GU J Sci 31(2): 643-658 (2018)

Gazi University

Journal of Science

http://dergipark.gov.tr/gujs

A Mixture Model of Two Bivariate Weibull Distributions

Ayça Hatice TÜRKAN[1,*], Nazif ÇALIŞ[2]

[1]Afyon Kocatepe University, Department of Statistics, 03200, Afyonkarahisar, Turkey

[2]Adıyaman University, Department of Management, 02040, Adıyaman, Turkey

Article Info

Received: 27/10/2017
Accepted: 26/02/2018

Keywords

Block-Basu bivariate Weibull (BBBW)
Marshall-Olkin bivariate Weibull (MOBW)
Maximum Likelihood Estimation
Mixture distribution model

Abstract

In this paper, we propose a mixture model containing bivariate Weibull distributions—the Marshall-Olkin bivariate Weibull (MOBW) and the Block-Basu bivariate Weibull (BBBW)—because each of these distributions alone is inadequate for explaining a data set when certain special situations occur in bivariate lifetime data sets. We refer to the proposed model as Mix_BW. To estimate the model parameters, we use the expectation-maximisation (EM) algorithm in an adapted form we term the Mix_EM algorithm. We provide illustrative examples with real and simulated data sets to demonstrate the applicability of the proposed Mix_BW model.

1. INTRODUCTION

In lifetime and reliability data analyses, due to their flexibility and definitions for continuous positive random variables, the exponential, Weibull and gamma distributions are widely used to model univariate homogeneous data. Mixtures of these distributions are also used to model univariate heterogeneous data [4]. However, each of these distributions by themselves is inadequate for multivariate lifetime or reliability data. Moreover, there has been insufficient research with respect to modelling multivariate lifetime or reliability data. In recent decades, new bivariate exponential models have been developed in response to the lack of effective models related to the structure of bivariate lifetime data. To model bivariate lifetime data, researchers Gumbel [7], Freund [6], and Marshall and Olkin [12] proposed bivariate exponential distributions. In subsequent studies, Block and Basu [1] obtained bivariate exponential distributions from the Marshall-Olkin bivariate exponential (MOBE) distribution by removing the singular aspects and retaining only the absolutely continuous aspects. Diawara and Carpenter [3] proposed a mixture of bivariate exponential distributions, investigated the properties of the associated parameters and predicted the mixture elements. Weibull models are more flexible than exponential models; so many studies have been conducted using Weibull models. Lu [11] proposed bivariate Weibull extensions of the MOBE distribution. Various bivariate Weibull models were examined by Han [8], who also proposed a location-scale bivariate Weibull model for the lifetime modelling. Kundu and Dey [9] considered the Marshall-Olkin bivariate Weibull distribution and discussed the application of the EM algorithm for computing maximum likelihood estimators. Kundu and Gupta [10] extended the BBBE model to the Weibull model known as the Block-Basu bivariate Weibull and derived the EM algorithm for computing the maximum likelihood estimators of the unknown parameters.

*Corresponding author, e-mail: aturkan@aku.edu.tr

The MOBW and BBBW distributions can be used on bivariate lifetime data, which occurs in many fields, including medicine, biology, engineering and demography. However, these distributions are inadequate when the data are heterogeneous. For heterogeneous data sets, mixture distribution models are appropriate tools for modelling a wide variety of random phenomena. With this motivation, Çalış et al. [2] considered a mixture of MOBW distributions and calculated maximum likelihood estimators for the mixture using the EM algorithm. However, the mixture of MOBW distributions is only appropriate for the data sets including identical observations in each component. When there are heterogeneous data sets with two or more components, one of the components may have identical observations whereas the other may not. In this case, the mixtures of MOBW or BBBW distributions do not yield appropriate distributions, so mixtures of MOBW and BBBW distributions may be a solution to this problem. In the present paper, we propose a mixture model of bivariate Weibull distributions, including MOBW and BBBW distributions. We refer to the proposed mixture distribution as Mix_BW. The remainder of this paper is organised as follow: In Section 2, we provide detailed information about the proposed Mix_BW. In Section 3, we perform parameter estimations of the Mix_BW distribution and present the steps of the EM algorithm. We demonstrate the applicability of the proposed Mix_BW distribution on sample data sets with respect to different situations, as well as a real data set, in Section 4. We conduct a simulation study to demonstrate our results in Section 5 and we draw our conclusions in Section 6.

2. MIXTURE OF BIVARIATE WEIBULL DISTRIBUTIONS

To flexibly model data, mixture distribution models are used in situations where a single distribution is insufficient or there is evidence of multimodality [5]. In the current study, we propose the Mix_BW distribution to model non-homogeneous bivariate lifetime data.

The univariate Weibull distribution with the shape parameter $\alpha\rangle 0$ and the scale parameter $\theta\rangle 0$ has the following probability density function (pdf) [13]:

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

A Weibull distribution with the pdf given in Eq.(1) is denoted by $WE(\alpha,\theta)$. Suppose U_0, U_1 and U_2, respectively, are independent $WE(\alpha,\lambda_0)$, $WE(\alpha,\lambda_1)$, and $WE(\alpha,\lambda_2)$ random variables. Define $X_1 = \min\{U_0,U_1\}$ and $X_2 = \min\{U_0,U_2\}$. Then, the bivariate vector (X_1,X_2) has the MOBW distribution with the parameters α, λ_0, λ_1, λ_2. The joint pdf of X_1 and X_2 can be written as follows:

$$f_{MOBW}(x_1,x_2) = \begin{cases} f_{WE}(x_1;\alpha,\lambda_1)f_{WE}(x_2;\alpha,\lambda_0+\lambda_2) & 0\langle x_1\langle x_2\langle\infty \\ f_{WE}(x_1;\alpha,\lambda_0+\lambda_1)f_{WE}(x_2;\alpha,\lambda_2) & 0\langle x_2\langle x_1\langle\infty \\ \dfrac{\lambda_0}{\lambda_0+\lambda_1+\lambda_2} f_{WE}(x;\alpha,\lambda_0+\lambda_1+\lambda_2) & 0\langle x_1 = x_2 = x\langle\infty \end{cases} \tag{2}$$

A MOBW distribution with the pdf given in Eq.(2) is denoted as $MOBW(\alpha,\lambda_0,\lambda_1,\lambda_2)$ [9], and the BBBW distribution can be obtained from the MOBW distribution by removing the singular aspect and keeping only the continuous aspects. Then the joint pdf of the BBBW distribution can be written as follows:

$$f_{BBBW}(x_1,x_2) = \begin{cases} cf_{WE}(x_1;\alpha,\lambda_1)f_{WE}(x_2;\alpha,\lambda_0+\lambda_2) & 0\langle x_1\langle x_2\langle\infty \\ cf_{WE}(x_1;\alpha,\lambda_0+\lambda_1)f_{WE}(x_2;\alpha,\lambda_2) & 0\langle x_2\langle x_1\langle\infty \end{cases} \tag{3}$$

where c is a normalising constant and $c = \dfrac{\lambda_0+\lambda_1+\lambda_2}{\lambda_1+\lambda_2}$. The pdf given in Eq.(3) is denoted as $BBBW(\alpha,\lambda_0,\lambda_1,\lambda_2)$.

In finite mixture models, it is assumed that the population consists of $g\ (\geq 2)$ distinct subgroups or subclasses. Furthermore, a finite mixture density function can be written as follows:

$$f(x \mid \psi) = \sum_{k=1}^{g} \pi_k f_k(x \mid \theta_k) \tag{4}$$

where the vector $\psi = (\pi, \theta)$ contains all unknown parameters $\pi = (\pi_1, \ldots, \pi_g)$ and $\theta = (\theta_1, \ldots, \theta_g)$. The function $f_k(x \mid \theta_k)$ is called the mixture component density function for the θ_k and π_k parameters, where π_k is the mixture weight of the k th class in which $\pi_k \in (0,1)$ and $\sum_{k=1}^{g} \pi_k = 1$. Thus, the pdf of the Mix_BW model $f_{Mix\_BW}(x_1, x_2)$ can be written in the following form:

$$f_{Mix\_BW}(x_1, x_2) = \sum_{j=1}^{g} \pi_j f_{BW(j)}\left(x_{1(j)}, x_{2(j)}\right) \tag{5}$$

where $f_{BW(j)}\left(x_{1(j)}, x_{2(j)}\right)$ are component densities and π_j are mixing proportions or weights, which are nonnegative quantities that sum to one. Here we refer to the density given in Eq.(5) as a g-component finite mixture density. When this density function is reorganised as a mixture of MOBW and BBBW distributions, the model for the population with two subgroups can be written as follows:

$$f_{Mix\_BW}(x_1, x_2) = \pi\, f_{MOBW}(x_1, x_2) + (1-\pi)\, f_{BBBW}(x_1, x_2) \tag{6}$$

in which the f_{MOBW} and f_{BBBW} are given, respectively, as follows:

$$f_{MOBW}(x_1, x_2) = \begin{cases} f_{WE}\left(x_1; \alpha_{MO}, \lambda_{1(MO)}\right) f_{WE}\left(x_2; \alpha_{MO}, \lambda_{0(MO)} + \lambda_{2(MO)}\right) & 0 \langle x_1 \langle x_2 \langle \infty \\ f_{WE}\left(x_1; \alpha_{MO}, \lambda_{0(MO)} + \lambda_{1(MO)}\right) f_{WE}\left(x_2; \alpha_{MO}, \lambda_{2(MO)}\right) & 0 \langle x_2 \langle x_1 \langle \infty \\ \dfrac{\lambda_{0(MO)}}{\lambda_{0(MO)} + \lambda_{1(MO)} + \lambda_{2(MO)}} f_{WE}\left(x; \alpha_{MO}, \lambda_{0(MO)} + \lambda_{1(MO)} + \lambda_{2(MO)}\right) & 0 \langle x_1 = x_2 = x \langle \infty \end{cases} \tag{7}$$

and

$$f_{BBBW}(x_1, x_2) = \begin{cases} c f_{WE}\left(x_1; \alpha_{BB}, \lambda_{1(BB)}\right) f_{WE}\left(x_2; \alpha_{BB}, \lambda_{0(BB)} + \lambda_{2(BB)}\right) & 0 \langle x_1 \langle x_2 \langle \infty \\ c f_{WE}\left(x_1; \alpha_{BB}, \lambda_{0(BB)} + \lambda_{1(BB)}\right) f_{WE}\left(x_2; \alpha_{BB}, \lambda_{2(BB)}\right) & 0 \langle x_2 \langle x_1 \langle \infty \end{cases} \tag{8}$$

In Eqs.(7) and (8), *MO* and *BB* denote Marshall-Olkin and Block-Basu, respectively.

3. EM ALGORITHM FOR MIXTURE OF BIVARIATE WEIBULL DISTRIBUTIONS

In this section, we present the Mix_EM algorithm, comprising EM algorithms for the MOBW [9] and BBBW [10] distributions for computing the maximum likelihood estimators of the unknown parameters of the Mix_BW model. We define the bivariate data set as $\{(x_{11}, x_{21}), \ldots, (x_{1n}, x_{2n})\}$, which is clustered using the k-means algorithm. We individually fit the BW distributions to each cluster using the EM algorithm. After applying the k-means algorithm, we use the EM algorithm for MOBW, as given in [9], for the sub-cluster containing observations of equal value. For the sub-cluster containing no observations of equal value, we use the EM algorithm for the BBBW distribution given in [10]. We then appropriately recluster the data using the weighted pdfs of the BW distributions. We provide the EM algorithms for the MOBW and BBBW distributions in the following sections.

3.1. EM for the MOBW

To obtain the maximum likelihood estimators of the MOBW component, we group the data in the sub-cluster containing observations of equal value, as in the study of Kundu and Dey [9], using following notations:

$$I_{0(MO)} = \{i : x_{1i(MO)} = x_{2i(MO)} = x_{i(MO)}\}$$
$$I_{1(MO)} = \{i : x_{1i(MO)} \langle x_{2i(MO)}\}$$
$$I_{2(MO)} = \{i : x_{1i(MO)} \rangle x_{2i(MO)}\}$$
$$I_{3(MO)} = I_{1(MO)} \cup I_{2(MO)}$$
$$I_{(MO)} = I_{0(MO)} \cup I_{1(MO)} \cup I_{2(MO)}$$

where $|I_{0(MO)}| = n_{0(MO)}$, $|I_{1(MO)}| = n_{1(MO)}$, $|I_{2(MO)}| = n_{2(MO)}$ and $I_{(MO)} = I_{0(MO)} \cup I_{1(MO)} \cup I_{2(MO)}$. $|I_{l(MO)}|$ for $l = 0,1,2$ denotes the number of elements in the set $I_{l(MO)}$. We individually fit the Weibull distribution to each of the $X_{1(MO)}$, $X_{2(MO)}$ and $\min\{X_{1(MO)}, X_{2(MO)}\}$ values. We take the average value of the shape parameters of the fitted distributions as our initial guess for $\alpha_{(MO)}$. Our initial guess values for $\lambda_{0(MO)}$, $\lambda_{1(MO)}$ and $\lambda_{2(MO)}$ are 1.0, 1.0 and 1.0, respectively. We update the parameters $\lambda_{0(MO)}$, $\lambda_{1(MO)}$, and $\lambda_{2(MO)}$ using the notations given below:

$$u_{1(MO)} = \frac{\lambda_{0(MO)}}{\lambda_{0(MO)} + \lambda_{2(MO)}}, \qquad u_{2(MO)} = \frac{\lambda_{2(MO)}}{\lambda_{0(MO)} + \lambda_{2(MO)}},$$
$$v_{1(MO)} = \frac{\lambda_{0(MO)}}{\lambda_{0(MO)} + \lambda_{1(MO)}}, \qquad v_{2(MO)} = \frac{\lambda_{1(MO)}}{\lambda_{0(MO)} + \lambda_{1(MO)}}$$

Equations (9)-(11) show the updated equations:

$$\hat{\lambda}_{0(MO)}\left(\alpha_{(MO)}\right) = \frac{n_{0(MO)} + u_{1(MO)} n_{1(MO)} + v_{1(MO)} n_{2(MO)}}{\sum_{i \in I_{0(MO)}} x_{i(MO)}^{\alpha_{(MO)}} + \sum_{i \in I_2} x_{1i(MO)}^{\alpha_{(MO)}} + \sum_{i \in I_1} x_{2i(MO)}^{\alpha_{(MO)}}} \tag{9}$$

$$\hat{\lambda}_{1(MO)}\left(\alpha_{(MO)}\right) = \frac{n_{1(MO)} + v_{2(MO)} n_{2(MO)}}{\sum_{i \in I_{0(MO)}} x_{i(MO)}^{\alpha_{(MO)}} + \sum_{I_{3(MO)}} x_{1i(MO)}^{\alpha_{(MO)}}} \tag{10}$$

$$\hat{\lambda}_{2(MO)}\left(\alpha_{(MO)}\right) = \frac{n_{2(MO)} + u_{2(MO)} n_{1(MO)}}{\sum_{i \in I_{0(MO)}} x_{i(MO)}^{\alpha_{(MO)}} + \sum_{I_{3(MO)}} x_{2i(MO)}^{\alpha_{(MO)}}} \tag{11}$$

Using the updated values of $\hat{\lambda}_{0(MO)}$, $\hat{\lambda}_{1(MO)}$, and $\hat{\lambda}_{2(MO)}$, we obtain the following equation:

$$
\begin{aligned}
h_{(MO)}\left(\alpha_{(MO)}\right)=\Bigg[&\hat{\lambda}_{0(MO)}\left(\alpha_{(MO)}\right)\left(\sum_{i\in I_{0(MO)}} x_{i(MO)}^{\alpha_{(MO)}}\ln x_{i(MO)}+\sum_{i\in I_{2(MO)}} x_{1i(MO)}^{\alpha_{(MO)}}\ln x_{1i(MO)}+\sum_{i\in I_{1(MO)}} x_{2i(MO)}^{\alpha_{(MO)}}\ln x_{2i(MO)}\right)\\
&+\hat{\lambda}_{1(MO)}\left(\alpha_{(MO)}\right)\left(\sum_{i\in I_{0(MO)}} x_{i(MO)}^{\alpha_{(MO)}}\ln x_{i(MO)}+\sum_{I_{3(MO)}} x_{1i(MO)}^{\alpha_{(MO)}}\ln x_{1i(MO)}\right)\\
&+\hat{\lambda}_{2(MO)}\left(\alpha_{(MO)}\right)\left(\sum_{i\in I_{0(MO)}} x_{i(MO)}^{\alpha_{(MO)}}\ln x_{i(MO)}+\sum_{I_{3(MO)}} x_{2i(MO)}^{\alpha_{(MO)}}\ln x_{2i(MO)}\right)\\
&-\left(\sum_{i\in I_{0(MO)}}\ln x_{i(MO)}+\sum_{I_{3(MO)}}\left(\ln x_{1i(MO)}+\ln x_{2i(MO)}\right)\right)\Bigg]
\end{aligned}
\tag{12}
$$

By solving Eq.(12), we obtain $\hat{\alpha}_{(j)}$, as shown in Eq.(13) below:

$$
\hat{\alpha}_{(MO)}=\frac{\left(n_{0(MO)}+2n_{1(MO)}+2n_{2(MO)}\right)}{h_{(j)}\left(\alpha_{MO}\right)}
\tag{13}
$$

As the stopping criterion, we choose $\left|\alpha_{(MO)}^{(t)}-\alpha_{(MO)}^{(t+1)}\right|\langle 10^{-6}$, as in the study by Kundu and Dey [9], when updating the parameters. Here, $\alpha_{(MO)}^{(t)}$ denotes the estimation of $\alpha_{(MO)}$ at the t th iteration.

3.2. EM for the BBBW

For the unknown parameters of the BBBW component, which contains no observations of equal value, we use the same notations as those used in the study of Kundu and Gupta [10], as given below:

$$
\begin{aligned}
&I_{1(BB)}=\left\{i: x_{1i(BB)}\langle x_{2i(BB)}\right\}\\
&I_{2(BB)}=\left\{i: x_{1i(BB)}\rangle x_{2i(BB)}\right\}\\
&I_{(BB)}=I_{1(BB)}\cup I_{2(BB)}
\end{aligned}
$$

Here, $\left|I_{1(BB)}\right|=n_{1(BB)}$ and $\left|I_{2(BB)}\right|=n_{2(BB)}$. $\left|I_{l(BB)}\right|$ for $l=0,1,2$ denotes the number of elements in the set $I_{l(BB)}$. We individually fit the Weibull distribution to $X_{1(BB)}$, $X_{2(BB)}$ and $\min\left\{X_{1(BB)},X_{2(BB)}\right\}$. As an initial guess for $\alpha_{(BB)}$, we take the average value of the shape parameters of fitted distributions. As the initial guess values of $\lambda_{0(BB)}$, $\lambda_{1(BB)}$ and $\lambda_{2(BB)}$, we take 1.0, 1.0 and 1.0, respectively. Using the following notations:

$$
u_{1(BB)}=\frac{\lambda_{0(BB)}}{\lambda_{0(BB)}+\lambda_{2(BB)}},\qquad u_{2(BB)}=\frac{\lambda_{2(BB)}}{\lambda_{0(BB)}+\lambda_{2(BB)}},
$$

$$
v_{1(BB)}=\frac{\lambda_{0(BB)}}{\lambda_{0(BB)}+\lambda_{1(BB)}},\qquad v_{2(BB)}=\frac{\lambda_{1(BB)}}{\lambda_{0(BB)}+\lambda_{1(BB)}},
$$

$$
\tilde{n}_{0(BB)}=\left(n_{1(BB)}+n_{2(BB)}\right)\frac{\lambda_{0(BB)}}{\lambda_{1(BB)}+\lambda_{2(BB)}},\qquad a_{0(BB)}=\frac{1}{\left(\lambda_{0(BB)}+\lambda_{1(BB)}+\lambda_{2(BB)}\right)^{\frac{1}{\alpha_{(BB)}}}}\Gamma\left(\frac{1}{\alpha_{(BB)}}+1\right)
$$

we update the parameters $\lambda_{0(BB)}$, $\lambda_{1(BB)}$ and $\lambda_{2(BB)}$ as follows:

$$\hat{\lambda}_{0(BB)}\left(\alpha_{(BB)}\right)=\frac{\tilde{n}_{0(BB)}+u_{1(BB)}n_{1(BB)}+v_{1(BB)}n_{2(BB)}}{\tilde{n}_{0(BB)}a_{0(BB)}^{\alpha_{(BB)}}\sum_{i\in I_{2(BB)}}x_{1i(BB)}^{\alpha_{(BB)}}+\sum_{i\in I_{1(BB)}}x_{2i(BB)}^{\alpha_{(BB)}}} \tag{14}$$

$$\hat{\lambda}_{1(BB)}\left(\alpha_{(BB)}\right)=\frac{n_{1(BB)}+v_{2(BB)}n_{2(BB)}}{\tilde{n}_{0(BB)}a_{0(BB)}^{\alpha_{(BB)}}+\sum_{I_{(BB)}}x_{1i(BB)}^{\alpha_{(BB)}}} \tag{15}$$

$$\hat{\lambda}_{2(BB)}\left(\alpha_{(BB)}\right)=\frac{n_{2(BB)}+u_{2(BB)}n_{1(BB)}}{\tilde{n}_{0(BB)}\alpha_{0(BB)}^{\alpha_{(BB)}}+\sum_{I_{(BB)}}x_{2i(BB)}^{\alpha_{(BB)}}} \tag{16}$$

Using the updated values of $\hat{\lambda}_0$, $\hat{\lambda}_1$ and $\hat{\lambda}_2$, we obtain the following equation:

$$\begin{aligned} h(\alpha)=\hat{\lambda}_0(\alpha)&\left[\tilde{n}_0 a_0^{\alpha}\ln a_0+\sum_{i\in I_1}x_{2i}^{\alpha}\ln x_{2i}+\sum_{i\in I_2}x_{1i}^{\alpha}\ln x_{1i}\right] \\ &+\hat{\lambda}_1(\alpha)\left[\tilde{n}_0 a_0^{\alpha}\ln a_0+\sum_{i\in I}x_{1i}^{\alpha}\ln x_{1i}\right]+\hat{\lambda}_2(\alpha)\left[\tilde{n}_0 a_0^{\alpha}\ln a_0+\sum_{i\in I}x_{2i}^{\alpha}\ln x_{2i}\right] \\ &-\left[\tilde{n}_0\ln a_0+\sum_{i\in I}\ln x_{1i}+\sum_{i\in I}\ln x_{2i}\right] \end{aligned} \tag{17}$$

By solving Eq.(17), we obtain $\hat{\alpha}_{(BB)}$ as follows:

$$\hat{\alpha}_{(BB)}=\frac{\left(n_{0(BB)}+2n_{1(BB)}+2n_{2(BB)}\right)}{h\left(\alpha_{(BB)}\right)} \tag{18}$$

As the stopping criterion, we choose $\left|\alpha_{(BB)}^{(t)}-\alpha_{(BB)}^{(t+1)}\right|\langle 10^{-6}$, as given in the study of Kundu and Gupta [10], for updating the parameters. Here, $\alpha_{(BB)}^{(t)}$ denotes the estimation of $\alpha_{(BB)}$ at the t th iteration.

After estimating the parameters of the MOBW and BBBW distributions for the subclusters obtained using the k-means algorithm, we update the label vectors. To update the label vectors, we consider the EM algorithm containing two EM algorithms for all the observations except those of equal value that use the pdfs of the sub-clusters.

As the stopping criterion for the log-likelihood value of the obtained parameters, we choose $\left|\left(l^k-l^{k-1}\right)/l^{k-1}\right|\langle 10^{-8}$. Here, l^k denotes the value of the log-likelihood function at the k th iteration for each component.

We obtain estimators of the mixing proportions as follows:

$$\pi_1=\frac{n_{0(MO)}+n_{1(MO)}+n_{2(MO)}}{n}\ \text{and}\ \pi_2=\frac{n_{1(BB)}+n_{2(BB)}}{n} \tag{19}$$

The log-likelihood value obtained using the Mix_EM algorithm for all the data is given in Eq.(20), as follows:

$$l=\pi_1\log\left(f_{MOBW}\left(x_{1(MO)},x_{2(MO)}\right)\right)+\pi_2\log\left(f_{BBBW}\left(x_{1(BB)},x_{2(BB)}\right)\right) \tag{20}$$

The iterations stops for the Mix_EM algorithm when $\left|l^{(i+1)}-l^{(i)}\right|\langle 10^{-3}$. Here, $l^{(i)}$ denotes the log-likelihood value obtained from i th iteration of the Mix_EM algorithm.

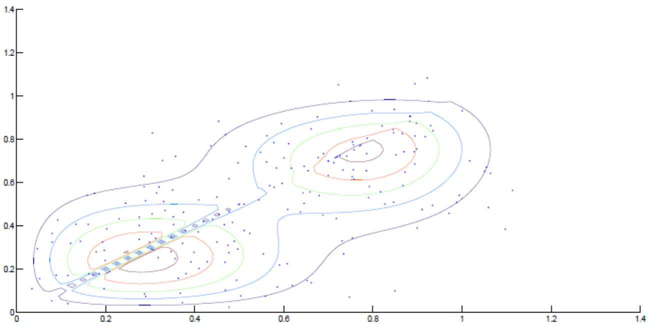
4. DATA ANALYSIS

For demonstration purpose, in this section, we present some results to verify how the proposed Mix_EM algorithm performs on generated data sets with different sample sizes and parameter values. We conduct analyses for three cases related to the sample sizes of MOBW and BBBW distributions, denoted as n_1 and n_2, respectively. We perform individual data generation processes for each of the distributions based on their structures, as defined in section 2, and then merge the separately generated data sets. It is then possible to apply the proposed EM algorithm for the MOBW distribution and that proposed for the mixture of the two MOBW (Mix_MOBW) distributions to the data set generated according to the Mix_BW distribution. In the application of the Mix_EM algorithm, we keep the initial values $\lambda_{0(j)}=1$, $\lambda_{1(j)}=1$, $\lambda_{2(j)}=1$, $(j=1,2)$ and consider the average value of the shape parameters to be the initial value of α. We take the stopping conditions provided in section 3. Tables 1, 2, 3 and 4 present the obtained estimation results for the generated data sets.

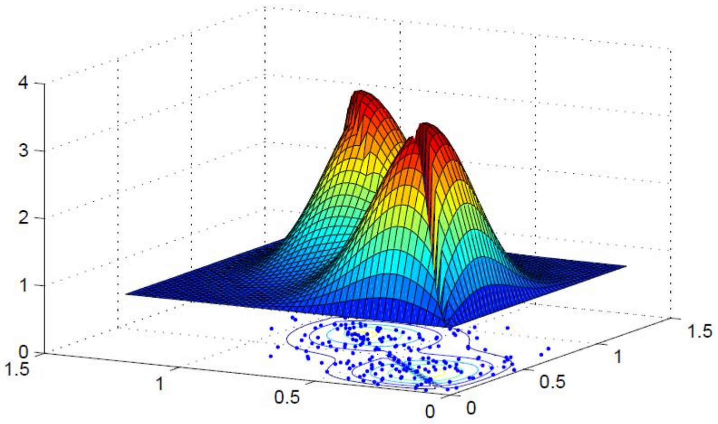
In this section, we also include an analysis of a real data set containing mice data from Sreeja [14]. This data set contains information for 300 rats divided into 50 male litters and 50 female litters, all which had a size of three. In the data, each observation pair represents lifetimes (in weeks) for a pair of mice.

Case 1 (n_1=100, n_2=100): First, we consider Case 1. To obtain the parameter estimations using the Mix_EM algorithm for the Mix_BW model, we generate a synthetic data set using the following parameter values: $\alpha_{(1)}$, $\lambda_{0(1)}$, $\lambda_{1(1)}$, $\lambda_{2(1)}$, and $\pi_{(1)}$ as 2, 2, 4, 6, and 0.50, respectively and $\alpha_{(2)}$, $\lambda_{0(2)}$, $\lambda_{1(2)}$, $\lambda_{2(2)}$, and $\pi_{(2)}$ as 5, 1, 2, 3, and 0.50, respectively. The contours and surface plots of the data set are given in Figs. 1(a) and 1(b), respectively.

The components of Mix_BW are close to each other, as shown in the Figs. 1(a) and 1(b), in which the synthetic data set is an example of a Mix_BW distribution. Table 1 shows the obtained estimation results for the generated data set.



(a)



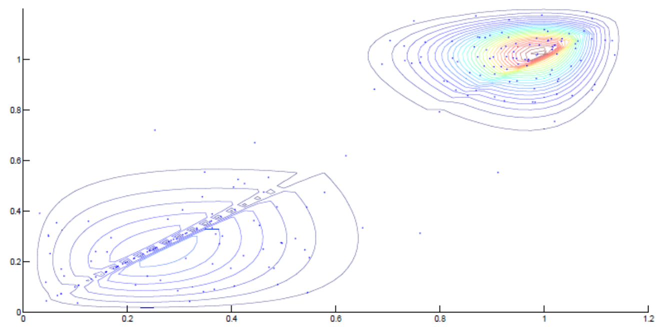
(b)

Figure 1. *(a) The contour plot of the generated data set with components close to each other (b) The surface plot of the generated data set with components close to each other*

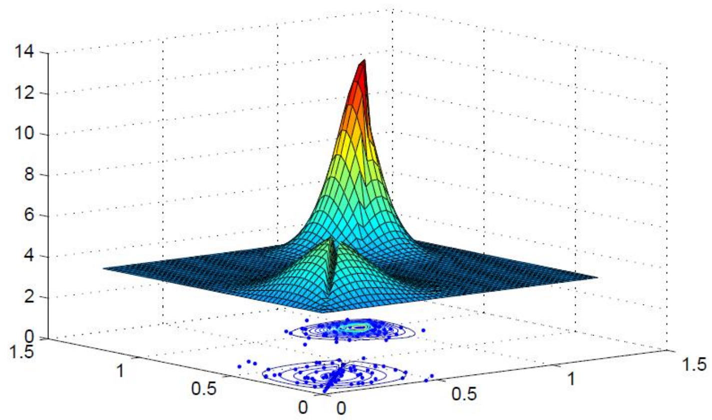
Table 1. *The estimation results for the Case 1*

Distribution	$\alpha_{(1)}$	$\lambda_{0(1)}$	$\lambda_{1(1)}$	$\lambda_{2(1)}$	$\pi_{(1)}$	$\alpha_{(2)}$	$\lambda_{0(2)}$	$\lambda_{1(2)}$	$\lambda_{2(2)}$	$\pi_{(2)}$	AIC ($\log L$)
MOBW	2.1599	0.5515	2.3516	2.8369	-	-	-	-	-	-	129.252 (-60.626)
Mix_BW	2.0917	1.8806	4.5798	8.1728	0.485	5.3506	0.9608	2.4471	2.8367	0.515	14.8312 (1.5844)

Next, we generate a data set in which the components of the Mix_BW distribution are far away from each other, in contrast to the data set presented above. The contours and surface plots of the data set are given in Figs. 2(a) and 2(b), respectively. From Figs. 2(a) and 2(b), we can see that components are far away from each other, with respect to the data set generated from the Mix_BW distribution with the parameter values $\alpha_{(1)}$, $\lambda_{0(1)}$, $\lambda_{1(1)}$, $\lambda_{2(1)}$, and $\pi_{(1)}$ being 2, 5, 5, 5, and 0.50, respectively, and $\alpha_{(2)}$, $\lambda_{0(2)}$, $\lambda_{1(2)}$, $\lambda_{2(2)}$, and $\pi_{(2)}$ being 10, 1, 0.3, 0.1, and 0.50, respectively. Table 2 shows the obtained estimation results for the generated data set.



(a)



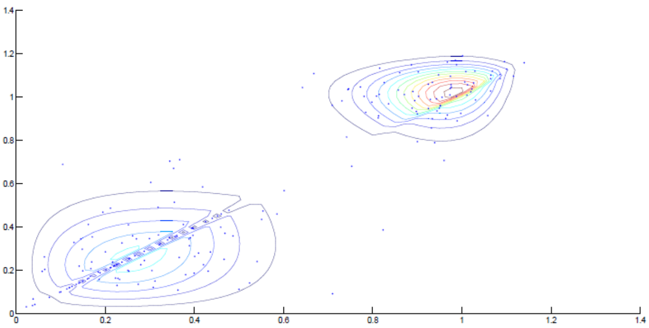
(b)

Figure 2. *(a) The contour plot of the generated data set with components far from each other (b) The surface plot of the generated data set with components far from each other*

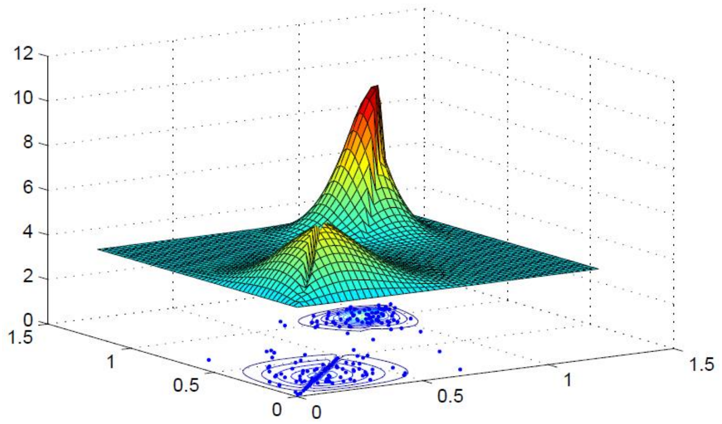
Table 2. *The estimation results for the data set with far from components*

Distribution	$\alpha_{(1)}$	$\lambda_{0(1)}$	$\lambda_{1(1)}$	$\lambda_{2(1)}$	$\pi_{(1)}$	$\alpha_{(2)}$	$\lambda_{0(2)}$	$\lambda_{1(2)}$	$\lambda_{2(2)}$	$\pi_{(2)}$	AIC (log L)
MOBW	1.8074	0.9465	1.4362	0.9658	-	-	-	-	-	-	390.500 (-191.25)
Mix_BW	1.8922	4.9146	5.0475	3.9824	0.5	11.7813	0.5321	1.0425	0.3109	0.5	-155.764 (86.8818)

Case 2 (n_1 =120, n_2 =80): We generate a third synthetic data set using the same parameter values as those used in the Case 1 to obtain components far away from each other with different sample sizes. Note that in the Cases 2 and 3, all the parameters are the same for both distributions although their proportions differ. The Case 2 is more likely a MOBW distribution than a BBBW distribution. The situation in the Case 3 is vice versa. Figure 3 shows the contours and surface plots of the data set for the Case 2, and Table 3 shows the obtained estimation results for the generated data set.



(a)



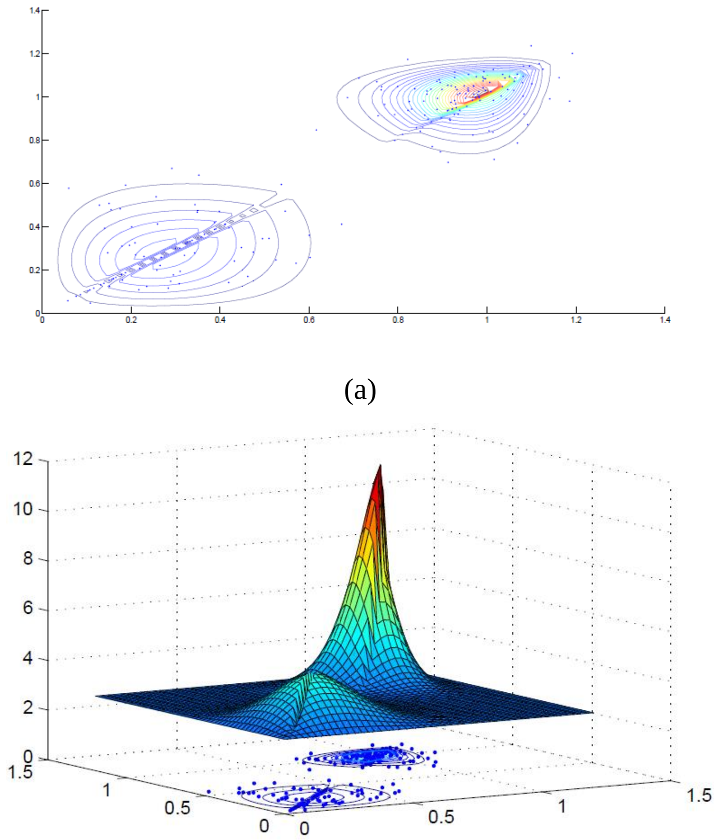
(b)

Figure 3. *(a) The contour plot for case 2 (b) The surface plot for case 2*

Table 3. *The estimation results for the Case 2*

Distribution	$\alpha_{(1)}$	$\lambda_{0(1)}$	$\lambda_{1(1)}$	$\lambda_{2(1)}$	$\pi_{(1)}$	$\alpha_{(2)}$	$\lambda_{0(2)}$	$\lambda_{1(2)}$	$\lambda_{2(2)}$	$\pi_{(2)}$	AIC ($\log L$)
MOBW	2.1537	0.7074	1.3114	0.9077	-	-	-	-	-	-	394.280 (-193.14)
Mix_BW	2.0379	5.0282	4.7744	5.0184	0.6	11.0154	0.7388	0.5613	0.1779	0.4	-167.153 (92.5766)

Case 3 (n_1 **=80,** n_2 **=120):** To obtain the synthetic data set, we choose the same parameter values as those in the Case 2. Figure 4 shows the contours and surface plots of this data set and Table 4 shows the obtained estimation results for the generated data set.



(a)

(b)

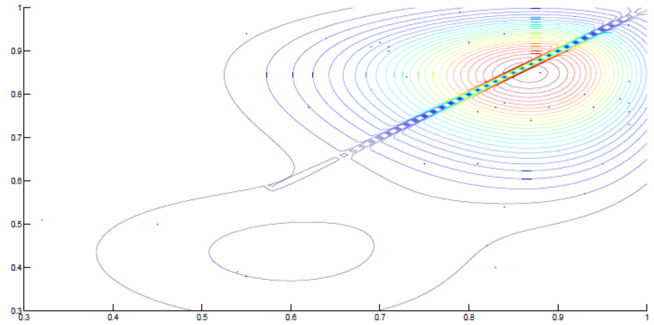
Figure 4. *(a) The contour plot for case 3 (b) The surface plot for case 3*

Table 4. *The estimation results for the Case 3*

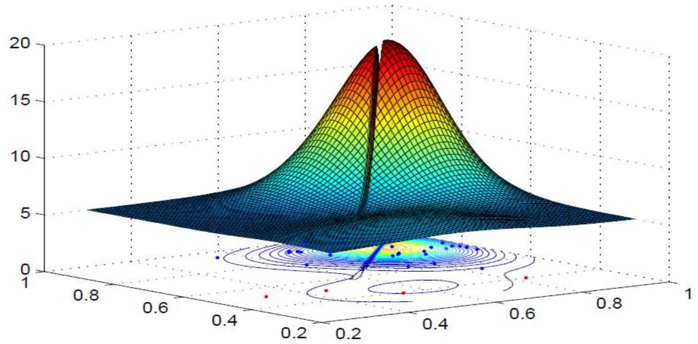
Distribution	$\alpha_{(1)}$	$\lambda_{0(1)}$	$\lambda_{1(1)}$	$\lambda_{2(1)}$	$\pi_{(1)}$	$\alpha_{(2)}$	$\lambda_{0(2)}$	$\lambda_{1(2)}$	$\lambda_{2(2)}$	$\pi_{(2)}$	AIC ($\log L$)
MOBW	1.5923	0.9551	1.5704	1.1404	-	-	-	-	-	-	369.760 (-180.88)
Mix_BW	2.0326	4.3063	6.3453	5.5713	0.40	11.7569	0.5799	0.5740	0.1501	0.60	-106.047 (62.0235)

From Tables 1, 2, 3 and 4, we can see that the performance of the proposed Mix_EM algorithm is satisfactory, based on the $\log L$ and AIC (Akaike information criterion) values, even when the Mix_BW components are far away from each other. If a data set is distributed using a Mix_BW distribution, more satisfactory results are obtained using the proposed Mix_EM algorithm to estimate the parameters. Applying the EM algorithm for a MOBW distribution will not generate better results with respect to the $\log L$ and AIC values. The EM algorithm for the Mix_MOBW distribution does not work for the generated data set since there are no observations for the singular aspect of the MOBW distribution in one of the components of the Mix_BW distribution.

Real Data (Mice Data): In this section, we analyse a data set from Sreeja [14]. The data represent the lifetimes (in weeks) of a pair of mice. We divide all the data points by 100 to generate parameters that are easy to understand. We obtain the maximum likelihood estimators and corresponding $\log L$ values for the MOBW and Mix_BW distributions for the mice data using the EM algorithm. Figure 5 shows the contours and surface plots of the mice data fitted to the Mix_BW distribution.



(a)



(b)

Figure 5. *(a) The contour plot of the mice data (b) The surface plot of the mice data*

Table 5 shows the parameter estimations and the $\log L$ and AIC values of the models. As shown in the table, the obtained $\log L$ and AIC values of the Mix_BW distribution are 34.7050 and -51.41, respectively, so we know that the Mix_BW distribution is more appropriate than the MOBW distribution for this data structure.

Table 5. *The estimation results for the mice data*

Distribution	$\alpha_{(1)}$	$\lambda_{0(1)}$	$\lambda_{1(1)}$	$\lambda_{2(1)}$	$\pi_{(1)}$	$\alpha_{(2)}$	$\lambda_{0(2)}$	$\lambda_{1(2)}$	$\lambda_{2(2)}$	$\pi_{(2)}$	AIC ($\log L$)
MOBW	6.4657	0.1908	2.6344	3.2433	-	-	-	-	-	-	-44.8776 (26.4388)
Mix_BW	4.3532	0.00004	7.2636	29.1190	0.13	9.1637	0.2125	3.0515	3.9731	0.87	-51.41 (34.7050)

5. SIMULATION STUDY

In this section, we present some simulation results to determine how well the Mix_EM algorithm performs for a mixture of MOBW and BBBW distributions with different sample sizes and parameter values. For this purpose, first, we generate samples from the mixture of the MOBW and BBBW distributions using the parameter values $\alpha_{(1)}$, $\lambda_{0(1)}$, $\lambda_{1(1)}$, and $\lambda_{2(1)}$ as 5, 1, 2, and 3, respectively, and $\alpha_{(2)}$, $\lambda_{0(2)}$, $\lambda_{1(2)}$, and $\lambda_{2(2)}$ as 2, 1.5, 5, and 7.5, respectively, for equal and unequal sample sizes, respectively. Next, we generate samples with the values $\alpha_{(1)}$, $\lambda_{0(1)}$, $\lambda_{1(1)}$, $\lambda_{2(1)}$, $\alpha_{(2)}$, $\lambda_{0(2)}$, $\lambda_{1(2)}$, and $\lambda_{2(2)}$ as 2, 5, 5, 5, 11, 0.7, 0.8, and 0.3, respectively. We obtain parameter estimators for the mixtures of MOBW and BBBW distributions using the Mix_EM algorithm. Tables 6 and 7 give the means and standard errors of the EM estimators. The standard error values of the estimators obtained from different sample sizes with different iteration numbers are remarkably close to zero. Moreover, as the sample size increases, the standard errors of the estimators decrease. These results indicate that the Mix_EM algorithm can successfully estimate the parameters.

Table 6. *The means (AV) and standard errors (SE) of the EM estimators for samples generated from mixture of the MOBW and BBBW distributions*

		Iteration number					
		100		**500**		**1000**	
	Parameters	**AV**	**SE**	**AV**	**SE**	**AV**	**SE**
BBBW Parameters (n_1=100)	$\pi_{(1)}$=0.5	0.5528	0.0099	0.5551	0.0043	0.5517	0.0029
	$\alpha_{(2)}$=2	2.1004	0.0190	2.1073	0.0083	2.0980	0.0060
	$\lambda_{0(2)}$=1.5	1.3385	0.0600	1.3064	0.0256	1.3234	0.0206
	$\lambda_{1(2)}$=5	4.9899	0.2092	4.9509	0.0919	4.9192	0.0634
	$\lambda_{2(2)}$=7.5	7.2543	0.3249	7.2954	0.1517	7.1900	0.1030
MOBW Parameters (n_2=100)	$\alpha_{(1)}$=5	5.7324	0.0977	5.7833	0.0422	5.7155	0.0312
	$\lambda_{0(1)}$=1	0.8169	0.0346	0.8134	0.0200	0.7905	0.0132
	$\lambda_{1(1)}$=2	2.3521	0.0614	2.3652	0.0392	2.3358	0.0174
	$\lambda_{2(1)}$=3	3.5526	0.0844	3.6336	0.0400	3.6223	0.0282

		Iteration number					
		100		**500**		**1000**	
	Parameters	**AV**	**SE**	**AV**	**SE**	**AV**	**SE**
BBBW Parameters (n_1=60)	$\pi_{(1)}$=0.3	0.3672	0.0095	0.3562	0.0041	0.3584	0.0029
	$\alpha_{(2)}$=2	2.0851	0.0260	2.0987	0.0124	2.0775	0.0087
	$\lambda_{0(2)}$=1.5	1.5937	0.0743	1.6539	0.0429	1.6206	0.0323
	$\lambda_{1(2)}$=5	4.3999	0.2886	4.8924	0.1487	4.6147	0.0965
	$\lambda_{2(2)}$=7.5	7.0760	0.5228	7.4678	0.2494	7.2795	0.1731
MOBW Parameters (n_2=140)	$\alpha_{(1)}$=5	5.4260	0.0553	5.3995	0.0293	5.4084	0.0174
	$\lambda_{0(1)}$=1	0.7910	0.0316	0.8435	0.0151	0.8613	0.0101
	$\lambda_{1(1)}$=2	2.2705	0.0418	2.2726	0.0182	2.2694	0.0123
	$\lambda_{2(1)}$=3	3.4978	0.0550	3.5129	0.0317	3.4601	0.0185

		Iteration number					
		100		**500**		**1000**	
	Parameters	**AV**	**SE**	**AV**	**SE**	**AV**	**SE**
BBBW Parameters (n_1=140)	$\pi_{(1)}$=0.7	0.7256	0.0077	0.7265	0.0033	0.7255	0.0024
	$\alpha_{(2)}$=2	2.1206	0.0158	2.1053	0.0065	2.1116	0.0046
	$\lambda_{0(2)}$=1.5	1.0832	0.0552	1.1531	0.0296	1.1346	0.0196
	$\lambda_{1(2)}$=5	5.2138	0.1362	5.1213	0.0566	5.2455	0.0441
	$\lambda_{2(2)}$=7.5	7.8316	0.2360	7.6224	0.1011	7.7047	0.0722
MOBW Parameters (n_2=60)	$\alpha_{(1)}$=5	5.9948	0.1311	5.9930	0.0509	5.9710	0.0413
	$\lambda_{0(1)}$=1	0.7679	0.0473	0.7717	0.0211	0.8506	0.0246
	$\lambda_{1(1)}$=2	2.4759	0.0901	2.4291	0.0314	2.3840	0.0231
	$\lambda_{2(1)}$=3	3.7338	0.1407	3.7422	0.0474	3.7180	0.0419

Table 7. *The means (AV) and standard errors (SE) of the EM estimators for samples generated from mixture of the MOBW and BBBW distributions*

		Iteration number					
		100		500		1000	
	Parameters	**AV**	**SE**	**AV**	**SE**	**AV**	**SE**
BBBW Parameters (n_1=100)	$\pi_{(1)}=0.5$	0.5055	0.0033	0.5092	0.0019	0.5080	0.0012
	$\alpha_{(2)}=11$	11.1301	0.1573	11.0735	0.0776	11.0389	0.0541
	$\lambda_{0(2)}=0.7$	0.6612	0.0143	0.6632	0.0063	0.6633	0.0043
	$\lambda_{1(2)}=0.8$	0.7786	0.0130	0.7723	0.0058	0.7708	0.0041
	$\lambda_{2(2)}=0.3$	0.2776	0.0088	0.2803	0.0038	0.2840	0.0029
MOBW Parameters (n_2=100)	$\alpha_{(1)}=2$	2.0169	0.0137	2.0514	0.0063	2.0410	0.0044
	$\lambda_{0(1)}=5$	5.0802	0.1148	5.1387	0.0501	5.1902	0.0389
	$\lambda_{1(1)}=5$	5.3346	0.1433	5.5824	0.0773	5.5369	0.0548
	$\lambda_{2(1)}=5$	5.2828	0.1301	5.6670	0.0808	5.5732	0.0583

		Iteration number					
		100		500		1000	
	Parameters	**AV**	**SE**	**AV**	**SE**	**AV**	**SE**
BBBW Parameters (n_1=60)	$\pi_{(1)}=0.3$	0.3085	0.0034	0.3119	0.0021	0.3087	0.0012
	$\alpha_{(2)}=11$	10.8660	0.2504	10.8612	0.1128	10.9285	0.0740
	$\lambda_{0(2)}=0.7$	0.9277	0.1049	0.8610	0.0446	0.7914	0.0281
	$\lambda_{1(2)}=0.8$	1.6423	0.3345	1.3187	0.1175	1.3477	0.0872
	$\lambda_{2(2)}=0.3$	1.2122	0.3523	0.9237	0.1371	0.9764	0.1107
MOBW Parameters (n_2=140)	$\alpha_{(1)}=2$	2.0332	0.0127	2.0264	0.0052	2.0337	0.0036
	$\lambda_{0(1)}=5$	4.9120	0.1282	4.8984	0.0492	4.9939	0.0365
	$\lambda_{1(1)}=5$	4.8379	0.1389	5.1488	0.0709	5.1159	0.0469
	$\lambda_{2(1)}=5$	4.9137	0.1540	5.1321	0.0763	5.0445	0.0481

		Iteration number					
		100		500		1000	
	Parameters	**AV**	**SE**	**AV**	**SE**	**AV**	**SE**
BBBW Parameters (n_1=140)	$\pi_{(1)}=0.7$	0.7026	0.0020	0.7023	0.0007	0.7023	0.0005
	$\alpha_{(2)}=11$	11.3163	0.1043	11.2349	0.0433	11.2419	0.0296
	$\lambda_{0(2)}=0.7$	0.6663	0.0121	0.6604	0.0049	0.6488	0.0036
	$\lambda_{1(2)}=0.8$	0.7665	0.0108	0.7687	0.0051	0.7626	0.0035
	$\lambda_{2(2)}=0.3$	0.2670	0.0074	0.2720	0.0033	0.2771	0.0023
MOBW Parameters (n_2=60)	$\alpha_{(1)}=2$	2.0330	0.0178	2.0685	0.0074	2.0477	0.0053
	$\lambda_{0(1)}=5$	5.3306	0.1683	5.4164	0.0615	5.3228	0.0452
	$\lambda_{1(1)}=5$	5.2655	0.1663	5.5546	0.0739	5.5561	0.0533
	$\lambda_{2(1)}=5$	5.3213	0.1433	5.6054	0.0726	5.5487	0.0602

6. CONCLUSIONS

In this paper, we use MOBW and BBBW distributions in a new finite mixture bivariate model. Although the data set structure that inspired this study is not wide enough, especially for data sets related to the health field and modelled by the MOBW distribution, our proposed finite mixture bivariate model could serve as an alternative for mixtures of the Marshall-Olkin bivariate Weibull distributions. We call the proposed mixture distribution the Mix_BW. To estimate the model parameters, the EM algorithm is adapted and it is called as the Mix_EM. We demonstrate the performance of the Mix_BW model and Mix_EM algorithm in the analysis of a real data example and in a simulation study. Both the simulation study and real data analysis confirm that the Mix_EM algorithm achieves satisfactory results.

CONFLICTS OF INTEREST

No conflict of interest was declared by the authors.

REFERENCES

[1] Block, H.W. and Basu, A.P., "A continuous bivariate exponential extension", Journal of the American Statistical Association, 69(348): 1031-1037, (1974).

[2] Çalış, N., Türkan, A.H., Erişoğlu, M., Erişoğlu, Ü. and Erol, H., "Em algorithm for the mixtures of the marshall - olkin bivariate weibull distributions", Advances and Applications in Statistics, 24(1): 35-48, (2011).

[3] Diawara, N. and Carpenter, M., "Mixture of bivariate exponential distributions", Communications in Statistics - Theory and Methods, 39(15): 2711-2720, (2010).

[4] Erişoğlu, Ü., Erişoğlu, M. and Erol, H., "Mixture model approach to the analysis of heterogeneous survival data", Pakistan Journal of Statistics, 28(1): 115-130, (2012).

[5] Everitt, B.S. and Hand, D.J., Finite mixture distributions, John Wiley & Sons, Ltd., Netherlands, (1981).

[6] Freund, J.E., "A bivariate extension of the exponential distribution", Journal of the American Statistical Association, 56(296): 971-977, (1961).

[7] Gumbel, E.J., "Bivariate exponential distributions", Journal of the American Statistical Association, 55(292): 698-707, (1960).

[8] Han, Y., "Location Scale Bivariate Weibull distributions for bivariate lifetime modeling", PhD thesis, Alabama: Auburn University, (2006).

[9] Kundu, D. and Dey, A.K., "Estimating the parameters of the marshall - olkin bivariate weibull distribution by em algorithm", Computational Statistics & Data Analysis, 53(4): 956-965, (2009).

[10] Kundu, D. and Gupta, R.D., "A class of absolutely continuous bivariate distributions", Statistical Methodology, 7(4): 464-477, (2010).

[11] Lu, J.C., "Weibull extensions of the freund and marshall - olkin bivariate exponential models", IEEE Transactions on Reliability, 38(5): 615-619, (1989).

[12] Marshall, A.W. and Olkin, I., "A multivariate exponential distribution", Journal of the American Statistical Association, 62(317): 30-44, (1967).

[13] Murthy, D.N.P., Xie, M. and Jiang, R., Weibull models, John Wiley & Sons, Hoboken, New Jersey, (2004).

[14] Sreeja, V.N. and Sankaran, P.G., “Regression models for bivariate survival data”, PhD thesis, Cochin University of Science and Technology, (2008).