Bayes Estimation for the Marshall-Olkin Bivariate Weibull Distribution†

Debasis Kundu[1] & Arjun K. Gupta[2]

Abstract

In this paper, we consider the Bayesian analysis of the Marshall-Olkin bivariate Weibull distribution. It is a singular distribution whose marginals are Weibull distributions. This is a generalization of the Marshall-Olkin bivariate exponential distribution. It is well known that the maximum likelihood estimators of the unknown parameters do not always exist. The Bayes estimators are obtained with respect to the squared error loss function and the prior distributions allow for prior dependence among the components of the parameter vector. If the shape parameter is known, the Bayes estimators of the unknown parameters can be obtained in explicit forms under the assumptions of independent priors. If the shape parameter is unknown, the Bayes estimators cannot be obtained in explicit forms. We propose to use importance sampling method to compute the Bayes estimators and also to construct associated credible intervals of the unknown parameters. The analysis of one data set is performed for illustrative purposes. Finally we indicate the analysis of data sets obtained from series and parallel systems.

Key Words and Phrases: Bivariate exponential model; maximum likelihood estimators; Importance sampling; Prior distribution; Posterior analysis; Credible intervals.

AMS 2000 Subject Classification: Primary 62F10; Secondary: 62H10

[1]Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Kanpur, Pin 208016, India. e-mail: kundu@iitk.ac.in.

[2]Department of Mathematics and Statistics, Bowling Green State University, Bowling Green, OH 43403 - 0221, USA.

† The paper is dedicated to Professor C.R. Rao on his 90-th birthday.

1 Introduction

Exponential distribution has been used extensively for analyzing the univariate lifetime data, mainly due to its analytical tractability. A huge amount of work on exponential distribution has been found in the statistical literature. Several books and book chapters have been written exclusively on exponential distribution, see for example Balakrishnan and Basu (1995), Johnson, Kotz and Balakrishnan (1995) etc. A variety of bivariate (multivariate) extensions of the exponential distribution also have been considered in the literature. These include the distributions of Gumbel (1960), Freund (1961), Henrich and Jensen (1995), Marshall and Olkin (1967), Downton (1970) as well as Block and Basu (1974), see for example Kotz, Balakrishnan and Johnson (2000).

Among several bivariate (multivariate) exponential distributions, Marshall-Olkin bivariate exponential (MOBE), see Marshall and Olkin (1967), has received the maximum attention. MOBE distribution is the only bivariate exponential distribution with exponential marginals and it also has the bivariate lack of memory property. It has a nice physical interpretation based on random shocks. Extensive work has been done in developing the inference procedure of the MOBE model and its characterization. Kotz, Balakrishnan and Johnson (2000) provided an excellent review on this distribution till that time, see also Karlis (2003) and Kundu and Dey (2009) in this respect.

MOBE distribution is a singular distribution, and due to this reason, it has been used quite successfully when there are ties in the data set. The marginals of the MOBE distribution are exponential distributions, and definitely that is one of the major limitations of the MOBE distribution. Since the marginals are exponential distributions, if it is known (observed) that the estimated probability density functions (PDFs) of the marginals are not decreasing functions or the hazard functions are not constant, then MOBE cannot be used.

Marshall and Olkin (1967) proposed a more flexible bivariate model, namely Marshall-Olkin bivariate Weibull (MOBW) model, where the marginals are Weibull models. Therefore, if it is observed empirically that the marginals are decreasing or unimodal PDFs and monotone hazard functions, then MOBW models can be used quite successfully. Further, MOBW can also be given a shock model interpretation.

Although, extensive work has been done on the MOBE model, MOBW model has not received much attention primarily due to the analytical intractability of the model, see for example Kotz, Balakrishnan and Johnson (2000). Lu (1992) considered the MOBW model and proposed the Bayes estimators of the unknown parameters. Recently, Kundu and Dey (2009) proposed an efficient estimation procedure to compute the maximum likelihood estimators (MLEs) using expectation maximization (EM) algorithm, which extends the EM algorithm proposed by Karlis (2003) to find the MLEs of the MOBE model. Although, the EM algorithm proposed by Kundu and Dey (2009) works very well even for small sample sizes, but it is well known that the MLEs do not always exist. Therefore, in those cases the method cannot be used.

The main aim of this paper is to develop Bayesian inference for the MOBW model. We want to compute the Bayes estimators and the associated credible intervals under proper priors. The paper is closely related to the paper by Pena and Gupta (1990). Pena and Gupta (1990) obtained the Bayes estimators of the unknown parameters of the MOBE model for both series and parallel systems under quadratic loss function. They have used very flexible Dirichlet-Gamma conjugate prior. Depending on the hyper-parameters, the Dirichlet-Gamma prior allows the stochastic dependence and independence among model parameters. Moreover, Jeffrey's non-informative prior also can be obtained as a limiting case. The Bayes estimators can be obtained in explicit forms, and Pena and Gupta (1990) provided a numerical procedure to construct the highest posterior density (HPD) credible intervals of the unknown parameters.

The MOBW model proposed by Marshall and Olkin (1967) has a common shape parameter. If the common shape parameter is known, the same Dirichlet-Gamma prior proposed by Pena and Gupta (1990) can be used as a conjugate prior, but if the common shape parameter is not known, then as expected the conjugate priors do not exist. We propose to use the same conjugate prior for the scale parameters, even when the common shape parameter is unknown. We do not use any specific form of prior on the shape parameter. It is only assumed that the PDF of the prior distribution is log-concave on $(0, \infty)$. It may be assumed that the assumption of log-concave PDF of the prior distribution is not very uncommon, see for example Berger and Sun (1993), Mukhopadhyay and Basu (1997), Patra and Dey (1999) or Kundu (2008). Moreover, many common distribution functions for example normal, log-normal, gamma and Weibull distributions have log-concave PDFs.

Based on the above prior distribution we obtain the joint posterior distribution of the unknown parameters. As expected the Bayes estimators cannot be obtained in closed form. We propose to use the importance sampling procedure to generate samples from the posterior distribution function and in turn use them to compute the Bayes estimators and also to construct the posterior density credible intervals of the unknown parameters. It is observed that the Bayes estimators exist even when the MLEs do not exist. We compare the Bayes estimators and the credible intervals with the MLEs and the corresponding confidence intervals obtained using the asymptotic distribution of the MLEs, when they exist. It is observed that when we have non-informative priors, then the performances of the Bayes estimators and the MLEs are quite comparable, but with informative priors, the Bayes estimators perform better than the MLEs as expected. For illustrative purposes, we have analyzed one data set.

Then we consider the Bayes estimators of the MOBW parameters, when random samples from series systems are available. If the two components are connected in series, and their lifetimes are denoted by X_1 and X_2 respectively, then the random vector observed on system failure is (Z, Δ), where

$$Z = \min\{X_1, X_2\}, \qquad \Delta = \begin{cases} 0 & \text{if} \quad X_1 = X_2 \\ 1 & \text{if} \quad X_1 < X_2 \\ 2 & \text{if} \quad X_1 > X_2. \end{cases} \tag{1}$$

Based on the assumption that (X_1, X_2) has a MOBW distribution, we develop the Bayesian inference of the unknown parameters, when we observe the data of the from (Z, Δ) as described above. It is observed that the Bayes estimators cannot be obtained in explicit forms, and we provide importance sampling procedure to compute the Bayes estimators and also to construct the associated HPD credible intervals.

We further consider the Bayes estimators of the MOBW parameters, when the data are obtained from a parallel system. If the system consists of two components and they are connected in parallel, and if it is assumed that the lifetimes of the two components are X_1 and X_2, then the random vector observed on system failure is (W, Δ), where

$$W = \max\{X_1, X_2\}, \qquad \Delta = \begin{cases} 0 & \text{if} \quad X_1 = X_2 \\ 1 & \text{if} \quad X_1 < X_2 \\ 2 & \text{if} \quad X_1 > X_2. \end{cases} \tag{2}$$

In this case also, based on the assumption that (X_1, X_2) has a MOBW distribution, we develop the Bayesian inference of the unknown parameters based on importance sampling.

Rest of the papers is organized as follows. In Section 2, we briefly describe the MOBW model. The necessary prior assumptions are presented in Section 3. Posterior analysis and Bayesian inference are presented in Section 4. The analysis of a data is presented in Section 5. In Section 6, we consider the Bayesian inference of the unknown parameters, when we observe the data from a series or from a parallel system. Finally we conclude the paper in Section 7.

2 Marshall-Olkin Bivariate Weibull Model

Marshall and Olkin (1967) proposed the MOBW model and it can be described as follows. Let us assume U_1, U_2 and U_0 are three independent random variables, and

$$U_1 \sim \text{WE}(\alpha, \lambda_1), \quad U_2 \sim \text{WE}(\alpha, \lambda_2), \quad U_0 \sim \text{WE}(\alpha, \lambda_0). \tag{3}$$

Here '$\sim$' means follows in distribution and $\text{WE}(\alpha, \lambda)$ means a Weibull distribution with the shape parameter and scale parameter as α and λ respectively, with the probability density function (PDF) as

$$f_{WE}(x; \alpha, \lambda) = \alpha \lambda x^{\alpha-1} e^{-\lambda x^{\alpha}}; \qquad x > 0, \tag{4}$$

and 0 otherwise. For a Weibull distribution with PDF (4), the cumulative distribution function (CDF) and the survival function (SF) will be

$$F_{WE}(x; \alpha, \lambda) = 1 - e^{-\lambda x^{\alpha}}, \quad \text{and} \quad S_{WE}(x; \alpha, \lambda) = e^{-\lambda x^{\alpha}}, \tag{5}$$

respectively.

If we define

$$X_1 = \min\{U_1, U_0\}, \qquad \text{and} \qquad X_2 = \min\{U_2, U_0\}, \tag{6}$$

then (X_1, X_2) is said to have a MOBW distribution with parameters $(\alpha, \lambda_1, \lambda_2, \lambda_0)$, and it will be denoted by $\text{MOBW}(\alpha, \lambda_1, \lambda_2, \lambda_0)$. If $(X_1, X_2) \sim \text{MOBW}(\alpha, \lambda_1, \lambda_2, \lambda_0)$, then the joint survival function of (X_1, X_2) can be written as

$$S_{X_1,X_2}(x_1, x_2) = P(X_1 > x_1, X_2 > x_2) = e^{-\lambda_1 x_1^{\alpha} - \lambda_2 x_2^{\alpha} - \lambda_0 \max\{x_1, x_2\}^{\alpha}}, \tag{7}$$

for $x_1 > 0, x_2 > 0$, and 0 otherwise. Therefore, if we define $z = \max\{x_1, x_2\}$, then (7) can be written as

$$S_{X_1,X_2}(x_1, x_2) \;\; = \;\; P(U_1 > x_1, U_2 > x_2, U_0 > z)$$

$$= S_{WE}(x_1;\alpha,\lambda_1)S_{WE}(x_2;\alpha,\lambda_2)S_{WE}(z;\alpha,\lambda_0)$$

$$= \begin{cases} S_{WE}(x_1;\alpha,\lambda_1)S_{WE}(x_2;\alpha,\lambda_0+\lambda_2) & \text{if} \quad x_1 < x_2 \\ S_{WE}(x_1;\alpha,\lambda_0+\lambda_1)S_{WE}(x_2;\alpha,\lambda_2) & \text{if} \quad x_1 > x_2 \\ S_{WE}(x;\alpha,\lambda_0+\lambda_1+\lambda_2) & \text{if} \quad x_1 = x_2 = x. \end{cases} \tag{8}$$

The joint PDF of (X_1, X_2) can be written as

$$f_{X_1,X_2}(x_1,x_2) = \begin{cases} f_1(x_1,x_2) & \text{if} \quad x_1 < x_2 \\ f_2(x_1,x_2) & \text{if} \quad x_1 > x_2 \\ f_0(x) & \text{if} \quad x_1 = x_2 = x, \end{cases} \tag{9}$$

where

$$\begin{aligned} f_1(x_1,x_2) &= f_{WE}(x_1;\alpha,\lambda_1)f_{WE}(x_2;\alpha,\lambda_0+\lambda_2) \\ f_2(x_1,x_2) &= f_{WE}(x_1;\alpha,\lambda_0+\lambda_1)f_{WE}(x_2;\alpha,\lambda_2) \\ f_0(x) &= \frac{\lambda_0}{\lambda_0+\lambda_1+\lambda_2} f_{WE}(x;\alpha,\lambda_0+\lambda_1+\lambda_2). \end{aligned}$$

The MOBW distribution has both an absolute continuous part and a singular part. The function $f_{X_1,X_2}(\cdot,\cdot)$ may be considered to be a density function for the MOBW distribution if it is understood that the first two terms are densities with respect to two-dimensional Lebesgue measure and the third function with respect to one dimensional Lebesgue measure, see for example Bemis, Bain and Higgins (1972) for a nice discussion on it.

3 Prior Assumptions and Available Data

3.1 Prior Assumptions

When the common shape parameter α is known, we assume the same conjugate prior on $(\lambda_0, \lambda_1, \lambda_2)$ as considered by Pena and Gupta (1990). It is assumed that $\lambda = \lambda_0 + \lambda_1 + \lambda_2$

has a Gamma(a,b) prior, say $\pi_0(\cdot|a,b)$. Here the PDF of Gamma(a,b) for $\lambda > 0$ is

$$\pi_0(\lambda|a,b) = \frac{b^a}{\Gamma(a)}\ \lambda^{a-1}e^{-b\lambda}; \tag{10}$$

and 0 otherwise.

Given λ, $\left(\frac{\lambda_1}{\lambda}, \frac{\lambda_2}{\lambda}\right)$ has a Dirichlet prior, say $\pi_1(\cdot|a_0,a_1,a_2)$, *i.e.*

$$\pi_1\left(\frac{\lambda_1}{\lambda}, \frac{\lambda_2}{\lambda}|\lambda, a_0, a_1, a_2\right) = \frac{\Gamma(a_0+a_1+a_2)}{\Gamma(a_0)\Gamma(a_1)\Gamma(a_2)}\left(\frac{\lambda_0}{\lambda}\right)^{a_0-1}\left(\frac{\lambda_1}{\lambda}\right)^{a_1-1}\left(\frac{\lambda_2}{\lambda}\right)^{a_2-1}, \tag{11}$$

for $\lambda_0 > 0, \lambda_1 > 0, \lambda_2 > 0$, where $\lambda_0 = \lambda - \lambda_1 - \lambda_2$. Here all the hyper parameters a, b, a_0, a_1 and a_2 are greater than 0. For known α it happens to be the conjugate prior also. After simplification, the joint prior of λ_0, λ_1 and λ_2 becomes

$$\begin{aligned}\pi_1(\lambda_0,\lambda_1,\lambda_2|a,b,a_0,a_1,a_2) &= \frac{\Gamma(a_0+a_1+a_2)}{\Gamma(a)}\ (b\lambda)^{a-a_0-a_1-a_2} \times \frac{b^{a_0}}{\Gamma(a_0)}\lambda_0^{a_0-1}e^{-b\lambda_0} \\ &\quad \times \frac{b^{a_1}}{\Gamma(a_1)}\lambda_1^{a_1-1}e^{-b\lambda_1} \times \frac{b^{a_2}}{\Gamma(a_2)}\lambda_2^{a_2-1}e^{-b\lambda_2}. \end{aligned} \tag{12}$$

If $\bar{a} = a_0 + a_1 + a_2$, then

$$\pi_1(\lambda_0,\lambda_1,\lambda_2|a,b,a_0,a_1,a_2) = \frac{\Gamma(\bar{a})}{\Gamma(a)}\ (b\lambda)^{a-\bar{a}} \times \prod_{i=0}^{2}\frac{b^{a_i}}{\Gamma(a_i)}\lambda_i^{a_i-1}e^{-b\lambda_i}. \tag{13}$$

This is a Gamma-Dirichlet distribution with parameters a, b, a_0, a_1, a_2, and will be denoted by GD(a,b,a_0,a_1,a_2). Clearly, in general λ_0, λ_1 and λ_2 will be dependent, but if we take $a = a_0 + a_1 + a_2$, then they will be independent. Therefore, independent priors also can be obtained as a special case of (13). Moreover, the correlation between λ_i and λ_j for $i \neq j$ can be positive or negative. It will be shown that when α is known, the above priors are the conjugate priors.

At this moment we do not assume any specific form of prior on α. It is simply assumed that it has non-negative support on $(0,\infty)$, and the PDF of the prior of α, say $\pi_2(\alpha)$ is log-concave. Moreover, the prior $\pi_2(\alpha)$ is independent of the joint prior on λ_0, λ_1, λ_2, $\pi_1(\lambda_0,\lambda_1,\lambda_2)$. From now on the joint prior of $\alpha, \lambda_0, \lambda_1, \lambda_2$ will be denoted by $\pi(\alpha,\lambda_0,\lambda_1,\lambda_2)$, and clearly $\pi(\alpha, \lambda_0, \lambda_1, \lambda_2) = \pi_2(\alpha)\pi_1(\lambda_0, \lambda_1, \lambda_2)$. It should be mentioned that for the given specific form of the prior $\pi_2(\alpha)$, the choice of the hyper-parameters is also very important for data analysis purposes. We do not address that issue in this paper.

3.2 Available Data

Bivariate Data

In this subsection we mention the kind of data available to us for analysis purposes. It is assumed that we have a bivariate sample of size n, from MOBW$(\alpha, \lambda_0, \lambda_1, \lambda_2)$, and it is as follows:

$$\mathcal{D}_1 = \{(x_{11}, x_{21}), \cdots, (x_{1n}, x_{2n})\}. \tag{14}$$

We will be using the following notations:

$$I_0 = \{i; x_{1i} = x_{2i} = x_i\}, \quad I_1 = \{i; x_{1i} < x_{2i}\} \quad I_2 = \{i; x_{1i} > x_{2i}\}, \quad I = \{1, \cdots, n\},$$

and $|I_0| = n_0$, $|I_1| = n_1$, and $|I_2| = n_2$, here $|I_j|$, for $j = 0, 1, 2$ denote the number of elements in the set I_j. It may be mentioned that if $n_j = 0$, for some $j = 0, 1, 2$, then the MLEs do not exist, see for example Arnold (1968).

Data from a Series System

In this subsection we provide the kind of data available to us from a series system. In this case it is assumed that we have a sample of size n of the form;

$$\mathcal{D}_2 = \{(z_1, \delta_1), \cdots, (z_n, \delta_n)\}. \tag{15}$$

Here (z_i, δ_i) for $i = 1, \cdots n$, are $i.i.d.$ samples from (Z, Δ) as defined in (1). We will be using the following notations:

$$J_0 = \{i; \delta_i = 0\}, \quad J_1 = \{i; \delta_i = 1\} \quad J_2 = \{i; \delta_2 = 2\},$$

and $|J_0| = m_0$, $|J_1| = m_1$, and $|J_2| = m_2$.

Data from a Parallel System

In this section we discuss the data available from a parallel system. It is assumed that in this case we have a sample of size n from a parallel system with two components and the data are coming form (W, Δ) as defined in (2). We have the data of the following form from a parallel system;

$$\mathcal{D}_3 = (w_1, \delta_1), \cdots, (w_n, \delta_n), \tag{16}$$

here (w_i, δ_i) for $i = 1, \cdots, n$ are *i.i.d.* samples from (W, Δ). Here also we denote J_0, J_1, J_2, m_0, m_1 and m_2 same as the series system.

4 Posterior Analysis and Bayesian Inference

In this section we provide the Bayes estimators of the unknown parameters, and the associated credible intervals when the data are of the form (14). We consider two cases separately, namely when (i) the common shape parameter is known, (ii) the common shape parameter is unknown. It is observed that in both the cases the Bayes estimators cannot be obtained in explicit forms in general. We propose to use the importance sampling procedure to compute the Bayes estimators and also to construct the associated credible intervals of the unknown parameters.

Based on the observations, the joint likelihood function of the observed data can be written as

$$l(\mathcal{D}_1|\alpha, \lambda_0, \lambda_1, \lambda_2) = \prod_{i \in I_0} f_0(x_i) \prod_{i \in I_1} f_1(x_{1i}, x_{2i}) \prod_{i \in I_2} f_2(x_{1i}, x_{2i}). \tag{17}$$

The MLEs of the unknown parameters can be obtained by maximizing (17) with respect to the unknown parameters, and as it has been already mentioned that the MLEs do not exist always.

4.1 Common Shape Parameter α is Known

In this case based on the priors $\pi_1(\cdot)$ on $(\lambda_0, \lambda_1, \lambda_2)$ as defined in (13), the posterior density function of $(\lambda_0, \lambda_1, \lambda_2)$ can be written as

$$\begin{aligned} l(\lambda_0, \lambda_1, \lambda_2|\alpha, \mathcal{D}_1) &\propto \lambda^{a-\bar{a}} \sum_{j=0}^{n_1}\sum_{k=0}^{n_2} \binom{n_1}{j}\binom{n_2}{k} \lambda_0^{a_{0jk}-1} e^{-\lambda_0(T_0(\alpha)+b)} \lambda_1^{a_{1k}-1} e^{-\lambda_1(T_1(\alpha)+b)} \\ &\quad \lambda_2^{a_{2j}-1} e^{-\lambda_2(T_2(\alpha)+b)} \\ &\propto \lambda^{a-\bar{a}} \sum_{j=0}^{n_1}\sum_{k=0}^{n_2} w_{jk} \text{ Gamma}(\lambda_0; a_{0jk}, b+T_0(\alpha)) \times \text{Gamma}(\lambda_1; a_{1k}, b+T_1(\alpha)) \\ &\qquad \times \text{ Gamma}(\lambda_2; a_{2j}, b+T_2(\alpha)), \end{aligned} \tag{18}$$

here $a_{0jk} = a_0 + n_0 + j + k$, $a_{1k} = a_1 + n_1 + n_2 - k$, $a_{2j} = a_2 + n_1 + n_2 - j$,

$$T_0(\alpha) = \sum_{i\in I_0} x_i^\alpha + \sum_{i\in I_2} x_{1i}^\alpha + \sum_{i\in I_1} x_{2i}^\alpha, \;\; T_1(\alpha) = \sum_{i\in I_0} x_i^\alpha + \sum_{i\in I_1\cup I_2} x_{1i}^\alpha, \;\; T_2(\alpha) = \sum_{i\in I_0} x_i^\alpha + \sum_{i\in I_1\cup I_2} x_{2i}^\alpha, \tag{19}$$

$$c_{jk} = \binom{n_1}{j}\binom{n_2}{k} \frac{\Gamma(a_{0jk})}{(b+T_0(\alpha))^{a_{0jk}}} \times \frac{\Gamma(a_{1k})}{(b+T_1(\alpha))^{a_{1k}}} \times \frac{\Gamma(a_{2j})}{(b+T_2(\alpha))^{a_{2j}}},$$

and $w_{jk} = \dfrac{c_{jk}}{\left(\sum_{j=0}^{n_1}\sum_{k=0}^{n_2} c_{jk}\right)}$.

Therefore, under the assumption of independence on λ_0, λ_1 and λ_2, *i.e.* when $a = \bar{a}$, it is possible to get the Bayes estimators of λ_0, λ_1 and λ_2 explicitly under the squared error loss function using (18), and they will be as follows:

$$\begin{aligned} \widehat{\lambda}_{0B} &= \frac{1}{b+T_0(\alpha)} \sum_{j=0}^{n_1}\sum_{k=0}^{n_2} w_{jk} a_{0jk}, \\ \widehat{\lambda}_{1B} &= \frac{1}{b+T_1(\alpha)} \sum_{j=0}^{n_1}\sum_{k=0}^{n_2} w_{jk} a_{1k}, \\ \widehat{\lambda}_{2B} &= \frac{1}{b+T_2(\alpha)} \sum_{j=0}^{n_1}\sum_{k=0}^{n_2} w_{jk} a_{2j}. \end{aligned}$$

If $a \neq \bar{a}$, then the Bayes estimators cannot be obtained explicitly. We need some numerical procedure to compute the Bayes estimators. We propose to use the importance sampling procedure, to compute the Bayes estimator of $\theta = \theta(\lambda_0, \lambda_1, \lambda_2)$, any function of λ_0, λ_1 and λ_2, and also to construct the associated HPD credible interval of θ.

Alternatively, the posterior density function of $(\lambda_0, \lambda_1, \lambda_2)$ can be written as

$$
\begin{aligned}
l(\lambda_0, \lambda_1, \lambda_2|\alpha, \mathcal{D}_1) \quad \propto \quad & \lambda^{a-\bar{a}} \times (\lambda_0 + \lambda_2)^{n_1} \times (\lambda_0 + \lambda_1)^{n_2} \times \text{Gamma}(\lambda_0; a_0 + n_0, T_0(\alpha) + b) \\
& \times \text{Gamma}(\lambda_1; a_1 + n_1, T_1(\alpha) + b) \times \text{Gamma}(\lambda_2; a_2 + n_2, T_2(\alpha) + b).
\end{aligned} \tag{20}
$$

Let us denote the right hand side of (20) as $l_N(\lambda_0, \lambda_1, \lambda_2|\alpha, \mathcal{D}_1)$. Therefore, $l_N(\lambda_0, \lambda_1, \lambda_2|\alpha, \mathcal{D}_1)$ and $l(\lambda_0, \lambda_1, \lambda_2|\alpha, \mathcal{D}_1)$ differ only by proportionality constant. The Bayes estimator of θ under the squared error loss function is

$$
\widehat{\theta}_B = \frac{\int_0^\infty \int_0^\infty \int_0^\infty \theta(\lambda_0, \lambda_1, \lambda_2) l_N(\lambda_0, \lambda_1, \lambda_2|\alpha, \mathcal{D}_1) d\lambda_0 d\lambda_1 d\lambda_2}{\int_0^\infty \int_0^\infty \int_0^\infty l_N(\lambda_0, \lambda_1, \lambda_2|\alpha, \mathcal{D}_1) d\lambda_0 d\lambda_1 d\lambda_2}. \tag{21}
$$

It is clear from (21) that to approximate $\widehat{\theta}_B(\lambda_0, \lambda_1, \lambda_2)$, using the importance sampling procedure one needs not compute the normalizing constant. We use the following procedure:

Step 1: Generate

$$
\lambda_0 \sim \text{Gamma}\left(\lambda_0; a_0 + n_0, T_0(\alpha)\right), \tag{22}
$$

$$
\lambda_1 \sim \text{Gamma}\left(\lambda_1; a_1 + n_1, T_1(\alpha)\right), \tag{23}
$$

$$
\lambda_2 \sim \text{Gamma}(\lambda_2; a_2 + n_2, T_2(\alpha)). \tag{24}
$$

Step 2: Repeat this procedure to obtain $\{(\lambda_{0i}, \lambda_{1i}, \lambda_{2i}); i = 1, \cdots N\}$.

Step 3: The approximate value of (21) can be obtained as

$$
\frac{\sum_{i=1}^N \theta_i h(\lambda_{0i}, \lambda_{1i}, \lambda_{2i})}{\sum_{i=1}^N h(\lambda_{0i}, \lambda_{1i}, \lambda_{2i})}, \tag{25}
$$

here $\theta_i = \theta(\lambda_{0i}, \lambda_{1i}, \lambda_{2i})$, and

$$
h(\lambda_0, \lambda_1, \lambda_2) = (\lambda_0 + \lambda_1 + \lambda_2)^{a-\bar{a}} \times (\lambda_0 + \lambda_2)^{n_1} \times (\lambda_0 + \lambda_1)^{n_2}. \tag{26}
$$

The estimator (25) is a consistent estimator of θ. Note that although $h(\cdot)$ is an unbounded function, if λ_0, λ_1 and λ_2 have distributions (22), (23) and (24) respectively, and they are independently distributed, then $E(h^k(\lambda_0, \lambda_1, \lambda_2)) < \infty$ for all $k = 1, 2, \cdots$.

The same method can be used to compute a HPD credible interval of θ, any function of λ_0, λ_1 and λ_2. Suppose for $0 < p < 1$, θ_p is $P[\theta \le \theta_p | \mathcal{D}_1] = p$. Now consider the following function

$$g(\lambda_0, \lambda_1, \lambda_2) = \begin{cases} 1 & \text{if} \quad \theta \le \theta_p \\ \\ 0 & \text{if} \quad \theta > \theta_p. \end{cases} \tag{27}$$

Clearly, $E(g(\lambda_0, \lambda_1, \lambda_2)|\mathcal{D}_1) = p$. Therefore, a consistent estimator of θ_p under the squared error loss function can be obtained from the generated sample $\{(\lambda_{0i}, \lambda_{1i}, \lambda_{2i}); i = 1, \cdots N\}$, as follows. Let

$$w_i = \frac{h(\lambda_{0i}, \lambda_{1i}, \lambda_{2i})}{\sum_{j=1}^{N} h(\lambda_{0j}, \lambda_{1j}, \lambda_{2j})}. \tag{28}$$

Rearrange, $\{(\theta_1, w_1), \cdots, (\theta_N, w_N)\}$ as $\{(\theta_{(1)}, w_{(1)}), \cdots, (\theta_{(N)}, w_{(N)})\}$, where $\theta_{(1)} < \cdots < \theta_{(N)}$, and $w_{(i)}$'s are not ordered, they are just associated with $\theta_{(i)}$. Then a consistent Bayes estimator of θ_p is $\widehat{\theta}_p = \theta_{(N_p)}$, where N_p is the integer satisfying

$$\sum_{i=1}^{N_p} w_{(i)} \le p < \sum_{i=1}^{N_p+1} w_{(i)}. \tag{29}$$

Now using the above procedure, a 100(1-γ)% credible interval of θ can be obtained as $(\widehat{\theta}_\delta, \widehat{\theta}_{\delta+1-\gamma})$, for $\delta = w_{(1)}, w_{(1)} + w_{(2)}, \cdots, \sum_{i=1}^{N_{1-\gamma}} w_{(i)}$. Therefore, a 100(1-$\gamma$)% HPD credible interval of θ becomes $(\widehat{\theta}_{\delta^*}, \widehat{\theta}_{\delta^*+1-\gamma})$, where δ^* satisfies

$$\widehat{\theta}_{\delta^*+1-\gamma} - \widehat{\theta}_{\delta^*} \le \widehat{\theta}_{\delta+1-\gamma} - \widehat{\theta}_\delta, \quad \text{for all} \ \ \delta.$$

4.2 Common Shape Parameter α is Unknown

In this case, the posterior density function of λ_0, λ_1, λ_2 and α can be written as

$$l(\lambda_0, \lambda_1, \lambda_2, \alpha | \mathcal{D}_1) = l(\lambda_0, \lambda_1, \lambda_2 | \alpha, \mathcal{D}_1) \times l(\alpha | \mathcal{D}_1). \tag{30}$$

In this case $l(\lambda_0, \lambda_1, \lambda_2 | \alpha, \mathcal{D}_1)$ can be written as

$$\begin{aligned} l(\lambda_0, \lambda_1, \lambda_2 | \alpha, \mathcal{D}_1) \ \ &\propto \ \ h(\lambda_0, \lambda_1, \lambda_2) \times \text{Gamma}(\lambda_0; a_0 + n_0, T_0(\alpha) + b) \\ &\quad \times \text{Gamma}(\lambda_1; a_1 + n_1, T_1(\alpha) + b) \times \text{Gamma}(\lambda_2; a_2 + n_2, T_2(\alpha) + b) \end{aligned}$$

and $l(\alpha|\mathcal{D}_1)$ can be written as

$$l(\alpha|\mathcal{D}_1) \propto \frac{\pi_2(\alpha) \times \alpha^{n_0+2n_1+2n_2} \left\{\prod_{I_0} x_i^{\alpha-1}\right\} \left\{\prod_{I_1 \cup I_2} x_{1i}^{\alpha-1} x_{2i}^{\alpha-1}\right\}}{(T_0(\alpha)+b)^{a_0+n_0} \times (T_1(\alpha)+b)^{a_1+n_1} \times (T_2(\alpha)+b)^{a_2+n_2}}. \quad (31)$$

Now if we denote

$$\begin{aligned} l_N(\lambda_0, \lambda_1, \lambda_2, \alpha|\mathcal{D}_1) \quad &= \quad h(\lambda_0, \lambda_1, \lambda_2) \times \text{Gamma}(\lambda_0; a_0+n_0, T_0(\alpha)) \\ &\times \text{Gamma}(\lambda_1; a_1+n_1, T_1(\alpha)) \times \text{Gamma}(\lambda_2; a_2+n_2, T_2(\alpha)) \\ &\times \frac{\pi_2(\alpha) \times \alpha^{n_0+2n_1+2n_2} \left\{\prod_{I_0} x_i^{\alpha-1}\right\} \left\{\prod_{I_1 \cup I_2} x_{1i}^{\alpha-1} x_{2i}^{\alpha-1}\right\}}{(T_0(\alpha)+b)^{a_0+n_0} \times (T_1(\alpha)+b)^{a_1+n_1} \times (T_2(\alpha)+b)^{a_2+n_2}}, \quad (32) \end{aligned}$$

the Bayes estimator of $\theta = \theta(\lambda_0, \lambda_1, \lambda_2, \alpha)$ under the squared error loss function is

$$\widehat{\theta}_B = \frac{\int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty \theta(\lambda_0, \lambda_1, \lambda_2, \alpha) l_N(\lambda_0, \lambda_1, \lambda_2, \alpha|\mathcal{D}_1) d\lambda_0 d\lambda_1 d\lambda_2 d\alpha}{\int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty l_N(\lambda_0, \lambda_1, \lambda_2, \alpha|\mathcal{D}_1) d\lambda_0 d\lambda_1 d\lambda_2 d\alpha}. \quad (33)$$

Even if the form of $\pi_2(\alpha)$ is known, it is not possible to compute explicitly (33) in general. In this case we use an importance sampling procedure to provide a consistent estimator of (33). We need the following lemma for further development.

LEMMA 1: If the PDF of $\pi_2(\alpha)$ is log-concave, then $l(\alpha|\mathcal{D}_1)$ is log-concave.

PROOF: The proof can be obtained along the same line as the proof of Theorem 2 of Kundu (2008), and therefore it is avoided. ∎

Now using Lemma 1, we provide an importance sampling method to provide a consistent estimator of (33) and the HPD associated credible as follows.

STEP 1: Generate α_1 from the log-concave density $l(\alpha|\mathcal{D}_1)$ as given in (31) using the method proposed by Devroye (1984). It should be mentioned that to apply Devroye's method one needs to know the mode of the density function.

STEP 2: Generate

$$\lambda_{01}|\alpha, \mathcal{D}_1 \sim \text{Gamma}(\lambda_0; a_0+n_0, T_0(\alpha)+b),$$

$$\lambda_{11}|\alpha, \mathcal{D}_1 \sim \text{Gamma}(\lambda_1; a_1 + n_1, T_1(\alpha) + b),$$

$$\lambda_{21}|\alpha, \mathcal{D}_1 \sim \text{Gamma}(\lambda_2; a_2 + n_2, T_2(\alpha) + b).$$

Step 3: Repeat Step 1 and Step 2, to obtain $\{(\alpha_i, \lambda_{0i}, \lambda_{1i}, \lambda_{2i}); i = 1, \cdots N\}$.

Step 4: A consistent estimator of (33) can be obtained as

$$\frac{\sum_{i=1}^{N} \theta_i h(\lambda_{0i}, \lambda_{1i}, \lambda_{2i})}{\sum_{j=1}^{N} h(\lambda_{0j}, \lambda_{1j}, \lambda_{2j})}, \tag{34}$$

here $\theta_i = \theta(\alpha_i, \lambda_{0i}, \lambda_{1i}, \lambda_{2i})$, and $h(\lambda_0, \lambda_1, \lambda_2)$ is same as defined in (26).

Exactly similar procedure can be used as described in the previous section, to compute the HPD credible interval for θ.

5 Data Analysis

We have analyzed one data set for illustrative purposes. The data set has been obtained from Meintanis (2007) and it is presented in Table 1. It represents the UEFA Champion's League (soccer) data, where at least one goal is scored by the home team and at least one goal scored directly from a penalty kick, foul kick or any other direct kick (all of them together will be called as *kick* goal) by any team have been considered. The soccer is a ninety minutes game. Here, we are interested about the two variables only, X_1 and X_2, where X_1 represents the time in minutes of the first *kick* goal scored by any team and X_2 represents the time of the first goal of any type scored by the home team. In this case all possibilities are open, for example $X_1 < X_2$, or $X_1 > X_2$ or $X_1 = X_2 = X$ (say), and they occur with non-zero probabilities. Therefore, some singular distribution function should be used to analyze this data set.

Before going to analyze the data using MOBW model, first we want to get an idea about the hazard functions of the marginals. We have presented in Figure 1 the scaled TTT

2005-2006	X_1	X_2	2004-2005	X1	X2
Lyon-Real Madrid	26	20	Internazionale-Bremen	34	34
Milan-Fenerbahce	63	18	Real Madrid-Roma	53	39
Chelsea-Anderlecht	19	19	Man. United-Fenerbahce	54	7
Club Brugge-Juventus	66	85	Bayern-Ajax	51	28
Fenerbahce-PSV	40	40	Moscow-PSG	76	64
Internazionale-Rangers	49	49	Barcelona-Shakhtar	64	15
Panathinaikos-Bremen	8	8	Leverkusen-Roma	26	48
Ajax-Arsenal	69	71	Arsenal-Panathinaikos	16	16
Man. United-Benfica	39	39	Dynamo Kyiv-Real Madrid	44	13
Real Madrid-Rosenborg	82	48	Man. United-Sparta	25	14
Villarreal-Benfica	72	72	Bayern-M. TelAviv	55	11
Juventus-Bayern	66	62	Bremen-Internazionale	49	49
Club Brugge-Rapid	25	9	Anderlecht-Valencia	24	24
Olympiacos-Lyon	41	3	Panathinaikos-PSV	44	30
Internazionale-Porto	16	75	Arsenal-Rosenborg	42	3
Schalke-PSV	18	18	Liverpool-Olympiacos	27	47
Barcelona-Bremen	22	14	M. Tel-Aviv-Juventus	28	28
Milan-Schalke	42	42	Bremen-Panathinaikos	2	2
Rapid-Juventus	36	52			

Table 1: UEFA Champion's League data

plot; see Aarset (1987). If a family has a survival function $S(y) = 1 - F(y)$, the scaled TTT transform is $g(u) = H^{-1}(u)/H^{-1}(1)$ for $0 < u < 1$, where $H^{-1}(u) = \int_0^{F^{-1}(u)} S(y)dy$. The corresponding empirical version of the scaled TTT transform is given by $g_n(r/n) = H_n^{-1}(r/n)/H_n^{-1}(1) = \left[\sum_{i=1}^{r} y_{i:n} + (n-r)y_{r:n}\right]$, where $r = 1, \cdots, n$, with $y_{i:n}$, $i = 1, \cdots, n$, being the order statistics of the sample. It has been shown by Aarset (1987) that the scaled TTT transform is convex (concave) if the hazard rate is decreasing (increasing) and for bathtub (unimodal) shaped hazard rate, the scaled TTT transform is first convex (concave) and then concave (convex). In this example, the scaled TTT transform of both X_1 and X_2 presented in Figure 1, shows that in both cases the scaled TTT transforms are concave; we can therefore conclude that both marginals have increasing hazard rates.

We would like to analyze the data using MOBW model also. In this case $n = 37$, n_0

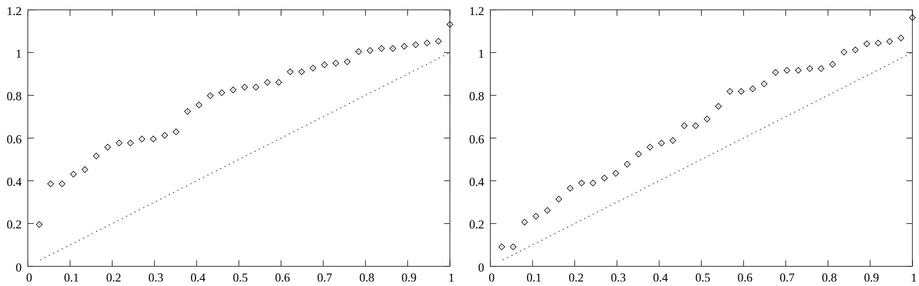

Figure 1: The scaled TTT transform of X_1 and X_2.

$= 14$, $n_1 = 6$, $n_2 = 17$. The sample mean (standard deviation) of X_1 and X_2 are 40.89 (20.14) and 32.86 (22.83) respectively. All the data points have been divided by 100 so that the shape and scale parameters are of the same order. It is not going to make any difference in any statistical inference, although it has helped in the calculation part. We have assumed that the prior distribution of α is Gamma(α; c, d). Since we do not have any prior information, we have taken all the hyper-parameters to be 0.001 (instead of 0.0), as suggested by Congdon (2001). It may be mentioned that for $c = 0.001$, and $d = 0.001$, the PDF of Gamma(α; c, d) is not log-concave, but still the posterior of PDF $l(\alpha|\mathcal{D}_1)$ as defined in (31) is still log-concave. First we have generated α as suggested in Kundu (2008). The histogram of the generated samples and the PDF of $l(\alpha|\mathcal{D}_1)$ are plotted in Figure 2. Clearly, they match very well.

Finally we compute consistent Bayes estimates of α, λ_0, λ_1 and λ_2 with respect to squared error loss functions and they are 1.7049, 2.0750, 0.9566, 3.0368 respectively. The corresponding credible intervals are (1.3280, 1.8920), (1.4852, 2.3571), (0.5045, 1.4816) and (2.1901, 3.5573) respectively.

Now the natural question is how good is this model? Unfortunately there does not exist a general bivariate goodness of fit test like in the univariate situations. We want to see how well the model fits the marginals and to the minimum. Although it is well known that

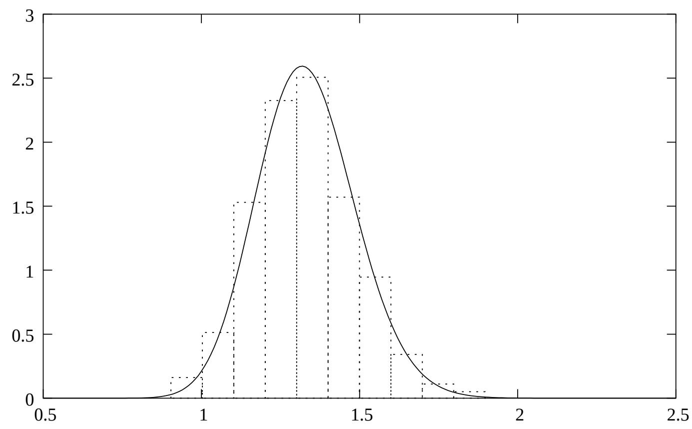

Figure 2: The histogram of the generated samples and the PDF of $l(\alpha|\mathcal{D}_1)$.

they are not sufficient, at least they are necessary. We compute the Kolmogorov-Smirnov statistics to the fitted marginals and to the $\min\{X_1, X_2\}$, they are 0.123, 0.152 and 0.136 the associated p values are 0.631, 0.358 and 0.498, respectively. Therefore, based on the above results, we can conclude that Weibull distribution can be used for the marginals and for the minimum.

For comparison purposes we have reported the maximum likelihood estimates of α, λ_0, λ_1, λ_2 and they are 1.6954, 2.1927, 1.1192, 2.8852 respectively. The associated 95% confidence intervals are as follows; (1.3284, 2.0623), (1.5001, 2.8754), (0.5411, 1.6973), (1.3023, 4.4681) respectively. Interestingly, the Bayes estimate of α and the MLE estimate of α match very well, although the length of the credible interval is smaller than the length of the confidence intervals.

6 Series and Parallel Systems

6.1 Series System

In this section it is assumed that we have a sample of size n, from a two-component series system. We will indicate how to develop the Bayesian inference of the unknown parameters under the assumption that $(X_1, X_2) \sim$ MOBW$(\alpha, \lambda_0, \lambda_1, \lambda_2)$, when X_1 and X_2 denote the lifetime of the Component 1 and Component 2 respectively. It may be mentioned that Pena and Gupta (1990) considered the same problem, under the assumptions that (X_1, X_2) follows MOBE distribution. Although they have obtained the Bayes estimators under the squared error loss function, they have not provided the HPD credible intervals.

Now based on the observations (15), the likelihood function can be written as

$$\begin{aligned} l(\mathcal{D}_2|\alpha, \lambda_0, \lambda_1, \lambda_2) &= \prod_{i \in J_0} \left\{\alpha\lambda_0 z_j^{\alpha-1} e^{-\lambda z_j^{\alpha}}\right\} \prod_{i \in J_1} \left\{\alpha\lambda_1 z_j^{\alpha-1} e^{-\lambda z_j^{\alpha}}\right\} \prod_{i \in J_2} \left\{\alpha\lambda_2 z_j^{\alpha-1} e^{-\lambda z_j^{\alpha}}\right\} \\ &= \alpha^n \lambda_0^{m_0} \lambda_1^{m_1} \lambda_2^{m_2} e^{-\lambda \sum_{i=1}^n z_j^{\alpha}} \prod_{i=1}^n z_j^{\alpha-1}, \end{aligned} \tag{35}$$

here $\lambda = \lambda_0 + \lambda_1 + \lambda_2$. Now based on the prior distribution as described in Section 3, the joint posterior density function of α, λ_0, λ_1 and λ_2 can be written as

$$\begin{aligned} l(\alpha, \lambda_0, \lambda_1, \lambda_2|\mathcal{D}_2) &= \alpha^n \lambda^{a-\bar{a}} \left\{\prod_{i=1}^n z_j^{\alpha-1}\right\} \lambda_0^{m_0+a_0-1} e^{-\lambda_0(b+\sum_{i=1}^n z_i^{\alpha})} \lambda_1^{m_1+a_1-1} e^{-\lambda_1(b+\sum_{i=1}^n z_i^{\alpha})} \\ &\quad \lambda_2^{m_2+a_2-1} e^{-\lambda_2(b+\sum_{i=1}^n z_i^{\alpha})}. \end{aligned} \tag{36}$$

The posterior density function (36) can be re-written as

$$l(\alpha, \lambda_0, \lambda_1, \lambda_2|\mathcal{D}_2) = l(\lambda_0, \lambda_1, \lambda_2|\mathcal{D}_2, \alpha) \times l(\alpha|\mathcal{D}_2), \tag{37}$$

where

$$l(\alpha|\mathcal{D}_2) \propto \frac{\alpha^n \prod_{i=1}^n z_j^{\alpha-1} \pi_2(\alpha)}{\left(b + \sum_{i=1}^n z_j^{\alpha}\right)}, \tag{38}$$

and

$$
\begin{aligned}
l(\lambda_0, \lambda_1, \lambda_2 | \alpha, \mathcal{D}_2) &= \frac{\Gamma(\bar{a} + m_0 + m_1 + m_2)}{\Gamma(a + m_0 + m_1 + m_2)} \times \left[\lambda(b + \sum_{j=1}^{n} z_j^{\alpha}) \right]^{a - \bar{a}} \\
& \quad \times \prod_{i=0}^{2} \text{Gamma}\left(\lambda_i; m_i + a_i, b + \sum_{j=1}^{n} z_j^{\alpha} \right) \\
&\sim \text{GD}(a + m_0 + m_1 + m_2, b + \sum_{j=1}^{n} z_j^{\alpha}, a_0 + m_0, a_1 + m_1, a_2 + m_2).
\end{aligned}
\tag{39}
$$

It is clear from (39) that for known α, using Theorem 2 of Pena and Gupta (1990), the Bayes estimators of λ_0, λ_1 and λ_2 with respect to squared error loss function are respectively

$$
c \times \frac{m_0 + a_0}{b + \sum_{j=1}^{n} z_j^{\alpha}}, \quad c \times \frac{m_1 + a_1}{b + \sum_{j=1}^{n} z_j^{\alpha}}, \quad \text{and} \quad c \times \frac{m_2 + a_2}{b + \sum_{j=1}^{n} z_j^{\alpha}}, \tag{40}
$$

where

$$
c = \frac{a + m_0 + m_1 + m_2}{\bar{a} + m_0 + m_1 + m_2}.
$$

Note that under the independent priors of λ_0, λ_1 and λ_2, *i.e.* when $a = a_1 + a_2 + a_3 = \bar{a}$, $c = 1$. Although for known α, the Bayes estimators can be obtained in explicit forms, the corresponding HPD credible intervals cannot be obtained explicitly. We propose to use the importance sampling procedure to compute the HPD credible intervals of the unknown parameters, the details will be explained later.

Now we consider the more important case, *i.e.*, when α is also unknown. In this case, the Bayes estimator of any function of α, λ_0, λ_1, λ_2 say $g(\alpha, \lambda_0, \lambda_1, \lambda_2)$, with respect to squared error loss function can be obtained as the posterior expectation of $g(\alpha, \lambda_0, \lambda_1, \lambda_2)$, *i.e.*

$$
\widehat{g}_B = E(g(\alpha, \lambda_0, \lambda_1, \lambda_2) | \mathcal{D}_2) = \frac{\int_0^{\infty} \int_0^{\infty} \int_0^{\infty} \int_0^{\infty} g(\alpha, \lambda_0, \lambda_1, \lambda_2) l(\alpha, \lambda_0, \lambda_1, \lambda_2) d\alpha d\lambda_0 d\lambda_1 d\lambda_2}{\int_0^{\infty} \int_0^{\infty} \int_0^{\infty} \int_0^{\infty} l(\alpha, \lambda_0, \lambda_1, \lambda_2) d\alpha d\lambda_0 d\lambda_1 d\lambda_2}. \tag{41}
$$

Clearly, (41) cannot be computed explicitly, in most of the cases even if we know $\pi_2(\alpha)$ explicitly. Although, importance sampling procedure can be used very effectively to provide a consistent estimator of (41), and also to construct HPD credible interval of $g(\alpha, \lambda_0, \lambda_1, \lambda_2)$. We need the following result for developing the importance sampling procedure.

LEMMA 2: If PDF of $\pi_2(\alpha)$ is log-concave, $l(\alpha|\mathcal{D}_2)$ is log-concave.

PROOF: It can be obtained following the same procedure as in Theorem 2 of Kundu (2008), and therefore it is avoided.

Using Lemma 2, and following the same procedure as in Section 4.2, random samples from the joint posterior distribution function can be generated, and they can be used to compute consistent Bayes estimator of $g(\alpha, \lambda_0, \lambda_1, \lambda_2)$ and also to construct the associated HPD credible interval.

6.2 Parallel System

In this section it is assumed that we have a sample of size n from a two components parallel system. It is further assumed that the lifetime distributions of the two components, $(X_1, X_2) \sim \text{MOBW}(\alpha, \lambda_0, \lambda_1, \lambda_2)$. We will consider the Bayesian inference of the unknown parameters, namely $\alpha, \lambda_0, \lambda_1, \lambda_2$. We will provide the Bayes estimators and the associated credible intervals based on the squared error loss function.

Based on the observed data as described in Section 3.2, the likelihood function can be written as

$$\begin{aligned} l(\mathcal{D}_3|\alpha, \lambda_0, \lambda_1, \lambda_2) &= \alpha^n \left\{ \prod_{i=1}^{n} w_i^{\alpha} \right\} \lambda_0^{m_0} (\lambda_0 + \lambda_1)^{m_2} (\lambda_0 + \lambda_2)^{m_1} e^{-\lambda_0 \sum_{i \in J} w_i^{\alpha}} e^{-\lambda_1 \sum_{i \in J_0 \cup J_2} w_i^{\alpha}} \\ & \quad e^{-\lambda_2 \sum_{i \in J_0 \cup J_1} w_i^{\alpha}} \times \prod_{i \in J_1} \left(1 - e^{-\lambda_1 w_i^{\alpha}}\right) \times \prod_{i \in J_2} \left(1 - e^{-\lambda_2 w_i^{\alpha}}\right). \end{aligned}$$

Now based on the prior distribution as described in Section 3, the joint posterior density function of α, λ_0, λ_1 and λ_2 can be written as

$$
\begin{aligned}
l(\alpha, \lambda_0, \lambda_1, \lambda_2 | \mathcal{D}_2) \propto \sum_{i=0}^{m_2} \sum_{k=0}^{m_1} p_{ik} l_{ik}(\alpha | \mathcal{D}_3) \quad & \times \quad \text{gamma}(\lambda_0; i+k+a_0, b + \sum_{j=1}^{n} w_j^{\alpha}) \\
& \times \quad \text{gamma}(\lambda_1; m_2 + a_1 - i, b + \sum_{j \in J_0 \cup J_2} w_j^{\alpha}) \\
& \times \quad \text{gamma}(\lambda_2; m_1 + a_2 - k, b + \sum_{j \in J_0 \cup J_1} w_j^{\alpha}) \\
& \times \quad h(\alpha, \lambda_1, \lambda_2 | \mathcal{D}_3), \qquad (42)
\end{aligned}
$$

here

$$
p_{ik} = \binom{m_2}{i} \binom{m_1}{k} \Gamma(i+k+a_0) \Gamma(m_2 + a_1 - i) \Gamma(m_1 + a_2 - k)
$$

$$
l_{ik}(\alpha | \mathcal{D}_3) = \left\{ \frac{\pi_2(\alpha) \alpha^n \prod_{i=1}^{n} w_i^{\alpha}}{\left(b + \sum_{j=1}^{n} w_j^{\alpha}\right)^{i+k+a_0} \left(b + \sum_{j \in J_0 \cup J_2}^{n} w_j^{\alpha}\right)^{m_2 + a_1 - i} \left(b + \sum_{j \in J_0 \cup J_1}^{n} w_j^{\alpha}\right)^{m_1 + a_2 - k}} \right\}
$$

and

$$
h(\alpha, \lambda_1, \lambda_2 | \mathcal{D}_3) = \prod_{i \in J_1} \left(1 - e^{-\lambda_1 w_i^{\alpha}}\right) \times \prod_{i \in J_2} \left(1 - e^{-\lambda_2 w_i^{\alpha}}\right).
$$

Therefore, the Bayes estimator of any function of α, λ_0, λ_1, λ_2 say $g(\alpha, \lambda_0, \lambda_1, \lambda_2)$, with respect to squared error loss function can be obtained as the posterior expectation of $g(\alpha, \lambda_0, \lambda_1, \lambda_2)$, *i.e.*

$$
\widehat{g}_B = E(g(\alpha, \lambda_0, \lambda_1, \lambda_2) | \mathcal{D}_3) = \frac{\int_0^{\infty} \int_0^{\infty} \int_0^{\infty} \int_0^{\infty} g(\alpha, \lambda_0, \lambda_1, \lambda_2) l(\alpha, \lambda_0, \lambda_1, \lambda_2) d\alpha d\lambda_0 d\lambda_1 d\lambda_2}{\int_0^{\infty} \int_0^{\infty} \int_0^{\infty} \int_0^{\infty} l(\alpha, \lambda_0, \lambda_1, \lambda_2) d\alpha d\lambda_0 d\lambda_1 d\lambda_2}. \qquad (43)
$$

In this case even if α and $\pi_2(\alpha)$ are known, (43) cannot be computed explicitly in most of the cases. We propose to use the importance sampling technique to compute (43). We need the following result for implementing importance sampling.

LEMMA 3: If $\pi_2(\alpha)$ is log-concave, $l_{ik}(\alpha | \mathcal{D}_3)$ is log-concave for all $i = 0, \cdots, m_1$ and $k = 0, \cdots, m_2$.

PROOF: It can be obtained similarly as the proof of Lemma 2. ∎

Now using Lemma 3, and using standard technique to generate samples from a mixture distribution, see Ross (2006), samples from the mixture distribution proportional to

$$\begin{aligned}\sum_{i=0}^{m_2}\sum_{k=0}^{m_1} p_{ik} l_{ik}(\alpha|\mathcal{D}_3) \quad &\times \quad \text{gamma}(\lambda_0; i+k+a_0, b+\sum_{j=1}^{n} w_j^{\alpha}) \\ &\times \quad \text{gamma}(\lambda_1; m_2+a_1-i, b+\sum_{j\in J_0\cup J_2} w_j^{\alpha}) \\ &\times \quad \text{gamma}(\lambda_2; m_1+a_2-k, b+\sum_{j\in J_0\cup J_1} w_j^{\alpha}),\end{aligned}$$

can be easily generated. These samples can be used to compute consistent Bayes estimator of $g(\alpha, \lambda_0, \lambda_1, \lambda_2)$ and also to construct the associated credible interval following the same procedure as it has been described in Section 4.2.

7 Conclusions

In this paper we have considered the Bayesian analysis of the unknown parameters of the Marshall-Olkin bivariate Weibull model. Marshall-Olkin Bivariate Weibull distribution is a singular distribution similarly as the Marshall-Olkin bivariate exponential model. If there are ties in the data, MOBW model can be used quite effectively to analyze the data set. We have used dependent prior on the scale parameters as suggested by Pena and Gupta (1990) in case of MOBE model, and we have taken independent prior on the shape parameter. The Bayes estimators cannot be obtained in explicit forms, and we have proposed to use the importance sampling technique to compute the Bayes estimators and also compute the associated HPD credible intervals. We have performed the analysis of a data set mainly for illustrative purposes, and it is observed that the proposed method works quite well.

Acknowledgements

The authors are thankful to two reviewers and the associate editor for their constructive comments which helped us to improve the earlier versions of the manuscript.

References

[1] Aarset, M.V. (1987), "How to identify a bathtub hazard rate?", *IEEE Transactions on Reliability*, vol. 36, 106 -108.

[2] Arnold, B. (1968), "Parameter estimation for a multivariate exponential distribution", *Journal of the American Statistical Association*, vol. 63, 848–852.

[3] Balakrishnan, N. and Basu, A.P. (1995), *Exponential Distribution: Theory, Methods and Applications*, John Wiley, New York.

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

[5] Berger, J.O. and Sun, D. (1993), "Bayesian analysis for Poly-Weibull distribution", *Journal of the American Statistical Association*, vol. 88, 1412 - 1418.

[6] Block, H. and Basu, A. P. (1974), "A continuous bivariate exponential extension", *Journal of the American Statistical Association*, vol. 69, 1031–1037.

[7] Congdon, P. (2001), *Bayesian Statistical Modeling*, John Wiley and Sons, West Sussex, England.

[8] Devroye, L. (1984), "A simple algorithm for generating random variables with log-concave density", *Computing*, vol. 33, 247 - 257.

[9] Downton, F. (1970), "Bivariate exponential distributions in reliability theory", *Journal of the Royal Statistical Society*, B, vol. 32, 508 - 417.

[10] Freund, J.E. (1961), "A bivariate extension of the exponential distribution", *Journal of the American Statistical Association*, vol. 56, 971 - 977.

[11] Gumbel, E.J. (1960), "Bivariate exponential distribution", *Journal of the American Statistical Association*, vol. 55, 698 - 707.

[12] Henrich, G. and Jensen, U. (1995), "Parameter estimation for a bivariate lifetime distribution in reliability with multivariate extension", *Metrika*, vol. 42, 49 - 65.

[13] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1995), *Continuous Univariate Distribution*, Vol. 1, John Wiley and Sons, New York.

[14] Karlis, D. (2003), "ML estimation for multivariate shock models via an EM algorithm", *Annals of the Institute of Statistical Mathematics*, 55, 817–830.

[15] Kotz, S., Balakrishnan, N. and Johnson, N.L. (2000), *Continuous multivariate distributions, models and applications*, John Wiley and Sons, New York.

[16] Kundu, D. (2008), "Bayesian inference and life testing plan for Weibull distribution in presence of progressive censoring", *Technometrics*, vol. 50, 144 - 154.

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

[18] Kundu, D. and Dey, A.K. (2009), "Estimating the Parameters of the Marshall Olkin Bivariate Weibull Distribution by EM Algorithm", *Computational Statistics and Data Analysis*, vol. 53, 956 - 965, 2009.

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

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

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

[22] Mukhopadhyay, C. and Basu, A.P. (1997), "Bayesian analysis of incomplete time and cause of failure time", *Journal of Statistical Planning and Inference*, vol. 59, 79 - 100.

[23] Patra, K. and Dey, D.K. (1999), "A multivariate mixture of Weibull distributions in reliability modeling", *Statistics and Probability Letters*, vol. 45, 225–235.

[24] Pena, A. and Gupta, A.K. (1990), "Bayes estimation for the Marshall-Olkin exponential distribution", *Journal of the Royal Statistical Society*, Ser B, vol. 52, 379-389.

[25] Ross, S.M. (2006), *Simulation*, 4-th edition, Elsevier, New York.