Asian Journal of Applied Sciences (ISSN: 2321 – 0893)
Volume 05 – Issue 06, December 2017

Parameters Estimation of the Bivariate Weibull Distribution by EM Algorithm Based on Censored Samples

Radwan E. Bedar

Department of Statistics, Institute for Computer and Management Sciences,
Thebes Academy, Cairo, Egypt
Email: radwanbeda [AT] yahoo.com

ABSTRACT---- ***In this paper, the parameters of Marshall-Olkin bivariate Weibull distribution are estimated. The maximum likelihood estimators under type I censoring are estimated. Also, the expectation-maximization algorithm is considered. Numerical studies are considered to compare the estimates. Hanagal and Ahmadi [6] results may be considered as special case from our results.***

Keywords--- Bivariate Weibull, Censored sample, EM algorithm, Maximum likelihood estimation, Missing data

---

1. INTRODUCTION

Hanagal and Ahmadi [6] estimated the parameters Marshall-Olkin bivariate exponential [7] under censored samples using maximum likelihood estimators (MLEs) and compared them by that expectation-maximization (EM) algorithm. We follow the same arguments to estimate the parameters of Marshall-Olkin bivariate Weibull (MOBW) distribution, in which the marginals are Weibull distributions

The censoring time (T) is assumed to be independent of the life times (X, Y) of the two components. The bivariate density function of (X, Y) is denoted by $f_{X,Y}(x, y)$ The considered situation occurs for example in medical studies of paired organs like kidneys, eyes, lungs, or any other paired organs of an individual as a two components system which works under interdependency circumstances. Failure of an individual may censor failure of either one of the paired organ or both. This scheme of censoring is right censoring.

There are similar situation in engineering science whenever sub-systems are considered having two components with life times (X, Y) being independent of the life time (T) of the entire system. However, failure of the main system may censor failure of either one component or both. Hanagal [4,5] derived maximum likelihood estimators of the parameters of the bivariate exponential distribution under right censoring samples.

Censoring may also occur in other ways. Patients may be lost to follow up during the study. They may decide to move elsewhere or may become non-cooperative therefore the experimenter cannot follow them or her again. Such cases are called withdrawal from the study. A patient with censored data contributes valuable information and should therefore not be omitted from the analysis.

Because of lack of data of real processes, the data in this study are generated from the MOBW of Marshall-Olkin [7] using Mathcad software. Subsequently the EM algorithm is used to estimate the parameters.

2. MARSHALL-OLKIN BIVARIATE WEIBULL DISTRIBUTION

Marshall-Olkin [7] suggested a more flexible bivariate MOBW distribution with many interesting properties like joint probability density function, marginal probability density functions and joint survival function of the bivariate Weibull distribution given for $x, y > 0, \lambda_1, \lambda_2, \lambda_3 > 0$ and $\alpha > 0$ as follows:

$$\bar{F}_{X,Y}(x, y) = e^{-\lambda_1 x^{\alpha} - \lambda_2 y^{\alpha} - \lambda_3 max(x,y)} \quad (1)$$

The distribution (1) is not absolutely continuous with respect to the Lebesque measure in R^2. It has singularities on the diagonal $X = Y$. The parameter λ_3 reflects the dependence between the two lifetimes X and Y and it is also related to the simultaneous failures of the two components. When $\lambda_3 = 0$, the lifetimes X and Y are independent. Simultaneous failure occurs of X and Y have the same lifetime. The probability of simultaneous failure is given by

$$P_{X,Y}(\{(x,y)|x=y\}) = \frac{\lambda_3}{(\lambda_1+\lambda_2+\lambda_3)} \quad (2)$$

which is also equal to the correlation coefficient between X and Y.

The density functions of $X|\{(x,y) \mid x > y\}$, $Y \mid \{(x,y) \mid y > x\}$ and $Z = min(X,Y)$ are given as follows:

$$\begin{aligned} f_{X|\{(x,y) \mid x>y\}}(x) &= (\lambda_1+\lambda_3)\alpha x^{\alpha-1}e^{-(\lambda_1+\lambda_3)(x^\alpha-y^\alpha)} \\ f_{Y|\{(x,y) \mid y>x\}}(y) &= (\lambda_2+\lambda_3)\alpha y^{\alpha-1}e^{-(\lambda_2+\lambda_3)(y^\alpha-x^\alpha)} \\ f_Z(z) &= (\lambda_1+\lambda_2+\lambda_3)\alpha z^{\alpha-1}e^{-(\lambda_1+\lambda_2+\lambda_3)z^\alpha} \end{aligned} \quad (3)$$

This paper aims at deriving an estimation method for the parameters of a bivariate Weibull distribution of Marshall-Olkin by the EM algorithm. In Section 2, the EM algorithm is presented, while in Section 3, the parameters are estimated by applying the EM algorithm. Finally in Section 4, we present the results of a simulation study.

3. THE EM ALGORITHM

The EM algorithm was introduced by Dempster et al. [2], we follow the same steps of Hanagal and Ahmadi [6]. The algorithm is an iterative procedure for finding the maximum likelihood estimates for incomplete, missing, unobserved or censored data. Because it is easy to implement, the impact of the EM algorithm has been far reaching, not only as computational tool but also as a way of solving difficult statistical problems. The basic idea behind the method is to transform a set of incomplete data into a complete data problem for which the required maximization is computationally more tractable and numerically stable. Each iteration increases the likelihood, which finally converges almost always to a local maximum. One can view the complete data set x as consisting of the vectors (t, t^*), where t is the observed, i.e., incomplete data, and t^* is the missing data.

3.1 The Iterations

The objective is to draw inferences about the parameter vectors $\underline{\lambda} = \lambda_1, \lambda_2, \lambda_3$ and α. We will use $L_c(\underline{\lambda}|t)$ to denote the likelihood function where t is the vector of observed data. Let t^* represent the vector of missing data. Starting with a guessed value for the parameter $\underline{\lambda}$, carry out the following iterations.

- Replace the missing data t^* by their expectation given the guessed value of the parameter vector and the observed data. Let this conditional expectation be denoted by $\tilde{t}^*$.
- Maximize $L_c(\underline{\lambda}, t^*|t)$ with respect to $\lambda_1, \lambda_2, \lambda_3$ and α replacing the missing data t^* by their expected values. This is equivalent to maximizing $L_c(\underline{\lambda}, E[T^*|t])$.
- Re-estimate the missing values t^* using their conditional expectation based on the updated $\underline{\lambda}$.
- Re-estimate $\underline{\lambda}$ and continue until the difference between a new iterated value and the previous iterated value is less than 0.00001.

3.2 Execution of the Algorithm

- Consider the logarithm of the likelihood log $L_c(\underline{\lambda}, t^* |t)$.
- Let the conditional expectation of log $L_c(\underline{\lambda}, t^*|t)$ with respect to $t^*|(\underline{\lambda}, t)$ be denoted by $Q(\underline{\lambda}|\underline{\lambda}^{(k)})$, where $\underline{\lambda}^{(k)}$ is the current guess of $\underline{\lambda}$. Then the EM-steps are as follows:

1. **E-step:** Calculate $Q(\underline{\lambda}|\underline{\lambda}^{(k)})$, that is, the expectation of the log-likelihood with respect to the conditional distribution of the missing data, given the observed data and the current guess of $\underline{\lambda}$.
2. **M-step:** Maximize $Q(\underline{\lambda}|\underline{\lambda}^{(k)})$ with respect to $\underline{\alpha}$ and set the result equal to $\underline{\lambda}^{(k+1)}$, the new value of the parameter vector. Since $\underline{\lambda}^{(k+1)}$ maximizes $Q(\underline{\lambda}|\underline{\lambda}^{(k)})$, the M-Step results in $Q(\underline{\lambda}^{(k+1)}|\underline{\lambda}^{(k)}) \geq Q(\underline{\lambda}|\underline{\lambda}^{(k)})$ for all $\underline{\lambda}^{(k+1)}$, implying that $\underline{\lambda}^{(k+1)}$ is a solution to the equation, $\frac{\partial Q(\underline{\lambda}|\underline{\lambda}^{(k)})}{\partial \underline{\lambda}} = 0$

The two steps are repeated iteratively until the difference between two successive iterations is less than 0.00001. This iterative procedure leads to a monotonic increase of $\log L_c\left(\underline{\lambda}, E[T^*|t]\right)$:

$$\log L_c\left(\underline{\lambda}^{(k+1)}, E(T^*|t)\right) \geq \log L_c\left(\underline{\lambda}^{(k)}, E(T^*|t)\right) \quad \text{for } k = 1,2,\ldots \tag{4}$$

Since the likelihood increases in each step, the EM algorithm converges generally to a local maximum.

When there is no closed form solution of the M-step, a numerical algorithm as, for example, the Newton-Raphson procedure, may be used for iteratively computing $\underline{\lambda}^{(k)}$. In fact, in this paper the Newton-Raphson procedure is used to obtain maximum likelihood estimates of $\underline{\lambda}$ at the $(k+1)$-iteration as follows:

$$\underline{\lambda}^{(k+1)} = \underline{\lambda}^{(k)} - \left[\frac{\partial^2 Q\left(\underline{\lambda}|\underline{\lambda}^{(k)}\right)}{\partial\lambda\partial\underline{\lambda}'}\right]^{-1}_{\underline{\lambda}=\underline{\lambda}^{(k)}} \left[\frac{\partial Q\left(\underline{\lambda}|\underline{\lambda}^{(k)}\right)}{\partial\underline{\lambda}}\right]_{\underline{\lambda}=\underline{\lambda}^{(k)}} \tag{5}$$

The iterative procedure is carried out until the difference $\underline{\lambda}^{(k+1)} - \underline{\lambda}^{(k)} < 0.00001$.

4. PARAMETER ESTIMATION

The density function of (X, Y) is given by

$$f_{X,Y}(x,y) = \begin{cases} \lambda_1 \alpha x^{\alpha-1} e^{-\lambda_1 x^\alpha} (\lambda_2+\lambda_3)\alpha y^{\alpha-1} e^{-(\lambda_2+\lambda_3)y^\alpha} & \text{for } 0<x<y \\ (\lambda_1+\lambda_3)\alpha x^{\alpha-1} e^{-(\lambda_1+\lambda_3)x^\alpha} \lambda_2 \alpha y^{\alpha-1} e^{-\lambda_2 y^\alpha} & \text{for } 0<y<x \\ \lambda_3 \alpha x^{\alpha-1} e^{-(\lambda_1+\lambda_2+\lambda_3)x^\alpha} & \text{for } 0<x=y \end{cases} \tag{6}$$

For the bivariate life time distribution, we use the univariate censoring scheme given by Hanagal [4,5] because the individuals do not enter at the same time the study and withdrawal or death of an individual or termination of the study will censor both life times of the components. Here the censoring time is independent of the life times of both components. This is the standard univariate right censoring for both failure times X and Y.

Suppose that there are n independent pairs of components, for example, paired kidneys, lungs, eyes, ears in an individual under study and *i*-th pair of the components have life times (X_i, Y_i) and a censoring time (T_i) .The life times associated with *i*-th pair of the components are given by

$$(X_i, Y_i) = \begin{cases} (X_i, Y_i) & \text{for } \max(X_i, Y_i) < T_i \\ (X_i, T_{iy}) & \text{for } X_i < T_i < T_{iy} = Y_i \\ (T_{ix}, Y_i) & \text{for } Y_i < T_i < T_{ix} = X_i \\ (T_{ixy}, T_{ixy}) & \text{for } T_i < T_{ixy} = \min(X_i, Y_i) \end{cases} \tag{7}$$

where T_{ix}, T_{iy} and T_{ixy} represent the unobserved random variables.

The likelihood of the sample of size n after discarding factors which do not contain any of the parameters of interest is given as follows:

$$L_c\left(\underline{\lambda}, \alpha, t_{ix}, t_{iy}, t_{ixy} | x_i, y_i, t_i\right) = \prod_{j=1}^{6} \prod_{i=1}^{n_j} f_{X,Y}^{(j)}(x_i, y_i) \tag{8}$$

where

$$\begin{aligned} &f_{X,Y}^{(1)}(x,y) = \lambda_1 \alpha x^{\alpha-1} e^{-\lambda_1 x^\alpha} (\lambda_2+\lambda_3)\alpha y^{\alpha-1} e^{-(\lambda_2+\lambda_3)y^\alpha} && if\ 0<x<y<\infty \\ &f_{X,Y}^{(2)}(x,y) = (\lambda_1+\lambda_3)\alpha x^{\alpha-1} e^{-(\lambda_1+\lambda_3)x^\alpha} \lambda_2 \alpha y^{\alpha-1} e^{-\lambda_2 y^\alpha} && if\ 0<y<x<\infty \\ &f_{X,Y}^{(3)}(x,y) = \lambda_3 \alpha x^{\alpha-1} e^{-(\lambda_1+\lambda_2+\lambda_3)x^\alpha} && if\ 0<x=y<\infty \\ &f_{X,Y}^{(4)}(x,y) = \lambda_1 \alpha x^{\alpha-1} e^{-\lambda_1 x^\alpha} (\lambda_2+\lambda_3)\alpha (t_{iy})^{\alpha-1} e^{-(\lambda_2+\lambda_3)(t_{iy})^\alpha} && if\ 0<x<t<t_y=y \\ &f_{X,Y}^{(5)}(x,y) = (\lambda_1+\lambda_3)\alpha (t_{ix})^{\alpha-1} e^{-(\lambda_1+\lambda_3)(t_{ix})^\alpha} \lambda_2 \alpha y^{\alpha-1} e^{-\lambda_2 y^\alpha} && if\ 0<y<t<t_x=x \\ &f_{X,Y}^{(6)}(x,y) = f_{X,Y}(t_{xy}) = (\lambda_1+\lambda_2+\lambda_3)\alpha (t_{ixy})^{\alpha-1} \end{aligned} \tag{9}$$

$$\times e^{-(\lambda_1+\lambda_2+\lambda_3)(t_{ixy})^{\alpha}} \qquad\qquad if\;\; t < t_{xy} = min(x,y)$$

Let n_1, n_2, n_3, n_4, n_5 *and* n_6 the number of realizations falling in the range corresponding to $f_{X,Y}^{(1)}(x,y)$, $f_{X,Y}^{(2)}(x,y)$, $f_{X,Y}^{(3)}(x,y)$, $f_{X,Y}^{(4)}(x,y)$, $f_{X,Y}^{(5)}(x,y)$ and $f_{X,Y}^{(6)}(x,y)$ respectively. $f_{X,Y}^{(1)}(x,y), f_{X,Y}^{(2)}(x,y)$, $f_{X,Y}^{(4)}(x,y)$ and $f_{X,Y}^{(5)}(x,y)$ are density functions with respect to the Lebesque measure on R^2, while $f_{X,Y}^{(3)}(x,y)$ and $f_{X,Y}^{(6)}(x,y)$ are density functions with respect to the Lebesque measure on R.

The log-likelihood function $\log L_c(\underline{\alpha}, \lambda, t_{ix}, t_{iy}, t_{ixy} | x_i, y_i, t_i)$ can be written as follows:

$$\begin{aligned}
&lnL_c(\underline{\lambda}, \alpha, t_{ix}, t_{iy}, t_{ixy} | x_i, y_i, t_i) = (n_1+n_4)ln(\lambda_1) + (n_1+n_4)ln(\lambda_2+\lambda_3) + (n_2+n_5)ln(\lambda_2) \\
&+(n_2+n_5)ln(\lambda_1+\lambda_3) + (n_3)ln(\lambda_3) + n_6 ln(\lambda_1+\lambda_2+\lambda_3) \\
&+(2n_1+2n_2+n_3+2n_4+2n_5+n_6)ln(\alpha) \\
&+(\alpha-1)\left\{\begin{array}{l} \sum_{i=1}^{n_1} ln(x_i) + \sum_{i=1}^{n_1} ln(y_i) + \sum_{i=1}^{n_2} ln(x_i) + \sum_{i=1}^{n_2} ln(y_i) + \sum_{i=1}^{n_3} ln(x_i) + \sum_{i=1}^{n_4} ln(x_i) \\ \qquad + \sum_{i=1}^{n_4} ln(t_{iy}) + \sum_{i=1}^{n_5} ln(y_i) + \sum_{i=1}^{n_5} ln(t_{ix}) + \sum_{i=1}^{n_6} ln(t_{ixy}) \end{array}\right\} \\
&-\lambda_1\left\{\sum_{i=1}^{n_1}(x_i)^{\alpha} + \sum_{i=1}^{n_2}(x_i)^{\alpha} + \sum_{i=1}^{n_3}(x_i)^{\alpha} + \sum_{i=1}^{n_4}(x_i)^{\alpha} + \sum_{i=1}^{n_4}(t_{ix})^{\alpha} + \sum_{i=1}^{n_4}(t_{ixy})^{\alpha}\right\} \\
&-\lambda_2\left\{\sum_{i=1}^{n_1}(y_i)^{\alpha} + \sum_{i=1}^{n_2}(y_i)^{\alpha} + \sum_{i=1}^{n_3}(x_i)^{\alpha} + \sum_{i=1}^{n_4}(t_{iy})^{\alpha} + \sum_{i=1}^{n_4}(y_i)^{\alpha} + \sum_{i=1}^{n_4}(t_{ixy})^{\alpha}\right\} \\
&-\lambda_3\left\{\sum_{i=1}^{n_1}(y_i)^{\alpha} + \sum_{i=1}^{n_2}(x_i)^{\alpha} + \sum_{i=1}^{n_3}(x_i)^{\alpha} + \sum_{i=1}^{n_4}(t_{iy})^{\alpha} + \sum_{i=1}^{n_4}(t_{ix})^{\alpha} + \sum_{i=1}^{n_4}(t_{ixy})^{\alpha}\right\}
\end{aligned}$$

Thus, the following E-step and M-step are obtained:

- **E-step:**

The unobserved random variables, T_x, T_y *and* T_{xy} follow Weibull distributions respectively as stated in (3). Because of the loss of memory property of the Weibull distribution, the conditional distributions of $T_x|\{(t_x, t, y)|t_x > t > y\}$, $T_y|\{(t_y, t, x)|t_y > t > x\}$ and $T_{xy}|\{(t_{xy}, t)|t_{xy} > t\}$ still follow the Weibull distributions and the corresponding density functions are as follows:

$$f_{T_x|\{(t_x,t,y)|t_x>t>y\}}(t_x) = (\lambda_1+\lambda_3)\alpha(t_x)^{\alpha-1}e^{-(\lambda_1+\lambda_3)((t_x)^{\alpha}-t^{\alpha})}$$

$$f_{T_y|\{(t_y,t,x)|t_y>t>x\}}(t_y) = (\lambda_2+\lambda_3)\alpha(t_y)^{\alpha-1}e^{-(\lambda_2+\lambda_3)((t_y)^{\alpha}-t^{\alpha})}$$

$$f_{T_{xy}|\{(t_{xy},t)|t_{xy}>t\}}(t_{xy}) = (\lambda_1+\lambda_2+\lambda_3)\alpha(t_{ixy})^{\alpha-1}e^{-(\lambda_1+\lambda_2+\lambda_3)((t_{ixy})^{\alpha}-t^{\alpha})}$$

The values of the first moments of the conditional unobserved random variables, $T_x|\{(t_x, t, y)|t_x > t > y\}$, $T_y|\{(t_y, t, x)|t_y > t > x\}$ and $T_{xy}|\{(t_{xy}, t)|t_{xy} > t\}$ are as follows.

$$E[T_x^{\alpha}|\{(t_x, t, y)|t_x > t > y\}] = t^{\alpha} + \frac{1}{(\lambda_1+\lambda_3)} = t^{\alpha} + a$$

$$E\left[T_y^{\alpha}|\{(t_y,t,x)|t_y>t>x\}\right]=t^{\alpha}+\frac{1}{(\lambda_2+\lambda_3)}=t^{\alpha}+b$$

$$E\left[T_{xy}^{\alpha}|\{(t_{xy},t)|t_{xy}>t\}\right]=t^{\alpha}+\frac{1}{(\lambda_1+\lambda_2+\lambda_3)}=t^{\alpha}+c$$

The conditional expectation $Q(\underline{\lambda}|\underline{\lambda}^{(k)})$ is obtained as follows:

$$\begin{aligned}Q(\underline{\lambda}|\underline{\lambda}^{(k)}) &= (n_1+n_4)ln(\lambda_1)+(n_1+n_4)ln(\lambda_2+\lambda_3)+(n_2+n_5)ln(\lambda_2)\\ &+(n_2+n_5)ln(\lambda_1+\lambda_3)+(n_3)ln(\lambda_3)+n_6ln(\lambda_1+\lambda_2+\lambda_3)\\ &+(2n_1+2n_2+n_3+2n_4+2n_5+n_6)ln(\alpha)\\ &+(\alpha-1)\left\{\begin{array}{l}\sum_{i=1}^{n_1}ln(x_i)+\sum_{i=1}^{n_1}ln(y_i)+\sum_{i=1}^{n_2}ln(x_i)+\sum_{i=1}^{n_2}ln(y_i)+\sum_{i=1}^{n_3}ln(x_i)+\sum_{i=1}^{n_4}ln(x_i)\\ +\sum_{i=1}^{n_4}ln(t_{iy})+\sum_{i=1}^{n_5}ln(y_i)+\sum_{i=1}^{n_5}ln(t_{ix})+\sum_{i=1}^{n_6}ln(t_{ixy})\end{array}\right\}\\ &-\lambda_1\left\{\sum_{i=1}^{n_1}(x_i)^{\alpha}+\sum_{i=1}^{n_2}(x_i)^{\alpha}+\sum_{i=1}^{n_3}(x_i)^{\alpha}+\sum_{i=1}^{n_4}(x_i)^{\alpha}+\sum_{i=1}^{n_5}(t^{\alpha}+a^{(k)})+\sum_{i=1}^{n_6}(t^{\alpha}+c^{(k)})\right\}\\ &-\lambda_2\left\{\sum_{i=1}^{n_1}(y_i)^{\alpha}+\sum_{i=1}^{n_2}(y_i)^{\alpha}+\sum_{i=1}^{n_3}(x_i)^{\alpha}+\sum_{i=1}^{n_4}(t^{\alpha}+b^{(k)})+\sum_{i=1}^{n_5}(y_i)^{\alpha}+\sum_{i=1}^{n_6}(t^{\alpha}+c^{(k)})\right\}\\ &-\lambda_3\left\{\sum_{i=1}^{n_1}(y_i)^{\alpha}+\sum_{i=1}^{n_2}(x_i)^{\alpha}+\sum_{i=1}^{n_3}(x_i)^{\alpha}+\sum_{i=1}^{n_4}(t^{\alpha}+b^{(k)})+\sum_{i=1}^{n_5}(t^{\alpha}+a^{(k)})+\sum_{i=1}^{n_6}(t^{\alpha}+c^{(k)})\right\}\end{aligned}$$

where a_1, a_2 and a_3 equal to $\lambda_1{}^{(k)}$, $\lambda_2{}^{(k)}$ and $\lambda_3{}^{(k)}$ at k^{th} iteration.

- **M-Step:**

The following likelihood equations are obtained by equating the partial derivatives of $Q(\underline{\lambda}|\underline{\lambda}^{(k)})$ with respect to λ_1, λ_2 and λ_3 to zero:

$$\begin{aligned}\frac{\partial Q(\underline{\lambda}|\underline{\lambda}^{(k)})}{\partial\lambda_1} &= \frac{(n_1+n_4)}{\lambda_1}+\frac{(n_2+n_5)}{(\lambda_1+\lambda_3)}+\frac{n_6}{(\lambda_1+\lambda_2+\lambda_3)}\\ &-\left\{\sum_{i=1}^{n_1}(x_i)^{\alpha}+\sum_{i=1}^{n_2}(x_i)^{\alpha}+\sum_{i=1}^{n_3}(x_i)^{\alpha}+\sum_{i=1}^{n_4}(x_i)^{\alpha}+\sum_{i=1}^{n_4}(t^{\alpha}+a^{(k)})+\sum_{i=1}^{n_4}(t^{\alpha}+c^{(k)})\right\}\end{aligned}$$

$$\begin{aligned}\frac{\partial Q(\underline{\lambda}|\underline{\lambda}^{(k)})}{\partial\lambda_2} &= \frac{(n_1+n_4)}{(\lambda_2+\lambda_3)}+\frac{(n_2+n_5)}{\lambda_2}+\frac{n_6}{(\lambda_1+\lambda_2+\lambda_3)}\\ &-\left\{\sum_{i=1}^{n_1}(y_i)^{\alpha}+\sum_{i=1}^{n_2}(y_i)^{\alpha}+\sum_{i=1}^{n_3}(x_i)^{\alpha}+\sum_{i=1}^{n_4}(t^{\alpha}+b^{(k)})+\sum_{i=1}^{n_4}(y_i)^{\alpha}+\sum_{i=1}^{n_4}(t^{\alpha}+c^{(k)})\right\}\end{aligned}$$

$$\begin{aligned}\frac{\partial Q(\underline{\alpha}|\underline{\alpha}^{(k)})}{\partial\lambda_3} &= \frac{(n_1+n_4)}{(\lambda_2+\lambda_3)}+\frac{(n_2+n_5)}{(\lambda_1+\lambda_3)}+\frac{n_3}{\lambda_3}+\frac{n_6}{(\lambda_1+\lambda_2+\lambda_3)}\\ &-\left\{\sum_{i=1}^{n_1}(y_i)^{\alpha}+\sum_{i=1}^{n_2}(x_i)^{\alpha}+\sum_{i=1}^{n_3}(x_i)^{\alpha}+\sum_{i=1}^{n_4}(t^{\alpha}+b^{(k)})+\sum_{i=1}^{n_4}(t^{\alpha}+a^{(k)})+\sum_{i=1}^{n_4}(t^{\alpha}+c^{(k)})\right\}\end{aligned}$$

The above likelihood equations are solved for the maximum likelihood estimates $\left(\lambda_1^{(k)}, \lambda_2^{(k)}, \lambda_3^{(k)}\right)$ using the Newton-Raphson procedure. The observed symmetric information matrix are given below as follows

$$-\frac{\partial^2 Q\left(\underline{\lambda}|\underline{\lambda}^{(k)}\right)}{\partial \lambda_1^2} = \frac{(n_1+n_4)}{(\lambda_1)^2} + \frac{(n_2+n_5)}{(\lambda_1+\lambda_3)^2} + \frac{n_6}{(\lambda_1+\lambda_2+\lambda_3)^2}$$

$$-\frac{\partial^2 Q\left(\underline{\lambda}|\underline{\lambda}^{(k)}\right)}{\partial \lambda_1 \partial \lambda_2} = 0$$

$$-\frac{\partial^2 Q\left(\underline{\lambda}|\underline{\lambda}^{(k)}\right)}{\partial \lambda_1 \partial \lambda_3} = \frac{(n_2+n_5)}{(\lambda_1+\lambda_3)^2} + \frac{n_6}{(\lambda_1+\lambda_2+\lambda_3)^2}$$

$$-\frac{\partial Q\left(\underline{\lambda}|\underline{\lambda}^{(k)}\right)}{\partial \lambda_2^2} = \frac{(n_1+n_4)}{(\lambda_2+\lambda_3)^2} + \frac{(n_2+n_5)}{(\lambda_2)^2} + \frac{n_6}{(\lambda_1+\lambda_2+\lambda_3)^2}$$

$$-\frac{\partial^2 Q\left(\underline{\lambda}|\underline{\lambda}^{(k)}\right)}{\partial \lambda_2 \partial \lambda_3} = \frac{(n_1+n_4)}{(\lambda_2+\lambda_3)^2} + \frac{n_6}{(\lambda_1+\lambda_2+\lambda_3)^2}$$

$$-\frac{\partial Q\left(\underline{\lambda}|\underline{\lambda}^{(k)}\right)}{\partial \lambda_3^2} = \frac{(n_1+n_4)}{(\lambda_2+\lambda_3)^2} + \frac{(n_2+n_5)}{(\lambda_1+\lambda_3)^2} + \frac{n_3}{(\lambda_3)^2} + \frac{n_6}{(\lambda_1+\lambda_2+\lambda_3)^2}$$

5. SIMULATION STUDY AND CONCLUSIONS

The sample data in the first case are generated based on the following algorithms:

- Step 1: generate z_i using the Weibull distribution with $\lambda_1 = 0.08$, $\lambda_2 = 0.1$, $\lambda_3 = 0.0$ for $i = 1,2,3$ and $\alpha = 3$.
- Step 2: take $x = \min(z_1, z_3)$ and $y = \min(z_2, z_3)$ and, therefore, (X, Y) follows the bivariate Weibull distribution of Marshall-Olkin type.
- Step 3: generate t_i using the Weibull distribution with $\theta = 0.03$, where the $t_i s$ are the censoring time.

For two cases with respect to $\lambda_i s$, we generate 1000 set samples. Each set consisted of three samples $n = 50,\ 75, 100$. The corresponding maximum likelihood estimates are displayed in Table 1 together with the estimates and standard error (SE). The estimates denoted by MLE_{em} and SE_{em} are obtained by using the EM algorithm, while the estimates denoted by MLE and SE are obtained without EM algorithm.

The estimates MLE_{em} are close to the true parameter values and the SE_{em} decreases as the sample size increases. The estimates for both methods are obtained by taking the mean of the 1000 maximum likelihood estimates and the biased is the difference between the true parameter and the mean of the estimated parameter from the 1000 samples $n = 50,75,100$.

One can also extend the estimation procedure based on the EM algorithm for other types of bivariate Weibull distribution like those of Freund [3], Block and Basu [1], or Proschan and Sullo [9].

6. APPENDIX: COMPUTATIONS WITHOUT EM ALGORITHM

As stated above the observable life times associated with the *i*th pair of components are given as follows:

$$(X_i, Y_i) = \begin{cases} (X_i, Y_i) & \text{for } \max(X_i, Y_i) < T_i \\ (X_i, T_{iy}) & \text{for } X_i < T_i < T_{iy} = Y_i \\ (T_{ix}, Y_i) & \text{for } Y_i < T_i < T_{ix} = X_i \\ (T_{ixy}, T_{ixy}) & \text{for } T_i < T_{ixy} = \min(X_i, Y_i) \end{cases} \tag{9}$$

The likelihood function for a given sample of size *n* is:

$$L_c(\underline{\lambda}, \alpha | x_i, y_i, t_i) = \prod_{j=1}^{6} \prod_{i=1}^{n_j} f_{X,Y}^{(j)}(x_i, y_i) \tag{10}$$

where

$$f_{X,Y}^{(1)}(x,y) = \lambda_1 \alpha x^{\alpha-1} e^{-\lambda_1 x^\alpha} (\lambda_2+\lambda_3) \alpha y^{\alpha-1} e^{-(\lambda_2+\lambda_3) y^\alpha} \qquad if\ 0 < x < y < \text{t}$$

$$f_{X,Y}^{(2)}(x,y) = (\lambda_1+\lambda_3)\alpha x^{\alpha-1}e^{-(\lambda_1+\lambda_3)x^\alpha}\lambda_2\alpha y^{\alpha-1}e^{-\lambda_2 y^\alpha} \quad if\ \ 0<y<x<t$$
$$f_{X,Y}^{(3)}(x,y) = \lambda_3\,\alpha x^{\alpha-1}e^{-(\lambda_1+\lambda_2+\lambda_3)x^\alpha} \quad if\ \ 0<x=y<t$$
$$f_{X,Y}^{(4)}(x,y) = \lambda_1\,\alpha x^{\alpha-1}e^{-\lambda_1 x^\alpha}e^{-(\lambda_2+\lambda_3)t^\alpha} \quad if\ \ 0<x<t<y$$
$$f_{X,Y}^{(5)}(x,y) = e^{-(\lambda_1+\lambda_3)t^\alpha}\lambda_2\alpha y^{\alpha-1}e^{-\lambda_2 y^\alpha} \quad if\ \ 0<y<t<x$$
$$f_{X,Y}^{(6)}(x,y) = e^{-(\lambda_1+\lambda_2+\lambda_3)t^\alpha} \quad if\ \ t<t_i=min(x,y)$$

where $f_{X,Y}^{(1)}, f_{X,Y}^{(2)}$ are density functions with respect to the Lebesque measure on R^2, while $f_{X,Y}^{(3)}, f_{X,Y}^{(4)}$ and $f_{X,Y}^{(6)}$ are density functions with respect to The Lebesque measure on R.

$\theta = 0.03\ \alpha = 3$, $\theta = 0.03\ \ \alpha = 2$

	Case 1			Case 1		
Parameters	λ_1	λ_2	λ_3	λ_1	λ_2	λ_3
	0.08	0.1	0.03	0.1	0.12	0.04
	n=50					
MLE_{em}	0.081796	0.101429	0.029931	0.102119	0.122057	0.040519
SE_{em}	0.0003393	0.000363	0.00022	0.0004232	0.0004413	0.0002811
MLE	0.081639	0.102724	0.030519	0.102027	0.123128	0.040752
SE	0.0003427	0.0003841	0.00022	0.000425	0.0004623	0.0002815
	n=75					
MLE_{em}	0.081	0.099046	0.030233	0.101213	0.120058	0.040519
SE_{em}	0.0001794	0.0001925	0.0001188	0.0002189	0.0002431	0.000151
MLE	0.080878	0.100963	0.030514	0.101029	0.122003	0.040399
SE	0.0001797	0.0002141	0.0001219	0.0002199	0.0002564	0.0001512
	n=100					
MLE_{em}	0.080548	0.099368	0.03021	0.100885	0.120494	0.040198
SE_{em}	0.0001144	0.0001257	0.000079	0.0001407	0.0001586	0.0000987
MLE	0.080532	0.101133	0.030445	0.100789	0.121521	0.04046
SE	0.000115	0.0001383	0.000079	0.0001414	0.0001666	0.0000989

Table 1: Comparison of MLEs obtained using the EM algorithm and without EM algorithm.

The log-likelihood is given by:

$$lnL_c(\underline{\lambda},\alpha|x_i,y_i,t_i) = (n_1+n_4)ln(\lambda_1)+(n_1)ln(\lambda_2+\lambda_3)+(n_2+n_5)ln(\lambda_2)+(n_2)ln(\lambda_1+\lambda_3)$$
$$+n_3 ln(\lambda_3)+(2n_1+2n_2+n_3+n_4+n_5)ln(\alpha)$$
$$+(\alpha-1)\left\{\sum_{i=1}^{n_1}ln(x_i)+\sum_{i=1}^{n_1}ln(y_i)+\sum_{i=1}^{n_2}ln(x_i)+\sum_{i=1}^{n_2}ln(y_i)+\sum_{i=1}^{n_3}ln(x_i)+\sum_{i=1}^{n_4}ln(x_i)+\sum_{i=1}^{n_5}ln(y_i)\right\}$$
$$-\lambda_1\left\{\sum_{i=1}^{n_1}(x_i)^\alpha+\sum_{i=1}^{n_2}(x_i)^\alpha+\sum_{i=1}^{n_3}(x_i)^\alpha+\sum_{i=1}^{n_4}(x_i)^\alpha+\sum_{i=1}^{n_5}(t_i)^\alpha+\sum_{i=1}^{n_6}(t_i)^\alpha\right\}$$
$$-\lambda_2\left\{\sum_{i=1}^{n_1}(y_i)^\alpha+\sum_{i=1}^{n_2}(y_i)^\alpha+\sum_{i=1}^{n_3}(x_i)^\alpha+\sum_{i=1}^{n_4}(t_i)^\alpha+\sum_{i=1}^{n_5}(y_i)^\alpha+\sum_{i=1}^{n_6}(t_i)^\alpha\right\}$$
$$-\lambda_3\left\{\sum_{i=1}^{n_1}(y_i)^\alpha+\sum_{i=1}^{n_2}(x_i)^\alpha+\sum_{i=1}^{n_3}(x_i)^\alpha+\sum_{i=1}^{n_4}(t_i)^\alpha+\sum_{i=1}^{n_5}(t_i)^\alpha+\sum_{i=1}^{n_6}(t_i)^\alpha\right\}$$

The first order partial derivatives of the log-likelihood can be obtained as follows:

$$\frac{\partial lnL_c(\underline{\lambda},\alpha|x_i,y_i,t_i)}{\partial\lambda_1}=\frac{(n_1+n_4)}{\lambda_1}+\frac{n_2}{(\lambda_1+\lambda_3)}$$
$$-\left\{\sum_{i=1}^{n_1}(x_i)^\alpha+\sum_{i=1}^{n_2}(x_i)^\alpha+\sum_{i=1}^{n_3}(x_i)^\alpha+\sum_{i=1}^{n_4}(x_i)^\alpha+\sum_{i=1}^{n_5}(t_i)^\alpha+\sum_{i=1}^{n_6}(t_i)^\alpha\right\}$$
$$\frac{\partial lnL_c(\underline{\lambda},\alpha|x_i,y_i,t_i)}{\partial\lambda_2}=\frac{n_1}{(\lambda_2+\lambda_3)}+\frac{(n_2+n_5)}{(\lambda_2)}$$
$$-\left\{\sum_{i=1}^{n_1}(y_i)^\alpha+\sum_{i=1}^{n_2}(y_i)^\alpha+\sum_{i=1}^{n_3}(x_i)^\alpha+\sum_{i=1}^{n_4}(t_i)^\alpha+\sum_{i=1}^{n_5}(y_i)^\alpha+\sum_{i=1}^{n_6}(t_i)^\alpha\right\}$$
$$\frac{\partial lnL_c(\underline{\lambda},\alpha|x_i,y_i,t_i)}{\partial\lambda_3}=\frac{n_1}{(\lambda_2+\lambda_3)}+\frac{n_2}{(\lambda_1+\lambda_3)}+\frac{n_3}{\lambda_3}$$
$$-\left\{\sum_{i=1}^{n_1}(y_i)^\alpha+\sum_{i=1}^{n_2}(x_i)^\alpha+\sum_{i=1}^{n_3}(x_i)^\alpha+\sum_{i=1}^{n_4}(t_i)^\alpha+\sum_{i=1}^{n_5}(t_i)^\alpha+\sum_{i=1}^{n_6}(t_i)^\alpha\right\}$$

The second order partial derivatives can be obtained as follows:

$$-\frac{\partial^2 lnL_c(\underline{\lambda},\alpha|x_i,y_i,t_i)}{\partial\lambda_1^2}=\frac{(n_1+n_4)}{(\lambda_1)^2}+\frac{n_2}{(\lambda_1+\lambda_3)^2}$$
$$-\frac{\partial^2 lnL_c(\underline{\lambda},\alpha|x_i,y_i,t_i)}{\partial\lambda_1\partial\lambda_2}=0$$
$$-\frac{\partial^2 lnL_c(\underline{\lambda},\alpha|x_i,y_i,t_i)}{\partial\lambda_1\partial\lambda_3}=\frac{n_2}{(\lambda_1+\lambda_3)^2}$$
$$-\frac{\partial^2 lnL_c(\underline{\lambda},\alpha|x_i,y_i,t_i)}{\partial\lambda_2^2}=\frac{n_1}{(\lambda_2+\lambda_3)^2}+\frac{(n_2+n_5)}{(\lambda_2)^2}$$
$$-\frac{\partial^2 lnL_c(\underline{\lambda},\alpha|x_i,y_i,t_i)}{\partial\lambda_2\partial\lambda_3}=\frac{n_1}{(\lambda_2+\lambda_3)^2}$$
$$-\frac{\partial^2 lnL_c(\underline{\lambda},\alpha|x_i,y_i,t_i)}{\partial\lambda_3^2}=\frac{n_1}{(\lambda_2+\lambda_3)^2}+\frac{n_2}{(\lambda_1+\lambda_3)^2}+\frac{n_3}{(\lambda_3)^2}$$

7. REFERENCES

[1] Block, H. W. and Basu, A. P. (1974). A continuous bivariate exponential extension. J.Amer.Statist.Assoc. 69, 1031-37.

[2] Dempster, A.P., Laird, N.M., and Rubin, D.B. (1977). Maximum likelihood from incomplete data via de EM algorithm (with discussion). Journal of the Royal Statistical Society B 39, 1-38.

[3] Freund, J.E. (1961). A bivariate extension of the exponential distribution. J. Amer. Statist. Assoc. 56, 971-977.

[4] Hanagal, D. D. (1992). Some inference results in bivariate exponential distributions based on censored samples. Communications in Statistics-Theory & Methods 21(5), 1273-1295.

[5] Hanagal, D. D. (1992). Some inference results in modified Freund's bivariate exponential distribution. Biometrical Journal. 34, 745-56.

[6] Hanagal, D.D., Ahmadi, K.A. (2008). Parameter estimation for the bivariate exponential distribution by the EM algorithm based on censored samples. Economic Quality Control. 2, 257-266.

[7] Marshall, A. W. and Olkin, I. (1967). A multivariate exponential distribution. J.Amer.Statist.Assoc. 62, 30-44.

[8] McLachlan, G.J., and Krishnan, T. (1997). The EM algorithm and extensions. John Wiley & Sons, New York.

[9] Proschan, F. and Sullo, P. (1974). Estimating the parameters of a bivariate exponential distributions in several sampling situations. Reliability and Biometry. Edited by F. Proschan and R.J. Surfling, Philadelphia: SIAM; pp. 423-40.