \alpha_2, \alpha_3, \beta)$. Let us use the following notation for further development.

$$I_1 = \{i : x_{1i} < x_{2i}\},\ I_2 = \{i : x_{1i} > x_{2i}\},\ I_0 = \{i : x_{1i} = x_{2i}\},\ I = I_0 \cup I_1 \cup I_2,$$

$$n_j = |I_j|, j = 0, 1, 2, \text{ and } n = n_0 + n_1 + n_2.$$

The likelihood function is

$$\begin{aligned}
l(\theta) &= \prod_{i=1}^{n} f(x_{1i}, x_{2i}) \\
&= \prod_{i\in I_1} f_2(x_{1i}, x_{2i}) \prod_{i\in I_2} f_1(x_{1i}, x_{2i}) \prod_{i\in I_0} f_0(x_{1i}, x_{2i}) \\
&= \alpha_1^{n_1}\alpha_2^{n_2}\alpha_3^{n_0}\beta^{2n-n_0}(\alpha_2+\alpha_3)^{n_1}(\alpha_1+\alpha_3)^{n_2}e^{n(\alpha_1+\alpha_2+\alpha_3)}\left(\prod_{i\in I} x_{1i}\prod_{i\in I_1\cup I_2} x_{2i}\right)^{\beta-1} \\
&\quad \exp\left\{\sum_{i\in I_1\cup I_2} x_{2i}^{\beta} + \sum_{i\in I} x_{1i}^{\beta} - \alpha_1\sum_{i\in I} e^{x_{1i}^{\beta}} - \alpha_2\sum_{i\in I} e^{x_{2i}^{\beta}} - \alpha_3\left[\sum_{i\in I_1} e^{x_{2i}^{\beta}} + \sum_{i\in I_0\cup I_2} e^{x_{1i}^{\beta}}\right]\right\}
\end{aligned}$$

The logarithm of the likelihood function, $\mathcal{L}(\theta)$, is

$$\begin{aligned}
\mathcal{L}(\theta) &= n_1\ln(\alpha_1) + n_2\ln(\alpha_2) + n_0\ln(\alpha_3) + (2n_1 - n_0)\ln(\beta) + n_1\ln(\alpha_2+\alpha_3) \\
&+ n_2\ln(\alpha_1+\alpha_3) + n\sum_{j=1}^{3}\alpha_j + \sum_{i\in I_1\cup I_2}(x_{1i}^{\beta} + x_{2i}^{\beta}) + \sum_{i\in I_0} x_{1i}^{\beta} - \alpha_1\sum_{i\in I} e^{x_{1i}^{\beta}} - \alpha_2\sum_{i\in I} e^{x_{2i}^{\beta}} \\
&+ (\beta-1)\left\{\sum_{i\in I_1\cup I_2}(\ln x_{1i} + \ln x_{2i}) + \sum_{i\in I_0}\ln x_{1i}\right\} - \alpha_3\left[\sum_{i\in I_1} e^{x_{2i}^{\beta}} + \sum_{i\in I_0\cup I_2} e^{x_{1i}^{\beta}}\right]. \quad (3.5)
\end{aligned}$$

The following are the first partial derivatives of $\mathcal{L} = \mathcal{L}(\theta)$ with respect to the four parameters

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \alpha_1} &= n + \frac{n_1}{\alpha_1} + \frac{n_2}{\alpha_1 + \alpha_3} - \sum_{i \in I} e^{x_{1i}^{\beta}} \\
\frac{\partial \mathcal{L}}{\partial \alpha_2} &= n + \frac{n_2}{\alpha_2} + \frac{n_1}{\alpha_2 + \alpha_3} - \sum_{i \in I} e^{x_{2i}^{\beta}} \\
\frac{\partial \mathcal{L}}{\partial \alpha_3} &= n + \frac{n_0}{\alpha_3} + \frac{n_1}{\alpha_2 + \alpha_3} + \frac{n_2}{\alpha_1 + \alpha_3} - \sum_{i \in I_1} e^{x_{2i}^{\beta}} - \sum_{i \in I_0 \cup I_2} e^{x_{1i}^{\beta}} \\
\frac{\partial \mathcal{L}}{\partial \beta} &= \frac{2n - n_0}{\beta} + \sum_{i \in I_1 \cup I_2} (x_{1i}^{\beta} \ln x_{1i} + x_{2i}^{\beta} \ln x_{2i}) + \sum_{i \in I_0} x_{1i}^{\beta} \ln x_{1i} \\
&+ \sum_{i \in I_1 \cup I_2} (\ln x_{1i} + \ln x_{2i}) + \sum_{i \in I_0} \ln x_{1i} - \alpha_1 \sum_{i \in I} e^{x_{1i}^{\beta}} x_{1i}^{\beta} \ln x_{1i} - \alpha_2 \sum_{i \in I} e^{x_{2i}^{\beta}} x_{2i}^{\beta} \ln x_{2i} \\
&- \alpha_3 \left[\sum_{i \in I_1} e^{x_{2i}^{\beta}} + \sum_{i \in I_0 \cup I_2} e^{x_{1i}^{\beta}} \right].
\end{aligned}$$

The MLEs $(\hat{\alpha_1}, \hat{\alpha_2}, \hat{\alpha_3}, \hat{\beta})$ for $(\alpha_1, \alpha_2, \alpha_3, \beta)$ satisfy $\frac{\partial \mathcal{L}}{\partial \alpha_1} = 0$, $\frac{\partial \mathcal{L}}{\partial \alpha_2} = 0$, $\frac{\partial \mathcal{L}}{\partial \alpha_3} = 0$, and $\frac{\partial \mathcal{L}}{\partial \beta} = 0$ such that the associated variance-covariance matrix is positive definite. Clearly these equations cannot be solved analytically. We use numerical routines in Matlab and R to solve such system in four non-linear equations. The second order derivatives of the log-likelihood function, which are needed for the Fisher information matrix, can also be calculated, however we do not include these derivatives here. The matrix will be calculated using Matlab and R.

Let $\theta = (\alpha_1, \alpha_2, \alpha_3, \beta)$. The Fisher information matrix, denoted by $\mathcal{F}_{\hat{\theta}}$, is given by the matrix

$$\mathcal{F}_{\hat{\theta}} = \begin{pmatrix} \mathcal{L}_{11} & \mathcal{L}_{12} & \mathcal{L}_{13} & \mathcal{L}_{14} \\ \mathcal{L}_{21} & \mathcal{L}_{22} & \mathcal{L}_{23} & \mathcal{L}_{24} \\ \mathcal{L}_{31} & \mathcal{L}_{32} & \mathcal{L}_{33} & \mathcal{L}_{34} \\ \mathcal{L}_{41} & \mathcal{L}_{42} & \mathcal{L}_{43} & \mathcal{L}_{44} \end{pmatrix}$$

where $\mathcal{L}_{ij}$ is the expected value of $-\frac{\partial^2 \mathcal{L}}{\partial \theta_i \partial \theta_j}$. Since, such expected values are not obtained analytically, we can approximate them by $-\frac{\partial^2 \mathcal{L}}{\partial \theta_i \partial \theta_j}$ evaluated at the MLE $\theta_i = \hat{\theta}_i$, for $i, j = 1, 2, 3, 4$. The inverse of the information matrix is the variance-

covariance matrix of the point estimate of the vector of unknown parameters $\hat{\theta} = (\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\beta})$.

The distribution of $\hat{\theta}$ is asymptotically normal with a mean θ and variance-covariance matrix described above, i.e., $\hat{\theta} \sim \mathcal{AN}\left(\theta, \mathcal{F}_{\hat{\theta}}^{-1}\right)$. Using this asymptotic approach of the MLEs, we can calculate the asymptotic $(1-\vartheta)100\%$ confidence intervals of θ_i, for $i = 1, 2, 3, 4$ as

$$\hat{\theta}_i \pm z_{\vartheta/2}\sqrt{\text{var}(\hat{\theta}_i)} \tag{3.6}$$

where $z_{\vartheta/2}$ is the $(1-\frac{\vartheta}{2})100$th percentile of the standard normal distribution.

3.2 Expectation Maximization Algorithm

In the previous section we have seen that the MLEs of BCD can be obtained by solving a four dimensional optimization problem. To avoid that we propose to use the following Expectation Maximization (EM) algorithm, which can be obtained by combining the profile likelihood method along with the expectation maximization (EM) algorithm. It saves the computational burden significantly. We need the following result for further development.

Theorem 3.1 *If $(X_1, X_2) \sim BCD(\alpha_1, \alpha_2, \alpha_3, \beta)$, then (Y_1, Y_2) follows bivariate exponential distribution as defined by Marshall and Olkin (1967) with parameters α_1, α_2 and α_3, where*

$$Y_1 = \left(e^{X_1^{\beta}} - 1\right) \quad \text{and} \quad Y_2 = \left(e^{X_2^{\beta}} - 1\right).$$

Proof: The proof can be easily obtained by using the following transformation on X_1 and X_2 namely; $Y_1 = \left(e^{X_1^{\beta}} - 1\right)$ and $Y_2 = \left(e^{X_2^{\beta}} - 1\right)$, in the joint survival function.

Now we use the above theorem to compute the MLEs of the unknown parameters based on EM algorithm. It involves solving only one one-dimensional optimization

problem. The main idea is as follows. First it is assumed that β is known, and we transform the problem to find the MLEs of α_1, α_2 and α_3 based on EM algorithm. When β is not known, then using profile likelihood method the MLE of β can be obtained. The details are as follows. Form the given data $\{(x_{1i}, x_{2i}); i = 1, \ldots, n\}$, we obtain the transformed data as $\{(y_{1i}, y_{2i}); i = 1, \ldots, n\}$, where $y_{1i} = e^{x_{1i}^{\beta}} - 1$ and $y_{2i} = e^{x_{1i}^{\beta}} - 1$, for $i = 1, \ldots, n$. In case of Marshall-Olkin bivariate exponential distribution, Karlis (2003) proposed an EM algorithm to compute the MLEs of the unknown parameters, where at each 'E'-step the corresponding 'M'-step can be obtained in explicit form. Hence, one does not need to solve any optimization problem, to compute the MLEs of α_1, α_2 and α_3, for known β from the transformed data $\{(y_{1i}, y_{2i}); i = 1, \ldots, n\}$. When β is not known, first we obtain the MLEs of α_1, α_2 and α_3, for a given β by using the above method, and let us denote them as $\widehat{\alpha}_1(\beta)$, $\widehat{\alpha}_2(\beta)$ and $\widehat{\alpha}_3(\beta)$, respectively. Then the MLE of β can be obtained by maximizing (3.5) with respect to β, where α_1, α_2 and α_3 are replaced by $\widehat{\alpha}_1(\beta)$, $\widehat{\alpha}_2(\beta)$ and $\widehat{\alpha}_3(\beta)$, respectively. Once $\widehat{\beta}$, the MLE of β, is obtained, the MLEs of α_1, α_2 and α_3 can be obtained as $\widehat{\alpha}_1(\widehat{\beta})$, $\widehat{\alpha}_2(\widehat{\beta})$ and $\widehat{\alpha}_3(\widehat{\beta})$, respectively.

4 Bayesian analysis

In this section, we apply Bayesian method to estimate the model parameters. To do so, we assume that the model parameters are independent random variables that follow gamma prior distributions. That is, the joint prior pdf of $\theta = (\alpha_1, \alpha_2, \alpha_3, \beta)$ is

$$g_0(\theta) \propto \beta^{a_{41}-1} e^{-a_{42}\beta} \prod_{j=1}^{3} \alpha_j^{a_{j1}-1} e^{-a_{j2}\alpha_j}\,,\ \alpha_1, \alpha_2, \alpha_3, \beta > 0\,, \tag{4.1}$$

where the hyperparameters $a_{ij} > 0$, $i = 1, 2, 3, 4$ and $j = 1, 2$ are all known. It may be mentioned that the gamma priors are very flexible priors, and with the different choices of the hyperparameters, they can take variety of shapes. Analytically it

is quite tractable, moreover, the 'non-informative' priors can be obtained as limiting case. Hence, it is quite common to take gamma priors in case of non-negative unbounded parameters. The independence assumption has been made mainly for analytical purposes.

The log-prior density function is

$$\log(g_0(\theta)) \propto (a_{41}-1)\ln(\beta) - a_{42}\beta + \sum_{j=1}^{3}\left[(a_{j1}-1)\ln(\alpha_j) - a_{j2}\alpha_j\right] . \tag{4.2}$$

Using the log-prior (4.2) and log-likelihood function (3.5) in the Bayes Theorem, the joint posterior probability density function of θ, given data, can be written as

$$g(\theta|\text{data}) = \frac{1}{I_0}\exp\{\mathcal{L}(\theta) + \log(g_0(\theta))\}\,, \tag{4.3}$$

where $I_0 = \int_0^\infty \exp\{\mathcal{L}(\theta) + \log(g_0(\theta))\}d\theta$ is the normalizing constant. Obviously, I_0 cannot be calculated analytically and hence the joint posterior distribution of θ is analytically intractable. It is known that under the squared-error loss, Bayes estimators for any function of θ, say $w(\theta)$, is given as the ratio of two multidimensional integrals given by

$$\hat{w}(\theta) = \frac{I_w}{I_0} = \frac{\int_0^\infty w(\theta)\,\exp\{\mathcal{L}(\theta) + \log(g_0(\theta))\}d\theta}{\int_0^\infty \exp\{\mathcal{L}(\theta) + \log(g_0(\theta))\}d\theta}\,. \tag{4.4}$$

Again I_w does not have analytic solution, as a result, $\hat{w}(\theta)$ can not be computed analytically. We should use approximation methods to solve the ratio of the integrals. We will use Markov Chain Monte Carlo (MCMC) method to approximate the Bayes estimate of the function $w(\theta)$. One of the great advantages of the MCMC is that it does not require to calculate the integrals I_0 and I_w that are needed for $\hat{w}(\theta)$. The main idea in the MCMC method is to generate random draws from the joint posterior distribution by generating draws from an arbitrary distribution, named proposal distribution, which satisfies two conditions: (i) it mimics the posterior distribution,

and (*ii*) easy to simulate from. For each random draw obtained from the proposed distribution, we apply the accept-reject method to obtain a sequence of random draws from the posterior distribution. Note that the underlying parameters are all positive, this will limit our choices of the proposal distribution of θ. To overcome this problem, we use the log-transform of the parameters, $\phi = \log\theta = (\log\alpha_1, \log\alpha_2, \log\alpha_3, \log\beta)$. The joint pdf of the transformed parameters, ϕ, is

$$g_\phi(\phi|\text{data}) = \frac{1}{I_0}\exp\left\{\mathcal{L}\left(\theta = e^\phi\right) + \log\left(g_0\left(\theta = e^\phi\right)\right) + \sum_{j=1}^{4}\phi_j\right\}, \qquad (4.5)$$

where $\phi_j \in (-\infty, \infty)$, $j = 1, 2, 3, 4$. Transforming the support of the original parameters from the positive real line to the whole real line, will allow us to choose a proposal distribution for $g_\phi(\phi|\text{data})$ form a wider list of distributions. In this article, we use multivariate normal distribution as a proposal for $g_\phi(\phi|\text{data})$. The following steps are followed to generate random draws from the posterior distribution (4.3) without computing the normalizing constant I_0:

1. Set the size of the random draws we wish to generate as, say $M = 2M_0$, where M_0 is a whole number.

2. Choose an initial guess of ϕ, say $\phi^{(0)}$.

3. For $i = 1, 2, \cdots, M$, perform the following steps:

   (a) Generate ϕ^* from the multivariate normal with mean $\phi^{(i-1)}$ and variance-covariance Σ.

   (b) Compute the ratio $\kappa = \min\left\{1, \dfrac{g_\phi(\phi^*|\text{data})}{g(\phi^{(i-1)}|\text{data})}\right\}$.

   (c) Generate a random value u from uniform distribution on $(0, 1)$.

   (d) If $\kappa \geq u$ set $\phi^{(i)} = \phi^*$, otherwise set $\phi^{(i)} = \phi^{(i-1)}$.

We discard the early M_0 number of burn-in draws and use the remaining $M - M_0$, "$\phi^{(M_0+1)}, \phi^{(M_0+2)}, \cdots, \phi^{(M)}$", as the desired draws from (4.5). Finally, use the inverse transformation of these draws to get a set of random draws $\theta^{(1)} = e^{\phi^{(M_0+1)}}, \theta^{(2)} = e^{\phi^{(M_0+2)}}, \cdots, \theta^{(M_0)} = e^{\phi^{(M)}}$ from (4.3). Thus Bayes estimate of $w(\theta)$ is

$$\hat{w}(\theta) = \frac{1}{M_0} \sum_{i=1}^{M_0} w(\theta^{(i)}).$$

Furthermore, the lower and upper bounds of the $100(1-\vartheta)\%$, $0 < \vartheta < 1$, Bayesian probability interval (BPI) of θ_j are the $\frac{\vartheta}{2}\,100$th and $(1-\frac{\vartheta}{2})\,100$th percentiles of the sequence of the M_0 draws $\theta_j^{(1)}, \theta_j^{(2)}, \cdots, \theta_j^{(M_0)}$, respectively, for $j = 1, 2, 3, 4$.

5 Data Generation and Simulation Results

In this section first we show how a data set can be generated from the proposed BCD model. We have further performed some simulation experiments to see how the proposed methods work for different sample sizes and for different parameter values.

5.1 Data Generation from BCD

We can generate a random sample from the BCD$(\alpha_1, \alpha_2, \alpha_3, \beta)$ by applying the following Algorithm:

Algorithm 1:

1. Specify the model parameters namely; $\alpha_1, \alpha_2, \alpha_3$ and β.
2. Generate v_1, v_2, v_3 independently from uniform $(0, 1)$.
3. Obtain
$$u_i = \exp\left\{\frac{1}{\beta}\,\log\left(\log\left(1 - \frac{\log(v_i)}{\alpha_i}\right)\right)\right\}\ i = 1, 2, 3.$$
4. Define $x_1 = \min(u_1, u_3)$ and $x_2 = \min(u_2, u_3)$. Then (x_1, x_2) is a random sample from BCD$(\alpha_1, \alpha_2, \alpha_3, \beta)$.

5.2 Simulation Results

In this section we present some simulation results which we have performed to see how the proposed methods work for different sample sizes and for different parameter values. We have taken sample sizes as $n = 20, 30, \ldots, 100$ and three different sets of parameter values. We have reported the average estimates and the MSEs for both the MLEs and Bayes estimates based on 2000 replications.

In case of Bayes estimates we have assumed that the parameters are independent and follow gamma prior distribution with hyperparameters equal to 0.001. Furthermore, to compute the Bayes estimates, we have used the MCMC with a multinormal distribution with the mean equals the mode of the poster distribution and variance covariance matrix as the asymptotic variance-covariance of the MLEs. The results are reported in Tables 1 and 2. Some of the points are quite clear from these simulation results. It is observed that in all these cases as the sample size increases the biases and MSEs decrease for both MLEs and Bayes estimators. It indicates the consistency properties of both these methods. Further both the MLEs and Bayes estimators behave in a very similar manner both in terms of biases and MSEs.

6 Data Analysis

In this section we illustrate how the new model can be used to analyze a real data set. The Union of European Football Associations (UEFA) Champion's League data (Meintanis, 2007; Table 3) consist of times at which at least one goal was scored by the home team and at least one goal was scored through a kick resulting from a penalty, foul or other by any team. The data set was derived at the group stage for the year periods 2004-05 and 2005-06. The data has been analyzed in the literature by various methods. For example, Meintanis (2007) used the Marshall-Olkin bivariate exponential model, Kundu and Gupta (2009) used a bivariate generalized exponential

model, and Sarhan et al. (2011) used the bivariate generalized linear failure rate (BGLFR) model. Each of these models provided a better fit than the fit in the preceding paper.

As an application, we analyze the UEFA data reported in Meintanis (2007) by the proposed BCD developed in this paper. We denote the time in minutes of the first goal scored by any team through a kick by X_1, and let X_2 denote the time in minutes of the first goal of any type scored by the home team.

As a quick non-parametric test that can be used to choose which univariate model can be a good fit for the marginal data, the total time on test transform (TTT-Transform) plots are provided in Figure 1. It seems from this figure that the marginal distributions should have increasing failure rates. Also, the parametric TTT-Transform plots using the marginal distributions of the three models used here (BVGE, BGLFR and BCD) using the MLE of the model parameters are shown in Figure 1. It shows from the plots that, the three models exhibit increasing failure rate and therefore any of these three models can be a suitable fit for the underlying real data set.

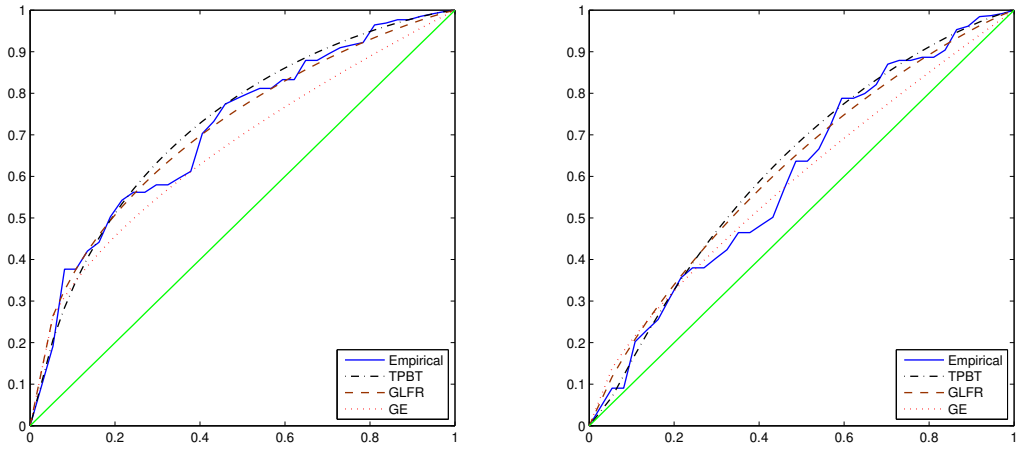

Figure 1: The non-parametric and parametric TTT-Transforms of the marginal data using the three models used in this paper.

To investigate if the BCD provides a better fit to the data set than the BVGE and the BGLFR distributions, we use the Akaike Information Criterion (AIC = -2$\hat{\mathcal{L}}$ + 2k, where k is the number of the model parameters). Table 6 provides the MLEs of the unknown parameters of the BVGE, BGLFR and BCD models and the corresponding $\hat{\mathcal{L}}$ and AIC. Based on the values of the AIC statistics, we could conclude that the BCD fits the data better than the BVGE and BGLFR distributions.

The inverse of the estimated Fisher-information matrix is

$$\mathfrak{F}^{-1}(\hat{\alpha},\hat{\beta}) = \begin{bmatrix} 2.6417\times10^{-6} & 3.2545\times10^{-6} & 2.8593\times10^{-6} & -2.4597\times10^{-5} \\ 3.2545\times10^{-6} & 9.3161\times10^{-6} & 7.0488\times10^{-6} & -5.4256\times10^{-5} \\ 2.8593\times10^{-6} & 7.0488\times10^{-6} & 9.0340\times10^{-6} & -5.4989\times10^{-5} \\ -2.4597\times10^{-5} & -5.4256\times10^{-5} & -5.4989\times10^{-5} & 4.1448\times10^{-4} \end{bmatrix}$$

The margin of error of the MLE of α_1 is 0.0032, which is larger than the MLE of $\alpha_1 = 0.0028$. This leads to a negative lower limit of the CI that contradicts the nature of the parameter $\alpha_1 > 0$. To overcome this problem, we used the log-transformation of the four parameters then we calculated the approximated confidence intervals of

the parameters. The approximations for the 95% confidence intervals of the model parameters along with their widths are given in Table 5.

In the Bayes analysis, gamma prior distributions are applied for the four unknown parameters using two scenarios of the hyperparameters. In the first (scenario I), we set all the hyperparameters equal to 0.001, while in the second (scenario II) we use the posterior means and variances from the results of the first scenario to be the prior means and variances then calculate the values of the hyperparameters as $a_{11} = 2.6$, $a_{12} = 1421.9$, $a_{21} = 4.3$, $a_{22} = 981.6$, $a_{31} = 3.8$, $a_{32} = 877.0$, $a_{41} = 519.0$ and $a_{42} = 1233.4$. We call the first scenario as non-informative prior while the second as informative prior.

Table 4 shows the summary statistics for the four parameters using Bayes method when gamma prior distributions with different values of hyperparameters are applied. Table 5 displays the 95% credible intervals for the model parameters using the two scenarios along with their widths. Based on the widths of the intervals, we can conclude that the Bayes estimates under scenario II are better than those under scenario I which are better than those obtained using the MLE, in the sense of having smaller widths.

	ML	Bayes	
Parameter		Non-informative	Informative
α_1	(0.000909, 0.008728)	(0.000465, 0.004743)	(0.0007312, 0.003370)
Width	0.0078	0.0043	0.0026
α_2	(0.002436, 0.016283)	(0.001484, 0.009526)	(0.0023894, 0.006883)
Width	0.0138	0.0080	0.0045
α_3	(0.002252, 0.016016)	(0.001442, 0.009573)	(0.0021605, 0.006912)
Width	0.0138	0.0081	0.0048
β	(0.365473, 0.445408)	(0.384471, 0.455917)	(0.4010465, 0.439018)
Width	0.0799	0.0714	0.0380

Table 5: 95% confidence and credible intervals for the model parameters along with their widths.

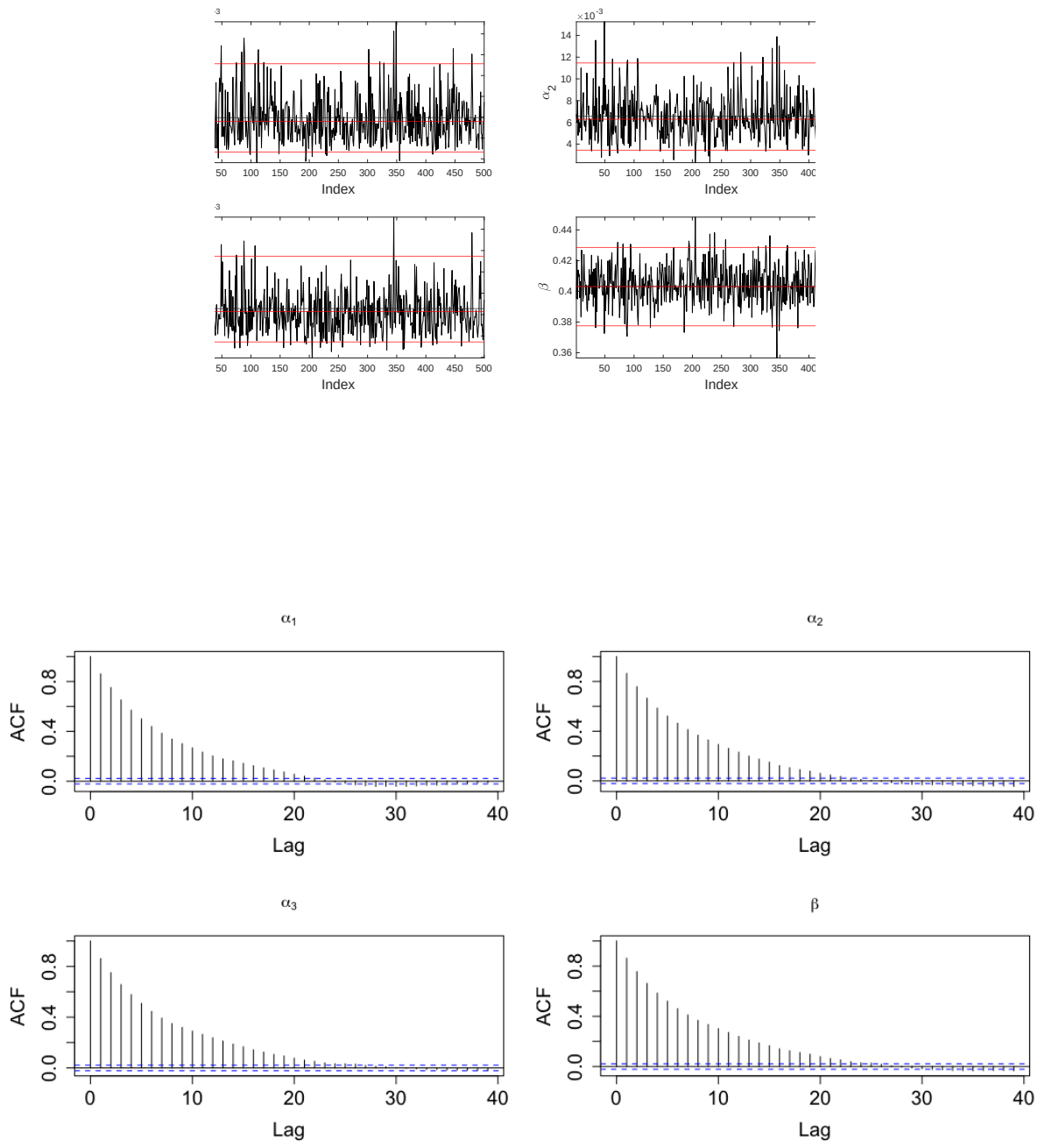

Figure 2: The trace plots for the simulations of the four parameters and the corresponding autocorrelations using the poor choice of the hyperparameters' values.

It seems from the plots in Figure 2 that there is good mix of the draws that implies that the draws are randomly generated from the posterior distribution and the autocorrelations go to zero rapidly which indicate that the random draws become more independent very fast. The marginal posterior densities using the two prior scenarios are given in Figure 3, which don't show big difference between the two scenarios.

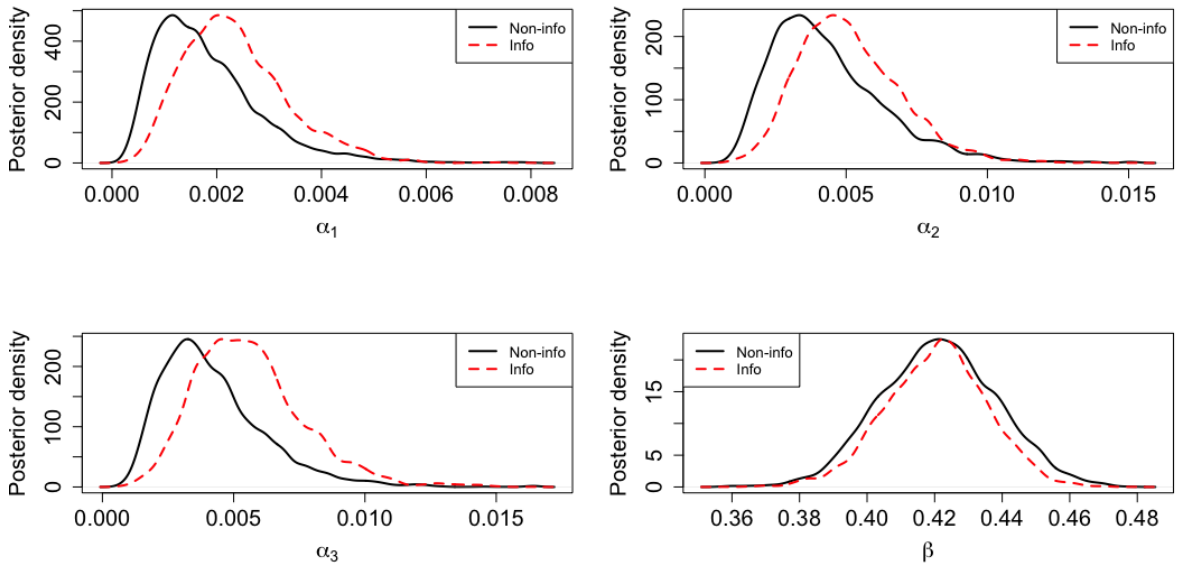

Figure 3: The marginal posterior densities of the four parameters using scenario I (solid-black) and scenario II (dashed-red) of the hyperparameters.

7 Competing Risks Application

In survival or in reliability analysis one often is interested in the assessment of one risk in presence of other risk factors. In the statistical analysis it is popularly known as the competing risk problem. The competing risk model assumes that the data consists of the failure time and the associated cause of failure. An extensive amount of study

has been carried out both under the parametric and non-parametric models, see for example the monograph by Crowder (2001) in this respect. There are mainly two different approaches to handle competing risks problem. One is known as the latent failure time model, originally proposed by Cox (1959), and the other is known as the cause specific hazard model proposed by Prentice et al. (1978). In case of exponential or Weibull lifetime model it has been shown by Kundu (2004) that both approaches provide the same likelihood function, although their interpretations are quite different. In this paper we have adopted the latent failure time model assumption and it has been assumed that there are only two possible causes of death, although the method can be easily generalized for any arbitrary number of causes also.

7.1 Model Assumptions

It is assumed a life testing experiment starts at time zero, with n number of identical units. We observe the time to failure and also the associated cause of failure. It is further observed that a unit can fail at time due to both causes. Let X_1 and X_2 be the random variables associated with the lifetime of the experimental unit due to Cause 1 and Cause 2, respectively. Hence, the observe failure time $T = \min\{X_1, X_2\}$, and the associate cause of failure

$$D = \begin{cases} 1 & \text{if} \quad X_1 < X_2 \\ 2 & \text{if} \quad X_2 < X_1 \\ 0 & \text{if} \quad X_1 = X_2. \end{cases}$$

In a competing risk model it is assumed that a sample of size n is obtained from (T, D). It may be mentioned that the assumption $X_1 = X_2$, is not very common. But there are competing risk data sets, where it has been observed that a unit fails due to both causes, hence we have made this assumption.

There exists an extensive amount of literature dealing with the analysis of competing risks data based on the specific continuous parametric distribution assumption on

X_1 and X_2, and assuming X_1 and X_2 to be independently distributed, see for example Crowder (2001). It is very clear that if both X_1 and X_2 are continuous distributions and they are independently distributed, then $P(X_1 = X_2) = 0$. Hence, if it is assumed that $P(D = 0) > 0$, then the independent assumption of X_1 and X_2 does not hold. Due to this reason, recently some work going on is based on the assumption that (X_1, X_2) has a specific bivariate distribution, see for example Feizjavdian and Hashemi (2015), Shen and Xu (2018) and Samanta and Kundu (2021). Importantly, in all these cases it has been assumed that (X_1, X_2) has a bivariate Marshall-Olkin Weibull distribution. Note that if (X_1, X_2) has the bivariate Marshall-Olkin Weibull distribution, then the marginals can have only monotone hazard functions. If the data indicate that the marginals are not monotone, it may not be proper to use this distribution. Due to this reason, we have provided the analysis based on the assumption that $(X_1, X_2) \sim \text{BCD}(\alpha_1, \alpha_2, \alpha_3, \beta)$. In that case the marginals may have non-monotone hazard functions. Hence, the practitioner has a choice of using another dependent competing risks model if it is observed that a failure can take place due to both causes.

It is assumed that the following competing risks data have been observed: $\{(t_1, d_1), \ldots, (t_n, d_n)\}$. Before, progressing further we provide the log-likelihood contribution of a typical data point (t, d), when $d = 0$, 1 or 2. The log-likelihood contribution of $(t, 0)$ is

$$\ln \alpha_3 + \ln \beta + (\beta - 1) \ln t + t^\beta - (\alpha_1 + \alpha_2 + \alpha_3)(e^{t^\beta} - 1).$$

Similarly, the log-likelihood contribution of $(t, 1)$ and $(t, 2)$ are, respectively,

$$\ln \alpha_1 + \ln \beta + (\beta - 1) \ln t + t^\beta - (\alpha_1 + \alpha_2 + \alpha_3)(e^{t^\beta} - 1),$$

and

$$\ln \alpha_2 + \ln \beta + (\beta - 1) \ln t + t^\beta - (\alpha_1 + \alpha_2 + \alpha_3)(e^{t^\beta} - 1).$$

Hence, the log-likelihood function can be written as

$$L(\theta) = n_0 \ln \alpha_3 + n_1 \ln \alpha_1 + n_2 \ln \alpha_2 + n \ln \beta + (\beta - 1) \sum_{i=1}^{n} \ln t_i + \sum_{i=1}^{n} t_i^{\beta} - (\alpha_1 + \alpha_2 + \alpha_3) \sum_{i=1}^{n} \left(e^{t_i^{\beta}} - 1 \right), \quad (7.1)$$

here $\theta = (\alpha_1, \alpha_2, \alpha_3, \beta)$ and $n_j = \#\{i : d_i = j\}$, for $j = 0, 1, 2$. The MLEs of the unknown parameters can be obtained by maximizing (7.1) with respect to the unknown parameters. For a given β, the MLEs of α_1, α_2 and α_3, say $\widehat{\alpha}_1(\beta)$, $\widehat{\alpha}_2(\beta)$ and $\widehat{\alpha}_3(\beta)$, can be obtained as

$$\widehat{\alpha}_1(\beta) = \frac{n_1}{\sum_{i=1}^{n} \left(e^{t_i^{\beta}} - 1 \right)}, \quad \widehat{\alpha}_2(\beta) = \frac{n_2}{\sum_{i=1}^{n} \left(e^{t_i^{\beta}} - 1 \right)} \quad \text{and} \quad \widehat{\alpha}_3(\beta) = \frac{n_0}{\sum_{i=1}^{n} \left(e^{t_i^{\beta}} - 1 \right)}.$$

Once, $\widehat{\alpha}_1(\beta)$, $\widehat{\alpha}_2(\beta)$ and $\widehat{\alpha}_3(\beta)$ are obtained, the MLE of β can be obtained by maximizing the profile log-likelihood of β, namely $L(\widehat{\alpha}_1(\beta), \widehat{\alpha}_2(\beta), \widehat{\alpha}_3(\beta), \beta)$. The associated confidence intervals can be obtained from the inverse of the observed Fisher information matrix.

7.2 Diabetic Retinopathy Data Analysis

The diabetic Retinopathy disease is a major cause of blindness and vision loss among diabetic patients. The present data set has been obtained from the National Eye Institute, where an attempt has been made to see the effect of laser treatment in reducing the risk of blindness. The experiment was conducted on seventy one patients, where all of them were diabetic. The experiment is as follows. For each patient, one eye was selected at random, where the laser treatment was given. The time to blindness since the laser treatment was given, and the indicator mentioning whether treated, untreated or both eyes became blind has been recorded. The main aim of this study is to examine whether the laser treatment has any effect in delaying (reducing) the blindness or not. The data are presented in Table 6.

The MLEs of the unknown parameters along with their 95% confidence intervals are provided within brackets:

$$\widehat{\alpha}_1 = 0.5338(\pm 0.1823), \qquad \widehat{\alpha}_2 = 0.6291(\pm 0.1898)$$

$$\widehat{\alpha}_3 = 0.1906(\pm 0.0657), \qquad \widehat{\beta} = 1.2004(\pm 0.5665).$$

The corresponding log-likelihood value is -121.7658. One natural question is whether BCD fits the competing risk data or not. We have obtained the Kolmogorv-Smirnov (KS) distance between the empirical cumulative distribution function (CDF) and the fitted CDF. The KS distance is 0.0948 and the associated p value is 0.5452. Hence, it provides a good fit. Note that due to complicated nature of the profile log-likelihood function of β, it is difficult to show analytically that it has a unique maximum. In Figure 4 we have provided the plot of the profile log-likelihood function β, and it is an unimodal function. Further, in Figure 4 we have also provided the empirical and fitted survival function of the competing risk data. It indicates that they match reasonably well.

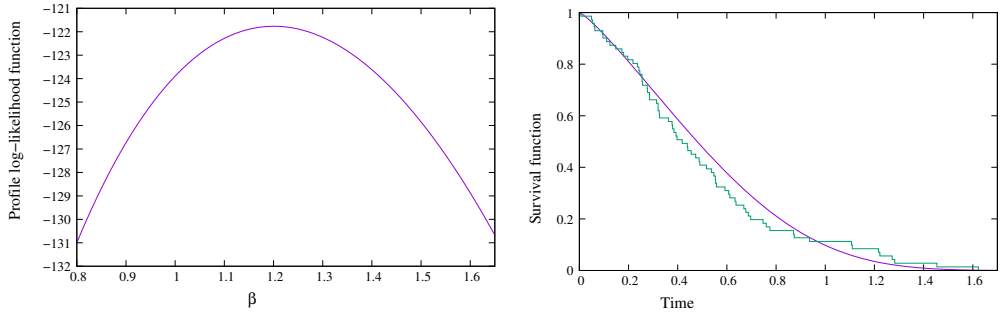

Figure 4: The profile log-likelihood function of β (left). The empirical and fitted survival function (right).

Now we would like to examine whether laser treatment has any significant effect or not. We want to test the following hypothesis

$$H_0 : \alpha_1 = \alpha_2 \quad vs. \quad H_1 : \alpha_1 \neq \alpha_2.$$

We would like to use the likelihood ratio test for this purpose. Let us assume $\alpha_1 = \alpha_2 = \alpha$, and under the null hypothesis, the log-likelihood function becomes

$$\begin{aligned} L_0(\theta_0) &= n_0 \ln \alpha_3 + (n_1 + n_2) \ln \alpha + n \ln \beta + (\beta - 1) \sum_{i=1}^{n} \ln t_i + \sum_{i=1}^{n} t_i^{\beta} \\ &\quad -(2\alpha + \alpha_3) \sum_{i=1}^{n} \left(e^{t_i^{\beta}} - 1 \right), \end{aligned} \tag{7.2}$$

Here $\theta_0 = (\alpha, \alpha_3, \beta)$. In this case also, the MLEs can be obtained along the same line as the unrestricted case. For a given β, the MLEs of α and α_3, say $\widehat{\alpha}(\beta)$, and $\widehat{\alpha}_3(\beta)$, can be obtained as

$$\widehat{\alpha}(\beta) = \frac{n_1 + n_2}{2 \sum_{i=1}^{n} \left(e^{t_i^{\beta}} - 1 \right)}, \quad \text{and} \quad \widehat{\alpha}_3(\beta) = \frac{n_0}{\sum_{i=1}^{n} \left(e^{t_i^{\beta}} - 1 \right)}.$$

Again, once, $\widehat{\alpha}(\beta)$ and $\widehat{\alpha}_3(\beta)$ are obtained, the MLE of β can be obtained by maximizing the profile log-likelihood of β. The MLEs of the unknown parameters under the null hypothesis, along with their 95% confidence intervals are provided within brackets:

$$\widehat{\alpha} = 0.5814(\pm 0.1857), \quad \widehat{\alpha}_3 = 0.1906(\pm 0.0649), \quad \widehat{\beta} = 1.2004(\pm 0.5638).$$

The associated log-likelihood value is -121.9709. Hence the value of the test statistic becomes -2(121.7658-121.9704) = 0.4102, and the associated p value becomes 0.521. Hence, we cannot reject H_0. It indicates that the laser treatment does not have any effect in delaying blindness.

8 Extension and Conclusion

In the previous sections we have discussed about the univariate and bivariate Chen distribution. In this section it has been extended to the multivariate (m-variate) case. The extension is quite straight forward, and it can be done along the same line. We

will call this new distribution the multivariate Chen distribution (MCD) and it will be denoted by MCD$(\alpha_1, \alpha_2, \cdots, \alpha_{m+1}, \beta)$. The MCD$(\alpha_1, \alpha_2, \cdots, \alpha_{m+1}, \beta)$ is defined as follows.

Definition 8.1 Let U_i, $i = 1, 2, \cdots, m+1$, be mutually independent random variables which follow UCD(α_i, β). Define $X_i = \min(U_i, U_{m+1})$, $i = 1, 2, \cdots, m$. Then, the random vector $\mathbf{X} = (X_1, X_2, \cdots, X_m)$ follows MCD$(\alpha_1, \alpha_2, \cdots, \alpha_{m+1}, \beta)$.

Theorem 8.1 *If $\boldsymbol{X} = (X_1, X_2, \cdots, X_m)$ follows MCD$(\alpha_1, \alpha_2, \cdots, \alpha_{m+1}, \beta)$, then*

1. *The joint survival function of $\boldsymbol{X}$ is*

$$S_{\boldsymbol{X}}(\boldsymbol{x}) = \exp\left\{-\sum_{i=1}^{m+1} \alpha_i \left(e^{x_i^\beta} - 1\right)\right\}, \tag{8.1}$$

   where $\boldsymbol{x} = (x_1, \cdots, x_m)$ and $x_{m+1} = \max(x_1, \cdots, x_m)$.

2. *The marginal distribution of X_i is UCD$(\alpha_i + \alpha_{m+1}, \beta)$, $i = 1, 2, \cdots, m$.*

3. *The joint pdf of $\boldsymbol{X}$ is*

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \begin{cases} f_j(\boldsymbol{x}) & \text{if } x_j = \max(x_1, \cdots, x_m) > 0,\ j = 1, 2, \cdots, m, \\ f_0(x) & \text{if } x_1 = x_2 = \cdots = x_m = x > 0, \end{cases} \tag{8.2}$$

   where

$$f_j(\boldsymbol{x}) = \left[1 + \frac{\alpha_{m+1}}{\alpha_j}\right] \beta^m \prod_{i=1}^{m} \alpha_i x_i^{\beta-1} \exp\left\{\sum_{i=1}^{m+1} \alpha_i - \alpha_{m+1} e^{x_j^\beta} - \sum_{i=1}^{m} \left[\alpha_i e^{x_i^\beta} - x_i^\beta\right]\right\},$$

   and

$$f_0(x) = \alpha_{m+1} \beta x^{\beta-1} \exp\left\{x^\beta - \left(\sum_{j=1}^{m+1} \alpha_j\right)\left(e^{x^\beta} - 1\right)\right\}.$$

Proof. The proof is similar to the proof of Theorem 2.1, and therefore, we omitted it.

The m variate MCD has $m + 2$ unknown parameters. It will be important to develop both the classical and Bayesian inference of the unknown parameters of this

399 model. Moreover, it will be also important to develop proper inference procedure in
400 case of dependent competing risks model when there are m causes of failures. More
401 work is needed along that direction.

402 Acknowledgements:

403 The authors would like to thank the unknown reviewers and the Associate Editor
404 for their constructive comments which have helped to improve the manuscript signif-
405 icantly.

406 References

407 [1] Chen, Z. (2000), A New Two-Parameter Lifetime Distribution with Bathtub
408 Shape or Increasing Failure Rate Function. *Statistics and Probability Letters*, 49,
409 155-161.

410 [2] Cox, D.R. (1959). The analysis of exponentially distributed lifetimes with two
411 types of failures. *Journal of the Royal Statistical Society Series B*, 21, 411 - 421.

412 [3] Crowder, M. (2001). *Classical Competing Risks Model*, Chapman & Hall, New
413 York.

414 [4] Feizjavdian, S.H. and Hashemi, R. (2015). Analysis of dependent competing
415 risks in presence of progressive hybrid censoring using Marshall-Olkin bivariate
416 Weibull distri-bution. *Computational Statistics and Data Analysis*, 82, 19–34.

417 [5] Gaver, D.P. and Acar, M. (1979). Analytical hazard representations for use
418 in reliability, mortality, and simulation studies. *Communications in Statistics-*
419 *Simulation and Computation*, 8(2), 91-111.

[6] Kundu, D. and Gupta, A.K. (2017), "On bivariate inverse Weibull distribution", *Brazilian Journal of Probability and Statistics*, vol. 31, no. 2, 275 – 302.

[7] Hjorth, U. (1980). A Reliability Distribution with Increasing, Decreasing, Constant and Bathtub-Shaped Failure Rates. *Technometrics*, 22(1), 99-107. doi:10.2307/1268388

[8] Karlis, D. (2003). ML estimation for multivariate shock models via an EM algorithm. *Ann. Inst. Statist. Math.*, 55, 817 - 830.

[9] Kundu, D. (2004). Parameter estimation of the partially complete time and type of failure data. *Biometrical Journal*, 46, 165–179.

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

[11] Leemis, L. (1986). Relationships Among Common Univariate Distributions. *The American Statistician*, 40,143-146.

[12] Lemonte, A.J. (2013). A new exponential-type distribution with constant, decreasing, increasing, upside-down bathtub and bathtub-shaped failure rate function. *Computational Statistics & Data Analysis*, 62(C), 149-170.

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

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

[15] Nelson, R.B. (2007), An Introduction to Copulas, 2nd ed. Springer Science and Business Media.

[16] Mudholkar, G.S. and Srivastava, D.K. (1993). Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE Transactions on Reliability*, 42(2), 299-302.

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

[18] Samanta, D. and Kundu, D. (2021). Bayesian Inference of a Dependent Competing Risk Data. *Journal of Statistical Computation and Simulation*, (to appear).

[19] Sarhan, A.M. and Balakrishnan, N. (2007). A new class of bivariate distributions and its mixture. *Journal of Multivariate Analysis*. 98. 1508-1527.

[20] 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*, 55, 644-654.

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

[22] Smith, R. and Bain, L. (1975). An Exponential Power Life-Testing Distribution. *Communications in Statistics-Theory and Methods*, 4, 469-481.

Table 1: The average of MLEs and Bayes estimates and corresponding MSE.

Case 1: $\alpha_1 = 0.1, \alpha_2 = 0.1, \alpha_3 = 0.3, \beta = 0.3$									
		Average point estimate				MSE			
n	Method	α_1	α_2	α_3	β	α_1	α_2	α_3	β
20	ML	0.102	0.101	0.303	0.317	0.00274	0.00282	0.00862	0.00224
20	Bayes	0.103	0.102	0.309	0.312	0.00284	0.00288	0.009	0.00214
30	ML	0.102	0.101	0.3	0.311	0.00183	0.00182	0.00542	0.00125
30	Bayes	0.102	0.101	0.303	0.307	0.00184	0.00182	0.00536	0.00119
40	ML	0.100	0.101	0.299	0.308	0.00128	0.00132	0.00425	0.00092
40	Bayes	0.101	0.101	0.302	0.306	0.0013	0.00135	0.00421	0.00088
50	ML	0.101	0.101	0.305	0.305	0.00107	0.00102	0.00325	0.00065
50	Bayes	0.101	0.101	0.307	0.303	0.00108	0.00103	0.00326	0.00063
60	ML	0.100	0.100	0.298	0.307	0.00089	0.00086	0.0027	0.00058
60	Bayes	0.100	0.100	0.300	0.305	0.00089	0.00086	0.00267	0.00055
70	ML	0.100	0.101	0.301	0.304	0.00077	0.00074	0.00226	0.00047
70	Bayes	0.101	0.101	0.303	0.303	0.00078	0.00074	0.00228	0.00045
80	ML	0.100	0.101	0.302	0.304	0.00064	0.00066	0.00201	4e-04
80	Bayes	0.100	0.101	0.304	0.303	0.00065	0.00067	0.00202	0.00039
90	ML	0.101	0.100	0.301	0.304	0.00061	0.00056	0.00181	0.00036
90	Bayes	0.101	0.100	0.302	0.302	0.00061	0.00056	0.00182	0.00036
100	ML	0.100	0.100	0.299	0.303	0.00048	0.00051	0.00162	0.00032
100	Bayes	0.100	0.101	0.301	0.302	0.00048	0.00051	0.00161	0.00031
Case 2: $\alpha_1 = 0.1, \alpha_2 = 0.2, \alpha_3 = 0.3, \beta = 0.3$									
20	ML	0.104	0.206	0.304	0.316	0.00334	0.00702	0.00929	0.00213
20	Bayes	0.105	0.207	0.310	0.312	0.00332	0.00734	0.0105	0.00207
30	ML	0.102	0.203	0.303	0.311	0.0022	0.00446	0.00634	0.00129
30	Bayes	0.103	0.203	0.306	0.308	0.0022	0.00469	0.00637	0.00123
40	ML	0.099	0.203	0.303	0.308	0.00149	0.00304	0.00479	9e-04
40	Bayes	0.100	0.203	0.305	0.306	0.0015	0.00304	0.00482	0.00086
50	ML	0.102	0.203	0.302	0.307	0.00127	0.00247	0.00335	0.00068
50	Bayes	0.102	0.203	0.304	0.305	0.00128	0.00248	0.00336	0.00065
60	ML	0.100	0.201	0.301	0.306	0.00099	0.002	0.00291	0.00058
60	Bayes	0.100	0.201	0.302	0.304	0.001	0.002	0.00291	0.00056
70	ML	0.100	0.202	0.300	0.305	0.00081	0.00175	0.0025	0.00045
70	Bayes	0.101	0.202	0.301	0.303	0.00082	0.00175	0.00249	0.00043
80	ML	0.100	0.200	0.302	0.304	0.00071	0.0015	0.00222	0.00039
80	Bayes	0.100	0.200	0.303	0.303	0.00071	0.00151	0.00223	0.00038
90	ML	0.101	0.201	0.301	0.303	0.00069	0.00133	0.00191	0.00033
90	Bayes	0.101	0.202	0.302	0.302	0.00069	0.00134	0.00189	0.00033
100	ML	0.101	0.201	0.3	0.304	0.00057	0.0012	0.00168	3e-04
100	Bayes	0.101	0.202	0.301	0.303	0.00057	0.00121	0.00168	3e-04

Table 2: The average of MLEs and Bayes estimates and corresponding MSE.

Case 3: $\alpha_1 = 0.1, \alpha_2 = 0.2, \alpha_3 = 0.3, \beta = 1.2$									
		Average point estimate				MSE			
n	Method	α_1	α_2	α_3	β	α_1	α_2	α_3	β
20	ML	0.103	0.206	0.301	1.27	0.00334	0.00693	0.00902	0.03446
20	Bayes	0.103	0.207	0.306	1.254	0.00332	0.00746	0.00953	0.03292
30	ML	0.103	0.203	0.304	1.236	0.00229	0.00432	0.00596	0.01904
30	Bayes	0.103	0.203	0.307	1.225	0.0023	0.00431	0.006	0.01813
40	ML	0.099	0.205	0.303	1.235	0.0015	0.00332	0.00461	0.01521
40	Bayes	0.100	0.205	0.305	1.227	0.00151	0.00332	0.00462	0.01447
50	ML	0.100	0.204	0.302	1.226	0.00118	0.00253	0.00347	0.01064
50	Bayes	0.101	0.204	0.303	1.218	0.00118	0.00253	0.00345	0.01023
60	ML	0.102	0.201	0.299	1.217	0.00102	0.00203	0.00282	0.00887
60	Bayes	0.102	0.201	0.301	1.211	0.00102	0.00203	0.00281	0.00862
70	ML	0.100	0.202	0.301	1.223	0.00083	0.00182	0.00235	0.00772
70	Bayes	0.1	0.202	0.302	1.218	0.00083	0.00182	0.00236	0.00749
80	ML	0.100	0.201	0.300	1.218	0.00071	0.00158	0.00211	0.00644
80	Bayes	0.1	0.201	0.301	1.213	0.00071	0.00158	0.0021	0.00628
90	ML	0.100	0.201	0.3	1.216	6e-04	0.00135	0.0018	0.00573
90	Bayes	0.100	0.201	0.301	1.212	6e-04	0.00136	0.00181	0.00559
100	ML	0.100	0.201	0.300	1.214	0.00058	0.00127	0.00169	0.00496
100	Bayes	0.100	0.202	0.301	1.210	0.00059	0.00128	0.00168	0.00485
Case 4: $\alpha_1 = 1.1, \alpha_2 = 1.2, \alpha_3 = 1.3, \beta = 1.2$									
20	ML	1.270	1.405	1.473	1.289	0.38411	0.45768	0.36021	0.05517
20	Bayes	1.271	1.408	1.477	1.282	0.39984	0.47789	0.37735	0.05348
30	ML	1.208	1.310	1.407	1.257	0.20021	0.22621	0.18342	0.03244
30	Bayes	1.206	1.309	1.408	1.253	0.20285	0.23	0.18529	0.03189
40	ML	1.181	1.279	1.381	1.242	0.12869	0.147	0.1291	0.02234
40	Bayes	1.179	1.278	1.381	1.239	0.12964	0.14824	0.13008	0.02199
50	ML	1.158	1.269	1.362	1.233	0.09598	0.10971	0.09532	0.018
50	Bayes	1.157	1.268	1.362	1.231	0.09646	0.11026	0.09559	0.01775
60	ML	1.143	1.25	1.354	1.226	0.07512	0.0885	0.07621	0.01369
60	Bayes	1.142	1.249	1.354	1.224	0.07561	0.08918	0.07663	0.01357
70	ML	1.149	1.251	1.347	1.228	0.0703	0.07662	0.06511	0.01183
70	Bayes	1.148	1.250	1.347	1.227	0.07069	0.07696	0.06554	0.01176
80	ML	1.136	1.244	1.336	1.221	0.05488	0.06522	0.05712	0.01026
80	Bayes	1.135	1.243	1.336	1.219	0.05503	0.06534	0.05711	0.01017
90	ML	1.138	1.243	1.332	1.218	0.04698	0.0604	0.05032	0.00842
90	Bayes	1.137	1.242	1.332	1.216	0.04705	0.06061	0.05044	0.00839
100	ML	1.135	1.234	1.325	1.22	0.04565	0.05195	0.04126	0.00823
100	Bayes	1.134	1.233	1.325	1.218	0.04588	0.05224	0.04139	0.00817

Model	MLEs	$\hat{\mathcal{L}}$	AIC
BVGE	$\hat{\alpha}_1 = 1.351, \hat{\alpha}_2 = 0.465, \hat{\alpha}_3 = 1.153,$ $\hat{\beta} = 0.039$	-296.935	601.870
BGLFR	$\hat{\alpha}_1 = 0.492, \hat{\alpha}_2 = 0.166, \hat{\alpha}_3 = 0.411,$ $\hat{\beta} = 2.013 \times 10^{-4}, \hat{\gamma} = 8.051 \times 10^{-4}$	-293.379	596.757
BCD	$\hat{\alpha}_1 = 2.817 \times 10^{-3}, \hat{\alpha}_2 = 6.298 \times 10^{-3},$ $\hat{\alpha}_3 = 6.006 \times 10^{-3}, \hat{\beta} = 0.4035$	-288.2341	584.4681

Table 3: The MLEs and the values of $\hat{\mathcal{L}}$ and AIC.

Statistic	α_1	α_2	α_3	β
Min.	9.314e-05	0.00081	0.00060	0.3300
1st Qu.	1.056e-03	0.00286	0.00273	0.4093
Median	1.601e-03	0.00401	0.00389	0.4213
Mean	1.838e-03	0.00441	0.00427	0.4208
3rd Qu.	2.343e-03	0.00557	0.00530	0.4326
Max.	1.600e-02	0.01791	0.02665	0.4779
Min.	0.0003374	0.00143	0.001268	0.3919
1st Qu.	0.0013286	0.00338	0.003277	0.4134
Median	0.0016916	0.00415	0.004002	0.4202
Mean	0.0017998	0.00425	0.004158	0.4200
3rd Qu.	0.0021995	0.00499	0.004870	0.4269
Max.	0.0052505	0.00911	0.010205	0.4515

Table 4: Summary statistics for the random draws from the posterior distribution using the two prior scenarios.

Table 6: Diabetic Retinopathy Data.

i	1	2	3	4	5	6	7	8	9	10	11	12
t_i	266	91	154	285	583	547	79	622	707	469	93	1313
δ_i	1	2	2	0	1	2	1	0	2	2	1	2
i	13	14	15	16	17	18	19	20	21	22	23	24
t_i	805	344	790	125	777	306	415	307	637	577	178	517
δ_i	1	1	2	2	2	1	1	2	2	2	1	2
i	25	26	27	28	29	30	31	32	33	34	35	36
t_i	272	1137	1484	315	287	1252	717	642	141	407	356	1653
δ_i	0	0	1	1	2	1	2	1	2	1	1	0
i	37	38	39	40	41	42	43	44	45	46	47	48
t_i	427	699	36	667	588	471	126	350	350	663	567	966
δ_i	2	1	2	1	2	0	1	2	1	0	2	0
i	49	50	51	52	53	54	55	56	57	58	59	60
t_i	203	84	392	1140	901	1247	448	904	276	520	485	248
δ_i	0	1	1	2	1	0	2	2	1	1	2	2
i	61	62	63	64	65	66	67	68	69	70	71	
t_i	503	423	285	315	727	210	409	584	355	1302	227	
δ_i	1	2	2	2	2	2	2	1	1	1	2	