مجله‌ی پژوهش‌های آماری ایران
J. Statist. Res. Iran **16** (2019): 343–378
دوره‌ی ۱۶، شماره‌ی ۲، پاییز و زمستان ۱۳۹۸، صص ۳۴۳–۳۷۸
پژوهشکده‌ی آمار
DOI: 10.52547/jsri.16.2.343
DOR: 20.1001.1.25385771.2020.16.2.6.6

Inference about the Marshal-Olkin Bivariate Burr Type III Distribution under Random Left Censoring

A. Azizi† and A. Sayyareh‡,*

† Razi University
‡ K.N. Toosi University of Technology

Received: 2020/19/10 Approved: 2021/12/02

Abstract. In this paper, a Marshal-Olkin bivariate model based on Burr *III* distribution is considered under random left censoring. The maximum likelihood estimator of the unknown parameters is obtained using the direct method and Expectation Conditional Maximization algorithm. We also obtained the Fisher information matrices. To discuss the properties of the estimators obtained iteratively, a simulation study is carried out. A real data set is used to illustrate the theoretical results.

Keywords. Marshal-Oklin bivariate distribution; Burr lll distribution; ECM algorithm; pseudo likelihood; random left censoring.

MSC 2010: 62H12, 62N02.

1 Introduction

Burr (1942) has developed a system of twelve types of distribution functions based on generating the Pearson differential equation. From the system of Burr distributions, the Burr XII distribution is widely used. The inverse distribution of Burr XII is Burr III. It is more flexible and includes a variety

* Corresponding author

of distributions with varying degrees of skewness and kurtosis. The Burr III distribution with two parameters c and k which is denoted by $BIII(c,k)$ has been used in a variety of setting for the purpose of statistical modeling.
The probability density function and the cumulative distribution function of $BIII(c,k)$ are given by, respectively

$$\begin{aligned} f_{BIII}(x;c,k) &= kcx^{-c-1}(1+x^{-c})^{-k-1}, \qquad x>0, c>0, k>0, \\ F_{BIII}(x;c,k) &= (1+x^{-c})^{-k}, \end{aligned}$$

where c and k are the shape parameters. Various fields of science used the Burr III distribution.
The modeling of lifetime is an important aspect of statistical work in a variety of scientific and technological fields. Many time the life/failure data of interest is bivariate in nature. Any study on twins or on failure data recorded twice on the same system naturally leads to bivariate data. For example, Houggard et al. (1992) studied data on lifelength of Danish twins and Lin et al. (1999) considered a data on patients of colon cancer where the paired data consist of the time from treatment to recurrence of the cancer and the time from treatment to death. Paired data could consist of blindness in the left/right eye, failure time of the left/right kidney or age at death of parent/child in a genetic study. So, In many practical situation, bivariate lifetime data arise frequently, and in these situations it is important to consider different bivariate models that could be used to model such bivariate lifetime data. In the recent years, bivariate lifetime data are often used to model reliability and survival data. For example, Sarhan and Balakrishnan (2007) studied Marshall and Olkin bivariate exponential distribution, Al-Khedhairi and El-Gohary (2008) presented a new class of bivariate Gompertz distributions, Kundu and Gupta (2009) proposed the bivariate generalized exponential distribution, Kundu and Dey (2009) studied an EM algorithm for computing maximum likelihood estimators of the parameters of the bivariate Weibull distribution in the case of complete data, Nandi and Dewan (2010) have considered the maximum likelihood estimators of parameters of bivariate Weibull distribution under random censoring and El-Sherpieny et al. (2013) have introduced a new bivariate generalized Gompertz distribution, Mirhosseini et al. (2015) proposed new absolutely continuous bivariate generalized exponential distribution, Asimit et al. (2016) used EM algorithm to estimate singular Marshall-Olkin bivariate Pareto distribution, Dey et al.

(2018) studied an EM algorithm for absolute continuous bivariate Pareto distribution, Azizi and Sayyareh (2019) presented estimating the Parameters of the Bivariate Burr Type III Distribution by EM Algorithm and Azizi et al (2019) have considered inference about the bivariate new extended Weibull distribution based on complete and censored data.

In many reliability and life-testing studies, the observed failure time data of items are often not wholly available. Lowering the expense and period associated with the tests is important in statistical tests with censored data. Among the censoring method, statistics of random left censored data is a new research field. Random left censoring is a situation when an item under study is lost or removed randomly from the experiment before its failure. In other words, some subjects in the study have not experienced the event of interest at the start of the study. In such cases, the exact survival time (or time to event of interest) of the subjects is unknown; therefore they are called random left censored observations. For some further examples, one may refer to Nandi and Dewan (2010) , Mitra and Kundu(2008), Balakrishnan (1989), Balakrishnan and Varadan (1991), Lee et al. (1980), etc. In bivariate lifetime distributions one or both components of the paired data could be subject to random censoring.

In this paper, we study the maximum likelihood estimators of the parameters of Marshal-Olkin bivariate Burr III distribution under random left censoring. In section 2 we state the joint distribution function and joint density of Burr III distribution. In section 3 we write the likelihood function under random left censoring and discuss the problem of computing the maximum likelihood estimators of the unknown parameters of Marshal-Olkin bivariate Burr III distribution based on a random sample. The observed Fisher information matrix is in section 4. The findings of the numerical experiments are reported in section 5. One real data set is analysed in section 6 and we conclude in section 7. Fisher's information matrices are given in the appendix.

2 Marshal-Olkin Bivariate Burr III Distribution

Consider the Burr III distribution ($BIII(c, k)$) with shape parameters $c > 0$ and $k > 0$. Suppose W_1, W_2, W_3, respectively, are independent $BIII(c, k_1)$, $BIII(c, k_2)$ and $BIII(c, k_3)$, random variables. Let $X_1 = \max(W_1, W_3)$ and $X_2 = \max(W_2, W_3)$. Then the pair (X_1, X_2) has Marshal-Olkin bivariate Burr III distribution, ($MOBBIII$), with shape parameters c, k_1, k_2 and k_3

is expressed as $MOBBIII(c, k_1, k_2, k_3)$, see Azizi and Sayyareh (2019).
Note that if $\max(W_1, W_2, W_3) = U_3$, then two random variable X_1 and X_2 are equal. For example, suppose a system has two components. Each component is subject to individual independent stress say W_1 and W_2 respectively. The system has an overall stress U_3 which has been transmitted to both the components equally, independent of their individual stresses. Therefore, the observed stress at the two components are $X_1 = \max(W_1, W_3)$ and $X_2 = \max(W_2, W_3)$ respectively. Now suppose the overall stress, W_3, is greater than individual stress W_1 and W_2, then $X_1 = X_2 = W_3$.
The joint distribution function of (X_1, X_2) can be written as follows:

$$F(x_1, x_2) = \begin{cases} F_{BIII}(x_1; c, k_1 + k_3) F_{BIII}(x_2; c, k_2) & \text{if} \quad x_1 < x_2 \\ F_{BIII}(x_1; c, k_1) F_{BIII}(x_2; c, k_2 + k_3) & \text{if} \quad x_1 > x_2 \\ F_{BIII}(x_1; c, k_1 + k_2 + k_3) & \text{if} \quad x_1 = x_2 \end{cases}$$

where $F_{BIII}(.)$ is the cumulative distribution function of the Burr III distribution.
The joint density function of (X_1, X_2) can be written as follows:

$$f(x_1, x_2) = \begin{cases} f_{BIII}(x_1; c, k_1 + k_3) f_{BIII}(x_2; c, k_2) & \text{if} \quad x_1 < x_2 \\ f_{BIII}(x_1; c, k_1) f_{BIII}(x_2; c, k_2 + k_3) & \text{if} \quad x_1 > x_2 \\ \frac{k_3}{k_1+k_2+k_3} f_{BIII}(x_1; c, k_1 + k_2 + k_3) & \text{if} \quad x_1 = x_2 \end{cases}$$

The following quantities are required to express the likelihood explicitly

$$\int_0^{x_2} f(x_1, u) du = \begin{cases} f_{BIII}(x_1; c, k_1 + k_3)[F_{BIII}(x_2; c, k_2) - F_{BIII}(x_1; c, k_2)] & \text{if} \quad x_1 < x_2 \\ f_{BIII}(x_1; c, k_1) F_{BIII}(x_2; c, k_2 + k_3) & \text{if} \quad x_1 > x_2 \\ f_{BIII}(x_1; c, k_1) F_{BIII}(x_1; c, k_2 + k_3) & \text{if} \quad x_1 = x_2 \end{cases}$$

and

$$\int_0^{x_1} f(\nu, x_2) d\nu = \begin{cases} F_{BIII}(x_1; c, k_1 + k_3) f_{BIII}(x_2; c, k_2) & \text{if} \quad x_1 < x_2 \\ f_{BIII}(x_2; c, k_2 + k_3)[F_{BIII}(x_1; c, k_1) - F_{BIII}(x_2; c, k_1)] & \text{if} \quad x_1 > x_2 \\ F_{BIII}(x_1; c, k_1 + k_3) f_{BIII}(x_1; c, k_2) & \text{if} \quad x_1 = x_2 \end{cases}$$

Suppose the pair (X_1, X_2) is subject to random left censoring by pair of random variables (Y_1, Y_2) which independent from (X_1, X_2). We observe $(T_1, \delta_1, T_2, \delta_2)$ where $T_1 = \max(X_1, Y_1), \delta_1 = I(X_1 > Y_1)$ and $T_2 = \max(X_2, Y_2), \delta_2 = I(X_2 > Y_2)$. Therefore, if $X_i < Y_i$, X_i for $i = 1, 2$ is censored.

3 Maximum Likelihood Estimation

In this section we address the problem of computing the maximum likelihood estimators (MLEs) of the unknown parameters of $MOBBIII$ distribution based on a random sample. In order to write down the likelihood, we note that, when $\delta_1 = \delta_2 = 1$, both failure times are observed and the contribution to the likelihood is $f(t_1, t_2)$. When $\delta_1 = 1 - \delta_2 = 1$, first component fails at t_1 and the second component is censored (fails before t_2) and the contribution to the likelihood is $\int_0^{t_2} f(t_1, x_2)dx_2$. Similarly, when $1 - \delta_1 = \delta_2 = 1$, first component is censored (fails before t_1) and the second component fails at t_2 and the contribution to the likelihood is $\int_0^{t_1} f(x_1, t_2)dx_1$. Finally, when $1 - \delta_1 = 1 - \delta_2 = 1$, both failure times are censored and the contribution to the likelihood is $F(t_1, t_2)$. Hence, the likelihood function, based on $(T_{1i}, \delta_{1i}, T_{2i}, \delta_{2i})$, $i = 1, 2, ..., n$ is given by

$$\begin{aligned} L &= \prod_{i=1}^{n} L(t_{1i}, \delta_{1i}, t_{2i}, \delta_{2i}) \\ &= \prod_{i=1}^{n} \big[f(t_{1i}, t_{2i})\big]^{\delta_{1i}\delta_{2i}} \big[\int_0^{t_{2i}} f(t_{1i}, x_2)dx_2\big]^{\delta_{1i}(1-\delta_{2i})} \\ &\quad \big[\int_0^{t_{1i}} f(x_1, t_{2i})dx_1\big]^{(1-\delta_{1i})\delta_{2i}} \big[F(t_{1i}, t_{2i})\big]^{(1-\delta_{1i})(1-\delta_{2i})} \end{aligned}$$

Let I_0, I_1, I_2, denote the following sets

$$I_0 = \{i | t_{1i} = t_{2i} = t_i\} \qquad I_1 = \{i | t_{1i} < t_{2i}\} \qquad I_2 = \{i | t_{1i} > t_{2i}\}$$

Let $L_k(t_{1i}, \delta_{1i}, t_{2i}, \delta_{2i}) \equiv L_k(c, k_1, k_2, k_3; t_{1i}, \delta_{1i}, t_{2i}, \delta_{2i})$ is the contribution from I_k to the likelihood function, $k = 0, 1, 2$, they are given in Appendix A explicitly. Based on the above notation the likelihood function becomes

$$L = \prod_{i \in I_0} L_0(t_i, \delta_{1i}, \delta_{2i}) \prod_{i \in I_1} L_1(t_{1i}, \delta_{1i}, t_{2i}, \delta_{2i}) \prod_{i \in I_2} L_2(t_{1i}, \delta_{1i}, t_{2i}, \delta_{2i}) \qquad (1)$$

Let n_0, n_1, and n_2, respectively, denote the number of elements in the sets I_0, I_1, and I_2 and n_{ij} be the number of pairs for which $(\delta_1, \delta_2) = (i, j), i, j = 0, 1$. Then,

$$n = \sum_{i=0}^{1}\sum_{j=0}^{1} n_{ij} \quad \text{and} \quad n_k = \sum_{i=0}^{1}\sum_{j=0}^{1} n_{ij}^{k}, \quad k = 0, 1, 2$$

where n_{ij}^{k} denotes the number of individuals in I_k with $(\delta_1, \delta_2) = (i, j), i, j = 0, 1$ and $k = 0, 1, 2$.

We need to maximize (1) with respect to the four unknown parameters. The contribution to the pseudo-likelihood from I_0 is;

$$\begin{aligned}\prod_{i\in I_0} L_0(t_i, \delta_{1i}, \delta_{2i}) &= \prod_{i\in I_0} \big[\frac{k_3}{k_1+k_2+k_3} f(t_i; c, k_1+k_2+k_3)\big]^{\delta_{1i}\delta_{2i}} \\ &\quad \big[f(t_i; c, k_1) F(t_i; c, k_2+k_3)\big]^{\delta_{1i}(1-\delta_{2i})} \\ &\quad \big[F(t_i; c, k_1+k_3) f(t_i; c, k_2)\big]^{(1-\delta_{1i})\delta_{2i}} \\ &\quad \big[F(t_i; c, k_1+k_2+k_3)\big]^{(1-\delta_{1i})(1-\delta_{2i})} \\ &= \prod_{i\in I_0} \big[k_3 c t_i^{-c-1}(1+t_i^{-c})^{-k_1-k_2-k_3-1}\big]^{\delta_{1i}\delta_{2i}} \\ &\quad \big[k_1 c t_i^{-c-1}(1+t_i^{-c})^{-k_1-k_2-k_3-1}\big]^{\delta_{1i}(1-\delta_{2i})} \\ &\quad \big[k_2 c t_i^{-c-1}(1+t_i^{-c})^{-k_1-k_2-k_3-1}\big]^{(1-\delta_{1i})\delta_{2i}} \\ &\quad \big[(1+t_i^{-c})^{-k_1-k_2-k_3}\big]^{(1-\delta_{1i})(1-\delta_{2i})} \qquad (2)\end{aligned}$$

so; the pseudo log-likelihood from I_0 is given as follows;

$$\begin{aligned}&\sum_{i\in I_0} \delta_{1i}\delta_{2i}\Big[\log k_3 + \log c - (c+1)\log t_i - (k_1+k_2+k_3+1)\log(1+t_i^{-c})\Big] \\ &\quad +\delta_{1i}(1-\delta_{2i})\Big[\log k_1 + \log c - (c+1)\log t_i - (k_1+k_2+k_3+1)\log(1+t_i^{-c})\Big] \\ &\quad +(1-\delta_{1i})\delta_{2i}\Big[\log k_2 + \log c - (c+1)\log t_i - (k_1+k_2+k_3+1)\log(1+t_i^{-c})\Big] \\ &\quad -(1-\delta_{1i})(1-\delta_{2i})\Big[(k_1+k_2+k_3)\log(1+t_i^{-c})\Big]\end{aligned}$$

The contribution to the pseudo-likelihood from I_1 is

$$\begin{aligned}\prod_{i\in I_1} L_1(t_{1i},\delta_{1i},t_{2i},\delta_{2i}) &= \prod_{i\in I_1}\big[f(t_{1i};c,k_1+k_3)f(t_{2i};c,k_2)\big]^{\delta_{1i}\delta_{2i}}\\ &\quad \big[f(t_{1i};c,k_1+k_3)\big[F(t_{2i};c,k_2)-F(t_{1i};c,k_2)\big]\big]^{\delta_{1i}(1-\delta_{2i})}\\ &\quad \big[F(t_{1i};c,k_1+k_3)f(t_{2i};c,k_2)\big]^{(1-\delta_{1i})\delta_{2i}}\\ &\quad \big[F(t_{1i};c,k_1+k_3)F(t_{2i};c,k_2)\big]^{(1-\delta_{1i})(1-\delta_{2i})} \qquad (3)\end{aligned}$$

Further, the pseudo-log-likelihood from I_1 is given as follows

$$\begin{aligned}&\sum_{i\in I_1}\delta_{1i}\delta_{2i}\Big[\log(k_1+k_3)+\log c-(c+1)\log t_{1i}-(k_1+k_3+1)\log(1+t_{1i}^{-c})\\ &\quad +\log k_2+\log c-(c+1)\log t_{2i}-(k_2+1)\log(1+t_{2i}^{-c})\Big]\\ &\quad +\delta_{1i}(1-\delta_{2i})\Big[\log(k_1+k_3)+\log c-(c+1)\log t_{1i}-(k_1+k_3+1)\log(1+t_{1i}^{-c})\\ &\quad +\log\big[(1+t_{2i}^{-c})^{-k_2}-(1+t_{1i}^{-c})^{-k_2}\big]\Big]\\ &\quad -(1-\delta_{1i})\delta_{2i}\Big[(k_1+k_3)\log(1+t_{1i}^{-c})-\log k_2-\log c+(c+1)\log t_{2i}\\ &\quad +(k_2+1)\log(1+t_{2i}^{-c})\Big]\\ &\quad -(1-\delta_{1i})(1-\delta_{2i})\Big[(k_1+k_3)\log(1+t_{1i}^{-c-1})+k_2\log(1+t_{2i}^{-c})\Big].\end{aligned}$$

Finally, the contribution to the pseudo-likelihood from I_2 is

$$\begin{aligned}\prod_{i\in I_2} L_2(t_{1i},\delta_{1i},t_{2i},\delta_{2i}) &= \prod_{i\in I_2}\big[f(t_{1i};c,k_1)f(t_{2i};c,k_2+k_3)\big]^{\delta_{1i}\delta_{2i}}\\ &\quad \big[f(t_{1i};c,k_1)F(t_{2i};c,k_2+k_3)\big]^{\delta_{1i}(1-\delta_{2i})}\\ &\quad \big[f(t_{2i};c,k_2+k_3)\big[F(t_{1i};c,k_1)-F(t_{2i};c,k_1)\big]\big]^{(1-\delta_{1i})\delta_{2i}}\\ &\quad \big[F(t_{1i};c,k_1)F(t_{2i};c,k_2+k_3)\big]^{(1-\delta_{1i})(1-\delta_{2i})} \qquad (4)\end{aligned}$$

Further, the pseudo-log-likelihood from I_2 is given as follows

$$
\begin{aligned}
&\sum_{i\in I_2}\delta_{1i}\delta_{2i}\Big[\log k_1+\log c-(c+1)\log t_{1i}-(k_1+1)\log(1+t_{1i}^{-c})\\
&\quad+\log(k_2+k_3)+\log c-(c+1)\log t_{2i}-(k_2+k_3+1)\log(1+t_{2i}^{-c})\Big]\\
&\quad-\delta_{1i}(1-\delta_{2i})\Big[k_1\log(1+t_{1i}^{-c})-\log(k_2+k_3)-\log c+(c+1)\log t_{2i}\\
&\quad+(k_2+k_3+1)\log(1+t_{2i}^{-c})\Big]\\
&\quad+(1-\delta_{1i})\delta_{2i}\Big[\log(k_2+k_3)+\log c-(c+1)\log t_{2i}-(k_2+k_3+1)\log(1+t_{2i}^{-c})\\
&\quad+\log\big[(1+t_{1i}^{-c})^{-k_1}-(1+t_{2i}^{-c})^{-k_1}\big]\Big]\\
&\quad-(1-\delta_{1i})(1-\delta_{2i})\Big[k_1\log(1+t_{1i}^{-c})+(k_2+k_3)\log(1+t_{2i}^{-c})\Big].
\end{aligned}
$$

Hence, the pseudo-log-likelihood is given by

$$
\begin{aligned}
\ell &= M_0\log c+M_1\log k_1+M_2\log k_2\\
&\quad+M_{13}\log(k_1+k_3)+M_{23}\log(k_2+k_3)+n_{11}^0\log k_3\\
&\quad-k_1\big[\sum_{i\in I_0}\log(1+t_i^{-c})+\sum_{i\in I_1}\log(1+t_{1i}^{-c})+\sum_{i\in I_2}(1-\delta_{2i}+\delta_{1i}\delta_{2i})\log(1+t_{1i}^{-c})\big]\\
&\quad-k_2\big[\sum_{i\in I_0}\log(1+t_i^{-c})+\sum_{i\in I_1}(1-\delta_{1i}+\delta_{1i}\delta_{2i})\log(1+t_{2i}^{-c})+\sum_{i\in I_2}\log(1+t_{2i}^{-c})\big]\\
&\quad-k_3\big[\sum_{i\in I_0}\log(1+t_i^{-c})+\sum_{i\in I_1}\log(1+t_{1i}^{-c})+\sum_{i\in I_2}\log(1+t_{2i}^{-c})\big]\\
&\quad-(c+1)\big[\sum_{i\in I_0}(\delta_{1i}+\delta_{2i}-\delta_{1i}\delta_{2i})\log t_i+\sum_{i\in I_1\cup I_2}(\delta_{1i}\log t_{1i}+\delta_{2i}\log t_{2i})\big]\\
&\quad-\sum_{i\in I_0}(\delta_{1i}+\delta_{2i}-\delta_{1i}\delta_{2i})\log(1+t_i^{-c})\\
&\quad-\sum_{i\in I_1\cup I_2}\big[\delta_{1i}\log(1+t_{1i}^{-c})+\delta_{2i}\log(1+t_{2i}^{-c})\big]\\
&\quad+\sum_{i\in I_1}\delta_{1i}(1-\delta_{2i})\log\big[(1+t_{2i}^{-c})^{-k_2}-(1+t_{1i}^{-c})^{-k_2}\big]\\
&\quad+\sum_{i\in I_2}(1-\delta_{1i})\delta_{2i}\log\big[(1+t_{1i}^{-c})^{-k_1}-(1+t_{2i}^{-c})^{-k_1}\big],
\end{aligned}
\tag{5}
$$

where

$$n=\sum_{i=0}^{1}\sum_{j=0}^{1}n_{ij},\qquad n_k=\sum_{i=0}^{1}\sum_{j=0}^{1}n_{ij}^{k},\quad k=0,1,2$$

and

$$\begin{aligned}
M_0 &= n_{11}^0+n_{10}^0+n_{01}^0+2n_{11}^1+n_{10}^1+n_{01}^1+2n_{11}^2+n_{10}^2+n_{01}^2\\
M_1 &= n_{10}^0+n_{11}^2+n_{10}^2;\qquad M_2=n_{01}^0+n_{11}^1+n_{01}^1\\
M_{13} &= n_{11}^1+n_{10}^1;\qquad M_{23}=n_{11}^2+n_{01}^2.
\end{aligned}$$

We need to maximize the pseudo-log-likelihood equation w.r.t. c, k_1, k_2, and k_3. We denote the first derivatives of the pseudo-log-likelihood function as

$$\begin{aligned}
\frac{\partial\ell}{\partial c} = \frac{M_0}{c} &+ (k_1+k_2+k_3)\sum_{i\in I_0}\frac{\ln t_i t_i^{-c}}{1+t_i^{-c}}+(k_1+k_3)\sum_{i\in I_1}\frac{\ln t_{1i}t_{1i}^{-c}}{1+t_{1i}^{-c}}\\
&+(k_2+k_3)\sum_{i\in I_2}\frac{\ln t_{2i}t_{2i}^{-c}}{1+t_{2i}^{-c}}+k_2\sum_{i\in I_1}(1-\delta_{1i}+\delta_{1i}\delta_{2i})\frac{\ln t_{2i}t_{2i}^{-c}}{1+t_{2i}^{-c}}\\
&+k_1\sum_{i\in I_2}(1-\delta_{2i}+\delta_{1i}\delta_{2i})\frac{\ln t_{1i}t_{1i}^{-c}}{1+t_{1i}^{-c}}-\sum_{i\in I_0}(\delta_{1i}+\delta_{2i}-\delta_{1i}\delta_{2i})\log t_i\\
&-\sum_{i\in I_1\cup I_2}[\delta_{1i}\log t_{1i}+\delta_{2i}\log t_{2i}]+\sum_{i\in I_0}(\delta_{1i}+\delta_{2i}-\delta_{1i}\delta_{2i})\frac{\ln t_i t_i^{-c}}{1+t_i^{-c}}\\
&+\sum_{i\in I_1\cup I_2}[\frac{\delta_{1i}\ln t_{1i}t_{1i}^{-c}}{1+t_{1i}^{-c}}+\frac{\delta_{2i}\ln t_{2i}t_{2i}^{-c}}{1+t_{2i}^{-c}}]\\
&+\sum_{i\in I_1}\delta_{1i}(1-\delta_{2i})\{\frac{k_2\ln t_{2i}t_{2i}^{-c}(1+t_{2i}^{-c})^{-k_2-1}-k_2\ln t_{1i}t_{1i}^{-c}(1+t_{1i}^{-c})^{-k_2-1}}{(1+t_{2i}^{-c})^{-k_2}-(1+t_{1i}^{-c})^{-k_2}}\}\\
&+\sum_{i\in I_2}(1-\delta_{1i})\delta_{2i}\{\frac{k_1\ln t_{1i}t_{1i}^{-c}(1+t_{1i}^{-c})^{-k_1-1}-k_1\ln t_{2i}t_{2i}^{-c}(1+t_{2i}^{-c})^{-k_1-1}}{(1+t_{1i}^{-c})^{-k_1}-(1+t_{2i}^{-c})^{-k_1}}\}
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial \ell}{\partial k_1} &= \frac{M_1}{k_1} + \frac{M_{13}}{k_1+k_3} \\
&\quad -\big[\sum_{i\in I_0}\log(1+t_i^{-c}) + \sum_{i\in I_1}\log(1+t_{1i}^{-c}) - \sum_{i\in I_2}(1-\delta_{2i}+\delta_{1i}\delta_{2i})\log(1+t_{1i}^{-c})\big] \\
&\quad -\sum_{i\in I_2}(1-\delta_{1i})\delta_{2i}\big\{\frac{\ln(1+t_{1i}^{-c})(1+t_{1i}^{-c})^{-k_1} - \ln(1+t_{2i}^{-c})(1+t_{2i}^{-c})^{-k_1}}{(1+t_{1i}^{-c})^{-k_1} - (1+t_{2i}^{-c})^{-k_1}}\big\}, \\
\frac{\partial \ell}{\partial k_2} &= \frac{M_2}{k_2} + \frac{M_{23}}{k_2+k_3} \\
&\quad -\big[\sum_{i\in I_0}\log(1+t_i^{-c}) + \sum_{i\in I_2}\log(1+t_{2i}^{-c}) - \sum_{i\in I_1}(1-\delta_{1i}+\delta_{1i}\delta_{2i})\log(1+t_{2i}^{-c})\big] \\
&\quad -\sum_{i\in I_1}\delta_{1i}(1-\delta_{2i})\big\{\frac{\ln(1+t_{2i}^{-c})(1+t_{2i}^{-c})^{-k_2} - \ln(1+t_{1i}^{-c})(1+t_{1i}^{-c})^{-k_2}}{(1+t_{2i}^{-c})^{-k_2} - (1+t_{1i}^{-c})^{-k_2}}\big\}, \\
\frac{\partial \ell}{\partial k_3} &= \frac{n_{11}^0}{k_3} + \frac{M_{23}}{k_2+k_3} + \frac{M_{13}}{k_1+k_3} \\
&\quad -\big[\sum_{i\in I_0}\log(1+t_i^{-c}) + \sum_{i\in I_1}\log(1+t_{1i}^{-c}) + \sum_{i\in I_2}\log(1+t_{2i}^{-c})\big]
\end{aligned}
$$

3.1 ECM Alghorithm

In this case we suggest EM algorithm to compute the MLEs of the unknown parameters. We treat this problem as missing value problem. Assume that for the bivariate random vector (X_1, X_2), there is an associated random vector (Δ_1, Δ_2) as follows:

$$
\Delta_1 = \begin{cases} 1 & \text{if} \quad W_1 > W_3 \\ 3 & \text{if} \quad W_1 < W_3 \end{cases} \quad \text{and} \quad \Delta_2 = \begin{cases} 2 & \text{if} \quad W_2 > W_3 \\ 3 & \text{if} \quad W_2 < W_3 \end{cases}
$$

- If $(x_{1i}, x_{2i}) \in I_0$, then $\Delta_1 = \Delta_2 = 3$.
- If $(x_{1i}, x_{2i}) \in I_1$, Δ_1 is unknown and Δ_2 is known and the possible values of (Δ_1, Δ_2) are $(1, 2)$ or $(3, 2)$,
- If $(x_{1i}, x_{2i}) \in I_2$, then Δ_1 is known and Δ_2 is unknown the possible values of (Δ_1, Δ_2) are $(1, 2)$ or $(1, 3)$ with non-zero probability.

Now we provide the E-step and M-step of the ECM algorithm. In the E-step we treat the observations belong to I_0 as complete observation and keep them intact. If the observation belong to I_1 or I_2, we treat it as a missing observation. If $(x_{1i}, x_{2i}) \in I_1$, we form the pseudo-observation by dividing $(x_1, x_2, \mu_1(\gamma))$ and $(x_1, x_2, \mu_2(\gamma))$, respectively. Here $\gamma = (c, k_1, k_2, k_3)$ and the fractional mass $\mu_1(\gamma), \mu_2(\gamma)$ assigned to the pseudo-observation is the conditional probability that the random vector(Δ_1, Δ_2) takes the values $(1, 2)$ and $(3, 2)$, respectively, given $X_1 < X_2$. Similarly, if $(x_{1i}, x_{2i}) \in I_2$, we form the pseudo-observation of the form $(x_1, x_2, \nu_1(\gamma))$ and $(x_1, x_2, \nu_2(\gamma))$. Here the fractional mass $\nu_1(\gamma)$ or $\nu_2(\gamma)$ assigned to the pseudo-observation is the conditional probability that the random vector (Δ_1, Δ_2) takes the values $(1, 2)$ and $(1, 3)$, respectively, given $X_2 < X_1$. Since

$$
\begin{aligned}
P(W_1 < W_3 < W_2) &= \frac{k_2\, k_3}{(k_1 + k_2 + k_3)\,(k_1 + k_3)},\\
P(W_3 < W_1 < W_2) &= \frac{k_2\, k_1}{(k_1 + k_2 + k_3)\,(k_1 + k_3)},\\
P(W_2 < W_3 < W_1) &= \frac{k_1\, k_3}{(k_1 + k_2 + k_3)\,(k_2 + k_3)},\\
P(W_3 < W_2 < W_1) &= \frac{k_1\, k_2}{(k_1 + k_2 + k_3)\,(k_2 + k_3)},
\end{aligned}
$$

therefore

$$
\begin{aligned}
\mu_1(\gamma) &= P(W_1 < W_3 < W_2 | X_1 < X_2) = \frac{k_3}{K_1 + k_3}\\
\mu_2(\gamma) &= P(W_3 < W_1 < W_2 | X_1 < X_2) \frac{k_1}{k_1 + k_3},
\end{aligned} \tag{6}
$$

$$
\begin{aligned}
\nu_1(\gamma) &= P(W_2 < W_3 < W_1 | X_1 > X_2) = \frac{k_3}{k_2 + k_3},\\
\nu_2(\gamma) &= P(W_3 < W_2 < W_1 | X_1 > X_2) = \frac{k_2}{k_2 + k_3}.
\end{aligned} \tag{7}
$$

From now on, we write $\mu_1(\gamma), \mu_2(\gamma), \nu_1(\gamma)$ and $\nu_2(\gamma)$ as μ_1, μ_2, ν_1 and ν_2, respectively.
In order to identify all parameters uniquely, we write the E-step of the algorithm as follows. We form a pseudo-likelihood by replacing the log-likelihood contribution of the observed $(T_1, \delta_1, T_2, \delta_2)$ by its expected value. The log-

likelihood function of the pseudo-data has three parts corresponding to contributions from the sets I_0, I_1, I_2.
The contribution to the pseudo-log-likelihood from I_0 is

$$\sum_{i\in I_0}\Bigg[\delta_{1i}\delta_{2i}\Bigg\{\log c+\log k_3-(c+1)\log t_i-(k+1)\log(1+t_i^{-c})\Bigg\}$$
$$+\delta_{1i}(1-\delta_{2i})\Bigg\{\log c+\log k_1-(c+1)\log t_i-(k+1)\log(1+t_i^{-c})\Bigg\}$$
$$+(1-\delta_{1i})\delta_{2i}\Bigg\{\log c+\log k_2-(c+1)\log t_i-(k+1)\log(1+t_i^{-c})\Bigg\}$$
$$+(1-\delta_{1i})(1-\delta_{2i})\Bigg\{k\log(1+t_i^{-c})\Bigg\}\Bigg]$$

Further, given the order U_i for $i=1,2,3$, the contribution to the pseudo-log-likelihood from I_1 is given as follows

$$\log\prod_{i\in I_1}L(c,k_1,k_2,k_3)=\sum_{i\in I_1}\delta_{1i}\delta_{2i}\Bigg\{\log\Big[\big[f(t_{1i},c,k_1)F(t_{1i},c,k_3)f(t_{2i},c,k_2)\big]\mu_1$$
$$+\big[F(t_{1i},c,k_1)f(t_{1i},c,k_3)f(t_{2i},c,k_2)\big]\mu_2\Big]\Bigg\}$$
$$+\delta_{1i}(1-\delta_{2i})\Bigg\{\log\Big[f(t_{1i},c,k_1)F(t_{1i},c,k_3)\big[F(t_{2i},c,k_2)-F(t_{1i},c,k_2)\big]\Big]\mu_1$$
$$+\big[F(t_{1i},c,k_1)f(t_{1i},c,k_3)\big[F(t_{2i},c,k_2)-F(t_{1i},c,k_2)\big]\big]\mu_2\Big]\Bigg\}$$
$$+(1-\delta_{1i})\delta_{2i}\Bigg\{\log\Big[F(t_{1i},c,k_1)F(t_{1i},c,k_3)f(t_{2i},c,k_2)\Big]\Bigg\}$$
$$+(1-\delta_{1i})(1-\delta_{2i})\Bigg\{\log\Big[F(t_{1i},c,k_1)F(t_{1i},c,k_3)F(t_{2i},c,k_2)\Big]\Bigg\},$$

where μ_1 and μ_2 are defined as (6);

$$
\begin{aligned}
&\sum_{i\in I_1}\delta_{1i}\delta_{2i}\Big\{\Big[2\log c+\log k_2+\log k_3-(k_2+1)\log(1+t_{2i}^{-c})\\
&-(k_1+k_3+1)\log(1+t_{1i}^{-c})-(c+1)\log(t_{1i}t_{2i})\Big]\mu_1\\
&+\Big[2\log c+\log k_2+\log k_1-(k_2+1)\log(1+t_{2i}^{-c})\\
&-(k_1+k_3+1)\log(1+t_{1i}^{-c})-(c+1)\log(t_{1i}t_{2i})\Big]\mu_2\Big\}\\
&+\delta_{1i}(1-\delta_{2i})\Big\{\Big[\log c+\log k_3-(c+1)\log t_{1i}-(k_1+k_3+1)\log(1+t_{1i}^{-c})\\
&+\log[(1+t_{2i}^{-c})^{-k_2}-(1+t_{1i}^{-c})^{-k_2}]\Big]\mu_1
\end{aligned}
$$

$$
\begin{aligned}
&+\Big[\log c+\log k_1-(c+1)\log t_{1i}-(k_1+k_3+1)\log(1+t_{1i}^{-c})\\
&+\log[(1+t_{2i}^{-c})^{-k_2}-(1+t_{1i}^{-c})^{-k_2}]\Big]\mu_2\Big\}\\
&+(1-\delta_{1i})\delta_{2i}\Big\{\log c+\log k_2-(c+1)\log t_{2i}-(k_2+1)\log(1+t_{2i}^{-c})\\
&-(k_1+k_3)\log(1+t_{1i}^{-c})\Big\}\\
&+(1-\delta_{1i})(1-\delta_{2i})\Big\{(k_1+k_3)\log(1+t_{1i}^{-c})+k_2\log(1+t_{2i}^{-c})\Big\},
\end{aligned}
$$

and given the order U_i for $i=1,2,3$, the contribution to the pseudo-log-likelihood from I_2 is given as follows

$$
\begin{aligned}
\log\prod_{i\in I_1}L(c,k_1,k_2,k_3)=&\sum_{i\in I_1}\delta_{1i}\delta_{2i}\Big\{\log\Big[\big[f(t_{1i},c,k_1)F(t_{1i},c,k_3)f(t_{2i},c,k_2)\big]\mu_1\\
&+\big[F(t_{1i},c,k_1)f(t_{1i},c,k_3)f(t_{2i},c,k_2)\big]\mu_2\Big]\Big\}\\
&+\delta_{1i}(1-\delta_{2i})\Big\{\log\big[f(t_{1i},c,k_1)F(t_{1i},c,k_3)\big[F(t_{2i},c,k_2)-F(t_{1i},c,k_2)\big]\big]\mu_1\\
&+\big[F(t_{1i},c,k_1)f(t_{1i},c,k_3)\big[F(t_{2i},c,k_2)-F(t_{1i},c,k_2)\big]\big]\mu_2\Big]\Big\}\\
&+(1-\delta_{1i})\delta_{2i}\Big\{\log\big[F(t_{1i},c,k_1)F(t_{1i},c,k_3)f(t_{2i},c,k_2)\big]\Big\}\\
&+(1-\delta_{1i})(1-\delta_{2i})\Big\{\log\big[F(t_{1i},c,k_1)F(t_{1i},c,k_3)F(t_{2i},c,k_2)\big]\Big\},
\end{aligned}
$$

where μ_1 and μ_2 are defined as (6);

$$\sum_{i\in I_1}\delta_{1i}\delta_{2i}\bigg\{\Big[2\log c+\log k_2+\log k_3-(k_2+1)\log(1+t_{2i}^{-c})$$
$$-(k_1+k_3+1)\log(1+t_{1i}^{-c})-(c+1)\log(t_{1i}t_{2i})\Big]\mu_1$$
$$+\Big[2\log c+\log k_2+\log k_1-(k_2+1)\log(1+t_{2i}^{-c})$$
$$-(k_1+k_3+1)\log(1+t_{1i}^{-c})-(c+1)\log(t_{1i}t_{2i})\Big]\mu_2\bigg\}$$
$$+\delta_{1i}(1-\delta_{2i})\bigg\{\Big[\log c+\log k_3-(c+1)\log t_{1i}-(k_1+k_3+1)\log(1+t_{1i}^{-c})$$

$$+\log[(1+t_{2i}^{-c})^{-k_2}-(1+t_{1i}^{-c})^{-k_2}]\Big]\mu_1$$
$$+\Big[\log c+\log k_1-(c+1)\log t_{1i}-(k_1+k_3+1)\log(1+t_{1i}^{-c})$$
$$+\log[(1+t_{2i}^{-c})^{-k_2}-(1+t_{1i}^{-c})^{-k_2}]\Big]\mu_2\bigg\}$$
$$+(1-\delta_{1i})\delta_{2i}\bigg\{\log c+\log k_2-(c+1)\log t_{2i}-(k_2+1)\log(1+t_{2i}^{-c})$$
$$-(k_1+k_3)\log(1+t_{1i}^{-c})\bigg\}$$
$$+(1-\delta_{1i})(1-\delta_{2i})\bigg\{(k_1+k_3)\log(1+t_{1i}^{-c})+k_2\log(1+t_{2i}^{-c})\bigg\},$$

and given the order U_i for $i=1,2,3$, the contribution to the pseudo-log-likelihood from I_2 is given as follows

$$\log\prod_{i\in I_2}L(c,k_1,k_2,k_3)=\sum_{i\in I_2}\delta_{1i}\delta_{2i}\Big\{\log\Big[\big[f(t_{1i},c,k_1)F(t_{2i},c,k_3)f(t_{2i},c,k_2)\big]\nu_1$$
$$+\big[f(t_{1i},c,k_1)f(t_{2i},c,k_3)F(t_{2i},c,k_2)\big]\nu_2\Big]\Big\}$$
$$+\delta_{1i}(1-\delta_{2i})\Big\{\log\big[f(t_{1i},c,k_1)F(x_{2i},c,k_3)F(t_{2i},c,k_2)\big]\Big\}$$
$$+(1-\delta_{1i})\delta_{2i}\Big\{\log\big[F(t_{2i},c,k_3)f(t_{2i},c,k_2)\big[F(t_{1i},c,k_1)-F(t_{2i},c,k_1)\big]\big]\nu_1$$
$$+\big[F(t_{2i},c,k_2)f(t_{2i},c,k_3)\big[F(t_{1i},c,k_1)-F(t_{2i},c,k_1)\big]\big]\nu_2\Big]\Big\}$$
$$+(1-\delta_{1i})(1-\delta_{2i})\Big\{\log\big[F(t_{1i},c,k_1)F(t_{2i},c,k_3)F(t_{2i},c,k_2)\big]\Big\},$$

where ν_1 and ν_2 are defined as (7) ;

$$\sum_{i\in I_2}\delta_{1i}\delta_{2i}\Big\{\Big[2\log c+\log k_1+\log k_3-(k_1+1)\log(1+t_{1i}^{-c})$$
$$-(k_2+k_3+1)\log(1+t_{2i}^{-c})-(c+1)\log(t_{1i}t_{2i})\Big]\nu_1+\Big[2\log c+\log k_1+\log k_2$$
$$-(k_1+1)\log(1+t_{1i}^{-c})-(k_2+k_3+1)\log(1+t_{2i}^{-c})-(c+1)\log(t_{1i}t_{2i})\Big]\nu_2\Big\}$$
$$+(1-\delta_{1i})\delta_{2i}\Big\{\Big[\log c+\log k_3-(c+1)\log t_{2i}-(k_2+k_3+1)\log(1+t_{2i}^{-c})$$
$$+\log[(1+t_{1i}^{-c})^{-k_1}-(1+t_{2i}^{-c})^{-k_1}]\Big]\nu_1+\Big[\log c+\log k_2-(c+1)\log t_{2i}$$
$$-(k_2+k_3+1)\log(1+t_{2i}^{-c})+\log[(1+t_{1i}^{-c})^{-k_1}-(1+t_{2i}^{-c})^{-k_1}]\Big]\nu_2\Big\}$$
$$+\delta_{1i}(1-\delta_{2i})\Big\{\log c+\log k_1-(c+1)\log t_{1i}-(k_1+1)\log(1+t_{1i}^{-c})$$
$$-(k_2+k_3)\log(1+t_{2i}^{-c})\Big\}$$
$$+(1-\delta_{1i})(1-\delta_{2i})\Big\{k_1\log(1+t_{1i}^{-c})+(k_2+k_3)\log(1+t_{2i}^{-c})\Big\}.$$

Let

$$\begin{aligned}
N_0 &= n_{11}^0+n_{10}^0+n_{01}^0+2n_{11}^1+n_{10}^1+n_{01}^1+2n_{11}^2+n_{10}^2+n_{01}^2\\
N_1 &= n_{10}^0+\mu_2(n_{11}^1+n_{10}^1)+(n_{11}^2+n_{10}^2)\\
N_2 &= n_{01}^0+n_{11}^1+n_{01}^1+\nu_2(n_{11}^2+n_{01}^2)\\
N_3 &= n_{11}^0+\mu_1(n_{11}^1+n_{10}^1)+\nu_1(n_{11}^2+n_{01}^2), \qquad (8)
\end{aligned}$$

hence, the pseudo-log-likelihood is given by

$$N_0\log c+N_1\log k_1+N_2\log k_2+N_3\log k_3$$
$$-k_1\Big[\sum_{i\in I_0}\log(1+t_i^{-c})+\sum_{i\in I_1}\log(1+t_{1i}^{-c})+\sum_{i\in I_2}(1-\delta_{2i}+\delta_{1i}\delta_{2i})\log(1+t_{1i}^{-c})\Big]$$
$$-k_2\Big[\sum_{i\in I_0}\log(1+t_i^{-c})+\sum_{i\in I_1}(1-\delta_{1i}+\delta_{1i}\delta_{2i})\log(1+t_{2i}^{-c})+\sum_{i\in I_2}\log(1+t_{2i}^{-c})\Big]$$
$$-k_3\Big[\sum_{i\in I_0}\log(1+t_i^{-c})+\sum_{i\in I_1}\log(1+t_{1i}^{-c})+\sum_{i\in I_2}\log(1+t_{2i}^{-c})\Big]$$

$$
\begin{aligned}
&-(c+1)\big[\sum_{i\in I_0}(\delta_{1i}+\delta_{2i}-\delta_{1i}\delta_{2i})\log t_i+\sum_{i\in I_1\cup I_2}(\delta_{1i}\log t_{1i}+\delta_{2i}\log t_{2i})\big]\\
&-\sum_{i\in I_0}(\delta_{1i}+\delta_{2i}-\delta_{1i}\delta_{2i})\log(1+t_i^{-c})-\sum_{i\in I_1\cup I_2}\left[\delta_{1i}\log(1+t_{1i}^{-c})+\delta_{2i}\log(1+t_{2i}^{-c})\right]\\
&+\sum_{i\in I_1}\delta_{1i}(1-\delta_{2i})\log\left[(1+t_{2i}^{-c})^{-k_2}-(1+t_{1i}^{-c})^{-k_2}\right]\\
&+\sum_{i\in I_2}(1-\delta_{1i})\delta_{2i}\log\left[(1+t_{1i}^{-c})^{-k_1}-(1+t_{2i}^{-c})^{-k_1}\right].
\end{aligned}
$$

In order to implement the M-step of the EM algorithm, we need to maximize the pseudo-log-likelihood equation w.r.t. c, k_1, k_2, and k_3. We denote the first derivatives of the pseudo-log-likelihood function as

$$
\begin{aligned}
\frac{\partial\ell}{\partial k_1}=\frac{N_1}{k_1}&-\big[\sum_{i\in I_0}\log(1+t_i^{-c})+\sum_{i\in I_1}\log(1+t_{1i}^{-c})\\
&-\sum_{i\in I_2}(1-\delta_{2i}+\delta_{1i}\delta_{2i})\log(1+t_{1i}^{-c})\big]\\
&-\sum_{i\in I_2}(1-\delta_{1i})\delta_{2i}\big\{\frac{\ln(1+t_{1i}^{-c})(1+t_{1i}^{-c})^{-k_1}-\ln(1+t_{2i}^{-c})(1+t_{2i}^{-c})^{-k_1}}{(1+t_{1i}^{-c})^{-k_1}-(1+t_{2i}^{-c})^{-k_1}}\big\},\\
\frac{\partial\ell}{\partial k_2}=\frac{N_2}{k_2}&-\big[\sum_{i\in I_0}\log(1+t_i^{-c})+\sum_{i\in I_2}\log(1+t_{2i}^{-c})\\
&-\sum_{i\in I_1}(1-\delta_{1i}+\delta_{1i}\delta_{2i})\log(1+t_{2i}^{-c})\big]\\
&-\sum_{i\in I_1}\delta_{1i}(1-\delta_{2i})\big\{\frac{\ln(1+t_{2i}^{-c})(1+t_{2i}^{-c})^{-k_2}-\ln(1+t_{1i}^{-c})(1+t_{1i}^{-c})^{-k_2}}{(1+t_{2i}^{-c})^{-k_2}-(1+t_{1i}^{-c})^{-k_2}}\big\},\\
\frac{\partial\ell}{\partial k_3}=\frac{N_3}{k_3}&-\big[\sum_{i\in I_0}\log(1+t_i^{-c})+\sum_{i\in I_1}\log(1+t_{1i}^{-c})+\sum_{i\in I_2}\log(1+t_{2i}^{-c})\big]
\end{aligned}
$$

$$\begin{aligned}
\frac{\partial \ell}{\partial c} &= \frac{N_0}{c} + (k_1+k_2+k_3)\sum_{i\in I_0}\frac{\ln t_i t_i^{-c}}{1+t_i^{-c}} + (k_1+k_3)\sum_{i\in I_1}\frac{\ln t_{1i} t_{1i}^{-c}}{1+t_{1i}^{-c}} \\
&+(k_2+k_3)\sum_{i\in I_2}\frac{\ln t_{2i} t_{2i}^{-c}}{1+t_{2i}^{-c}} + k_2\sum_{i\in I_1}(1-\delta_{1i}+\delta_{1i}\delta_{2i})\frac{\ln t_{2i} t_{2i}^{-c}}{1+t_{2i}^{-c}} \\
&+k_1\sum_{i\in I_2}(1-\delta_{2i}+\delta_{1i}\delta_{2i})\frac{\ln t_{1i} t_{1i}^{-c}}{1+t_{1i}^{-c}} - \sum_{i\in I_0}(\delta_{1i}+\delta_{2i}-\delta_{1i}\delta_{2i})\log t_i \\
&- \sum_{i\in I_1\cup I_2}[\delta_{1i}\log t_{1i} + \delta_{2i}\log t_{2i}] + \sum_{i\in I_0}(\delta_{1i}+\delta_{2i}-\delta_{1i}\delta_{2i})\frac{\ln t_i t_i^{-c}}{1+t_i^{-c}} \\
&+ \sum_{i\in I_1\cup I_2}[\frac{\delta_{1i}\ln t_{1i} t_{1i}^{-c}}{1+t_{1i}^{-c}} + \frac{\delta_{2i}\ln t_{2i} t_{2i}^{-c}}{1+t_{2i}^{-c}}] \\
&+\sum_{i\in I_1}\delta_{1i}(1-\delta_{2i})\Big\{\frac{k_2\ln t_{2i}t_{2i}^{-c}(1+t_{2i}^{-c})^{-k_2-1} - k_2\ln t_{1i}t_{1i}^{-c}(1+t_{1i}^{-c})^{-k_2-1}}{(1+t_{2i}^{-c})^{-k_2}-(1+t_{1i}^{-c})^{-k_2}}\Big\} \\
&+\sum_{i\in I_2}(1-\delta_{1i})\delta_{2i}\Big\{\frac{k_1\ln t_{1i}t_{1i}^{-c}(1+t_{1i}^{-c})^{-k_1-1} - k_1\ln t_{2i}t_{2i}^{-c}(1+t_{2i}^{-c})^{-k_1-1}}{(1+t_{1i}^{-c})^{-k_1}-(1+t_{2i}^{-c})^{-k_1}}\Big\}.
\end{aligned}$$

where N_0, N_1, N_2 and N_3 are defined in equation (2).

4 Observed Fisher Information Matrix

In this section, we present the observed Fisher information matrix obtained using the idea of Louis (1982), which is used when the EM algorithm is applied to obtain the MLEs in the case of incomplete data problem. The observed information matrix can then be inverted to obtain the asymptotic covariance matrix of the MLEs determined from the EM algorithm. Let S denote the derivative vector and H the Hessian matrix of the pseudo-log-likelihood function defined in (1). The observed Fisher information matrix is given by $H - SS^T$.

Since the Marshal-Olkin bivariate Burr Type III is a regular family, the usual asymptotic normality result holds in this case, i.e.

$$\sqrt{n}(\underline{\hat{\theta}} - \underline{\theta}) \to N_4(0, I^{-1})$$

here I is the expected Fisher information matrix and $\underline{\theta} = (c, k_1, k_2, k_3)$. The $(1-p)$ 100 % approximate confidence intervals for the unknown parameters c and $k_i, i = 1, 2, 3$ is given by $\hat{c} \pm z_{\frac{p}{2}}\sqrt{var(\hat{c})}$ and $\hat{k_i} \pm z_{\frac{p}{2}}\sqrt{var(\hat{k_i})}, i = 1, 2, 3$ respectively, where $z_{\frac{p}{2}}$ is the $\frac{p}{2}$-th upper percentile of the standard Normal distribution.

5 Numerical Experiments

In this section, we present simulation studies of maximum likelihood estimators using the direct method and the EM algorithm for different sample sizes and different parameter values for censored data. The results of this section obtained using R software. For conducting the experiment, we assume that the pair of orginal random variables (X_1, X_2) is distributed as MOBBIII(c, k_1, k_2, k_3) and we assume that the pair of censoring variables (Y_1, Y_2) is distributed as MOBBIII(c, k_1^*, k_2^*, k_3^*). Note that in this case

$$\begin{aligned} P(X_1 < Y_1) &= \frac{k_1^* + k_3^*}{k_1 + k_3 + k_1^* + k_3^*}, \\ P(X_2 < Y_2) &= \frac{k_2^* + k_3^*}{k_2 + k_3 + k_2^* + k_3^*}, \end{aligned}$$

since both pairs (X_1, X_2) and (Y_1, Y_2) have the same parameter c, this ensures that the percentage of censoring does not depend on c.
We used the Fisher's information matrix whose given in the appendix, to construct asymptotic confidence intervals for the direct method and the EM algorithm method.
In these simulation studies, four different cases are considered as follows,
case 1: $c = 0.6$, $k_1 = 1.0$, $k_2 = 1.2$, $k_3 = 1.3, 1.4, 1.5$ and $c^* = 1.5$, $k_1^* = 0.3$ and $k_2^* = 0.4$ and $K_3^* = 0.5$ (The results of this case in the tables 1 - 2 Provided).
case 2: $c = 0.6$, $k_1 = 1.5$, $k_2 = 2.5$, $k_3 = 2, 3.5, 5$ and $c^* = 1.\ 2$, $k_1^* = 0.\ 5$ and $k_2^* = 0.8$ and $k_3^* = 1$ (The results provided in tables 3 - 4).
case 3: $c = 0.6$, $k_1 = 1.5$, $k_2 = 1.5$, $k_3 = 1.5$ and $c^* = 0.6$, $k_1^* = k_2^* = k_3^* = 0.5$ And $k_1^* = k_2^* = k_3^* = 1.0$ (The results provided in table 5)
case 4: $c = 0.8$, $k_1 = 1.0$, $k_2 = 1.1$, $k_3 = 1.0, 1.2, 1.4$ and $c^* = 1.2$, $k_1^* = 0.2$ and $k_2^* = 0.3$ and $k_3^* = 0.3$ (The results provided in shapes 1 - 2).

Percentage of censorship with change k_3 changes. Mark the censoring percentage of the variables X_1 and X_2 with S_1 and S_2, respectively. In **case 1** the vector (S_1, S_2) values $(26, 26)$, $(25, 26)$ and $(24, 25)$ and in **case 2** the values $(30, 28)$, $(23, 23)$ and $(18, 19)$. In **case 3**, taking into account the values of the parameters $k_1 = k_2 = k_3 = 1.5$ and $k_2^* = k_3^* = 0.5, 1.0$, X_1 and X_2 equally 25 and 40 % are censored.
The average estimate, mean squared error, length of confidence intervals and coverage probability obtained from the maximum likelihood method represent by (ML), (MSE), (LCI) and (CP), respectively and the average estimate, mean squared error, length of confidence intervals and coverage probability obtained from EM algorithm method represent by ,$(EMML)$, $(EMMSE)$, $(EMLCI)$ and $(EMCP)$, respectively, for sample size $n = 20, 50, 100, 250$ for 1000 simulations.

Some of the salient features of the numerical experiments based on Tables 1 - 6 are given below.
i) The average estimates of the parameters c, k_1, k_2 and k_3 from the direct method for all k_3 choices are close to the correct values of the parameters.
ii) The mean square error of the estimators decreases with increasing in sample size.
iii) The average length of confidence interval length for large n for all parameters is significantly shorter than the average confidence interval length for smaller n.
iv) By increasing k_3 as the correlation parameter, the mean square error and the length of confidence interval of parameters k_1, k_2 and k_3 increases.
v) The mean square error, the length of confidence interval, and the coverage probability increase with a percentage of censorship from 25 to 40 % increased.
vi) It is observed that the mean square error, the length of confidence interval and the coverage probability in the second method (EM algorithm) for the parameters c, k_1, k_2 and k_3, is more satisfactory than the estimates in the first method.
Also, according to the results obtained from simulation results in the Figurs 1 - 2, it can be said that the average estimation of parameters c, k_1 and k_2 for all k_3 choices are close to the correct values of the parameters. The results also show; increasing the sample size reduces the mean square error. It is observed that the mean square error of the maximum likelihood estimates in the second method (EM algorithm) for the parameters c, k_1 and

Table 1. Average estimates, mean squared errors, average length of the confidence interval and coverage probability for n=20, 50 and case 1.

n	k_3	(S_1, S_2)	AVE	$c = 0.6$	$k_1 = 1$	$k_2 = 1.2$	k_3
20							
			MLE	0.4003	1.1059	1.2614	1.3482
			MSE	0.0464	0.2690	0.2346	0.2196
			LCI	0.26	1.55	1.65	1.73
			CP	20	89	89	93
	1.3	$(26, 26)$	$EMMLE$	0.4811	1.3184	1.6779	1.4794
			$EMMSE$	0.0288	0.1328	0.1113	0.2903
			$EMLCI$	0.28	1.87	2.12	1.75
			$EMCP$	45	94	86	95
			MLE	0.3929	1.1047	1.2609	1.4519
			MSE	0.0489	0.2716	0.2317	0.2336
			LCI	0.25	1.58	1.67	1.82
			CP	18	89	90	92
	1.4	$(25, 26)$	$EMMLE$	0.4820	1.3660	1.7297	1.5983
			$EMMSE$	0.0299	0.1013	0.1319	0.1093
			$EMLCI$	0.28	1.96	2.21	1.87
			$EMCP$	58	98	91	93
			MLE	0.3813	1.0700	1.3216	1.5607
			MSE	0.0525	0.2319	0.3224	0.2770
			LCI	0.25	1.61	1.69	1.89
			CP	12	88	88	92
	1.5	$(24, 25)$	$EMMLE$	0.7404	1.0414	1.1259	1.3289
			$EMMSE$	0.0887	0.1036	0.1334	0.1271
			$EMLCI$	0.34	1.52	1.47	1.52
			$EMCP$	58	99	94	92
50							
			MLE	0.3895	1.0303	1.2763	1.3218
			MSE	0.0472	0.0775	0.1121	0.0870
			LCI	$(0.31, 0.47)$	$(0.56, 1.49)$	$(0.75, 1.80)$	$(0.78, 1.86)$
			CP	22	91	91	93
	1.3	$(26, 26)$	$EMMLE$	0.6711	0.9792	1.1176	1.2060
			$EMMSE$	0.0198	0.0377	0.0587	0.0466
			$EMLCI$	0.20	0.90	0.89	0.91
			$EMCP$	60	98	94	95
			MLE	0.3836	1.0296	1.2781	1.4242
			MSE	0.0486	0.0772	0.1167	0.0933
			LCI	0.16	0.95	1.07	1.13
			CP	19	90	90	93
	1.4	$(25, 26)$	$EMMLE$	0.6716	0.9875	1.1243	1.2771
			$EMMSE$	0.0201	0.0384	0.0597	0.0562
			$EMLCI$	0.20	0.92	0.90	0.95
			$EMCP$	59	98	92	94
			MLE	0.3780	1.0303	1.2796	1.5262
			MSE	0.0510	0.0817	0.1183	0.1008
			LCI	0.15	0.97	1.07	1.18
			CP	14	92	91	93
	1.5	$(24, 25)$	$EMMLE$	0.6729	0.9970	1.1302	1.3461
			$EMMSE$	0.0206	0.0385	0.0600	0.0686
			$EMLCI$	0.20	0.94	0.91	0.98
			$EMCP$	59	99	95	93

Table 2. Average estimates, mean squared errors, average length of the confidence interval and coverage probability for n=100, 250 and case 1.

n	k_3	(S_1, S_2)	AVE	$c = 0.6$	$k_1 = 1$	$k_2 = 1.2$	k_3
100							
			MLE	0.3858	1.0370	1.2705	1.2943
			MSE	0.0467	0.0422	0.0525	0.0420
			LCI	0.11	0.65	0.74	0.75
			CP	22	92	92	94
	1.3	(26, 26)	$EMMLE$	0.6592	0.9710	1.1217	1.2002
			$EMMSE$	0.0098	0.0202	0.0315	0.0293
			$EMLCI$	0.13	0.63	0.62	0.64
			$EMCP$	64	97	94	95
			MLE	0.3798	1.0351	1.2680	1.3950
			MSE	0.0492	0.0425	0.0536	0.0457
			LCI	0.11	0.67	0.75	0.79
			CP	20	92	91	93
	1.4	(25, 26)	$EMMLE$	0.6598	0.9785	1.1274	1.2697
			$EMMSE$	0.0100	0.0201	0.0315	0.0383
			$EMLCI$	0.14	0.65	0.63	0.67
			$EMCP$	64	98	93	94
			MLE	0.3744	1.0361	1.2688	1.4961
			MSE	0.0416	0.0431	0.0555	0.0498
			LCI	0.11	0.69	0.75	0.83
			CP	17	93	93	93
	1.5	(24, 25)	$EMMLE$	0.6609	0.9874	1.1334	1.3401
			$EMMSE$	0.0101	0.0204	0.0312	0.0490
			$EMLCI$	0.14	0.67	0.62	0.69
			$EMCP$	63	97	96	94
250							
			MLE	0.3855	1.0361	1.2548	1.2972
			MSE	0.0463	0.0169	0.0219	0.0160
			LCI	0.7	0.42	0.46	0.47
			CP	23	94	93	94
	1.3	(26, 26)	$EMMLE$	0.6550	0.9669	1.1114	1.1901
			$EMMSE$	0.0055	0.0088	0.0181	0.0196
			$EMLCI$	0.8	0.40	0.38	0.40
			$EMCP$	65	98	95	96
			MLE	0.3796	1.0358	1.2548	1.3978
			MSE	0.0488	0.0170	0.0224	0.0178
			LCI	0.7	0.43	0.47	0.50
			CP	22	94	92	94
	1.4	(25, 26)	$EMMLE$	0.6562	0.9759	1.1184	1.2598
			$EMMSE$	0.0056	0.0084	0.0171	0.0277
			$EMLCI$	0.11	0.40	0.29	0.42
			$EMCP$	69	98	94	94
			MLE	0.3739	1.0374	1.2553	1.4972
			MSE	0.0413	0.0175	0.0230	0.0200
			LCI	0.7	0.43	0.4	0.52
			CP	20	95	94	94
	1.5	(24, 25)	$EMMLE$	0.6571	0.9843	1.1243	1.3299
			$EMMSE$	0.0057	0.0082	0.0165	0.0378
			$EMLCI$	0.10	0.42	0.39	0.43
			$EMCP$	63	98	95	97

Table 3. Average estimates, mean squared errors, average length of the confidence interval and coverage probability for n=20, 50 and case 2.

n	k_3	(S_1, S_2)	AVE	$c = 0.6$	$k_1 = 1.5$	$k_2 = 2.5$	k_3
20							
			MLE	0.2953	1.5963	2.7442	2.0979
			MSE	0.0943	0.5291	1.0261	0.5836
			LCI	0.23	2.50	4.28	3.33
			CP	8	88	92	92
	2	$(30, 28)$	$EMMLE$	0.6880	1.3834	2.1057	1.7808
			$EMMSE$	0.0538	0.1782	0.5511	0.2056
			$EMLCI$	0.35	1.98	2.65	2.09
			$EMCP$	65	97	82	93
			MLE	0.2664	1.5929	2.7519	3.6480
			MSE	0.1123	0.6090	1.1913	1.2146
			LCI	0.32	3.37	5.79	7.05
			CP	9	90	76	67
	3.5	$(23, 23)$	$EMMLE$	0.6542	1.4107	2.1265	2.7777
			$EMMSE$	0.0194	0.0850	0.4215	0.5046
			$EMLCI$	0.23	1.72	2.03	2.09
			$EMCP$	65	98	76	67
			MLE	0.2452	1.5749	2.7223	5.2145
			MSE	0.1266	0.7522	1.3631	2.1795
			LCI	0.33	3.42	5.46	9.7
			CP	6	83	83	90
	5	$(18, 19)$	$EMMLE$	0.6585	1.4994	2.1686	3.7190
			$EMMSE$	0.0201	0.1122	0.4277	1.9743
			$EMLCI$	0.22	2.02	2.19	3.77
			$EMCP$	64	99	86	49
50							
			MLE	0.2945	1.5334	2.6492	2.0394
			MSE	0.0938	0.1828	0.3737	0.2197
			LCI	0.14	1.48	2.39	1.91
			CP	11	97	82	94
	2	$(30, 28)$	$EMMLE$	0.6306	1.3266	2.1558	1.8077
			$EMMSE$	0.0085	0.0812	0.2246	0.0893
			$EMLCI$	0.18	1.01	1.37	1.09
			$EMCP$	71	97	82	94
			MLE	0.2654	1.5327	2.6504	3.5748
			MSE	0.0923	0.2305	0.4803	0.4481
			LCI	0.18	2.10	3.44	4.0
			CP	10	93	80	68
	3.5	$(23, 23)$	$EMMLE$	0.6369	1.4360	2.2027	2.7732
			$EMMSE$	0.0089	0.0692	0.2221	0.6324
			$EMLCI$	0.14	1.30	1.50	1.56
			$EMCP$	68	98	83	69
			MLE	0.2440	1.52535	2.6381	5.1187
			MSE	0.1269	0.2850	0.5727	0.7624
			LCI	0.26	2.86	4.41	8.01
			CP	14	87	84	96
	5	$(18, 19)$	$EMMLE$	0.6379	1.4994	2.2106	3.3717
			$EMMSE$	0.0088	0.0757	0.2407	1.7767
			$EMLCI$	0.17	1.56	1.63	2.51
			$EMCP$	67	99	86	59

Table 4. Average estimates, mean squared errors, average length of the confidence interval and coverage probability for n=100, 250 and case 2.

n	k_3	(S_1, S_2)	AVE	$c = 0.6$	$k_1 = 1.5$	$k_2 = 2.5$	k_3
100							
			MLE	0.2926	1.5441	2.6367	1.9917
			MSE	0.0947	0.0936	0.1878	0.1077
			LCI	0.10	1.04	1.64	1.31
			CP	12	90	82	93
	2	$(30, 28)$	$EMMLE$	0.6256	1.3071	2.1344	1.7990
			$EMMSE$	0.0054	0.0720	0.2079	0.0784
			$EMLCI$	0.15	0.83	1.12	0.91
			$EMCP$	72	97	84	94
			MLE	0.2641	1.5425	2.6223	3.5122
			MSE	0.0829	0.1183	0.2175	0.2106
			LCI	0.12	1.42	2.16	2.54
			CP	11	94	82	69
	3.5	$(23, 23)$	$EMMLE$	0.6308	1.4162	2.1808	2.7536
			$EMMSE$	0.0057	0.0508	0.1887	0.6284
			$EMLCI$	0.15	1.08	1.22	1.28
			$EMCP$	71	98	85	70
			MLE	0.2430	1.5391	2.6064	5.0278
			MSE	0.1155	0.1400	0.2559	0.3584
			LCI	0.22	2.43	3.75	6.89
			CP	15	90	85	96
	5	$(18, 19)$	$EMMLE$	0.6306	1.4745	2.1938	3.6994
			$EMMSE$	0.0055	0.0539	0.1996	1.8294
			$EMLCI$	0.14	1.28	1.34	2.61
			$EMCP$	68	99	86	60
250							
			MLE	0.2927	1.5443	2.6083	1.9934
			MSE	0.0945	0.0378	0.0780	0.0401
			LCI	0.06	0.66	1.03	0.83
			CP	12	92	85	94
	2	$(30, 28)$	$EMMLE$	0.6236	1.3040	2.1195	1.7849
			$EMMSE$	0.0024	0.0518	0.1750	0.0608
			$EMLCI$	0.10	0.59	0.69	0.66
			$EMCP$	76	97	86	94
			MLE	0.2643	1.5359	2.5895	3.5075
			MSE	0.0796	0.0447	0.0865	0.0860
			LCI	0.08	0.87	1.29	1.53
			CP	11	96	83	70
	3.5	$(23, 23)$	$EMMLE$	0.6275	1.4276	2.1606	2.7328
			$EMMSE$	0.0025	0.0252	0.1502	0.6206
			$EMLCI$	0.10	0.69	0.86	0.90
			$EMCP$	73	98	86	72
			MLE	0.2430	1.5332	2.5766	5.0162
			MSE	0.1074	0.0511	0.1025	0.1453
			LCI	0.16	1.66	2.60	4.75
			CP	15	95	85	97
	5	$(18, 19)$	$EMMLE$	0.6280	1.4767	2.1703	3.6814
			$EMMSE$	0.0026	0.0223	0.1506	1.8175
			$EMLCI$	0.09	0.81	1.08	2.03
			$EMCP$	68	99	86	63

Table 5. Average estimates, mean squared errors, average length of the confidence interval and coverage probability for n=100, 250 and case 3.

n	(S_1, S_2)	AVE	$c = 0.6$	$k_1 = 1.5$	$k_2 = 1.5$	$k_3 = 1.5$
20	$(25, 25)$					
		MLE	0.3495	1.6110	1.6469	1.5668
		MSE	0.0658	0.3878	0.4452	0.3318
		LCI	$(0.22, 0.48)$	$(0.49, 2.83)$	$(0.51, 2.87)$	$(0.44, 2.66)$
		CP	8	78	77	89
		$EMMLE$	0.7108	1.2058	1.3779	1.4664
		$EMMSE$	0.0591	0.2244	0.1928	0.1184
		$EMLCI$	0.34	1.52	1.79	1.75
		$EMCP$	61	82	81	90
	$(40, 40)$					
		MLE	0.3217	1.6196	1.6582	1.5610
		MSE	0.0896	0.4205	0.4935	0.3615
		LCI	0.31	3.24	2.88	2.57
		CP	6	82	71	87
		$EMMLE$	0.7265	1.3853	1.2346	1.3636
		$EMMSE$	0.0784	0.1781	0.2681	0.1700
		$EMLCI$	0.36	1.86	1.58	1.72
		$EMCP$	56	84	74	90
50	$(25, 25)$					
		MLE	0.3476	1.5669	1.5896	1.5199
		MSE	0.0657	0.1520	0.1668	0.1223
		LCI	0.15	1.43	1.34	1.33
		CP	10	80	85	90
		$EMMLE$	0.6653	1.1619	1.3760	1.4942
		$EMMSE$	0.0177	0.1744	0.0861	0.0526
		$EMLCI$	0.21	0.89	1.10	1.13
		$EMCP$	62	82	89	91
	$(40, 40)$					
		MLE	0.3199	1.5672	1.5947	1.5203
		MSE	0.0892	0.1585	0.1756	0.1344
		LCI	0.18	1.90	1.64	1.57
		CP	8	85	77	91
		$EMMLE$	0.6700	1.2993	1.2973	1.3702
		$EMMSE$	0.0199	0.1081	0.1134	0.0712
		$EMLCI$	0.23	1.1	1.05	1.09
		$EMCP$	65	87	81	91

Table 6. Average estimates, mean squared errors, average length of the confidence interval and coverage probability for n=100, 250 and case 3.

n	(S_1, S_2)	AVE	$c = 0.6$	$k_1 = 1.5$	$k_2 = 1.5$	$k_3 = 1.5$
100	$(25, 25)$					
		MLE	0.3447	1.5629	1.5824	1.4890
		MSE	0.0656	0.0772	0.0766	0.0616
		LCI	0.10	1.01	0.93	0.93
		CP	12	82	87	92
		$EMMLE$	0.6548	1.1548	1.3796	1.4892
		$EMMSE$	0.0086	0.1518	0.0485	0.0273
		$EMLCI$	0.14	0.60	0.77	0.79
		$EMCP$	63	85	90	95
	$(40, 40)$					
		MLE	0.3173	1.5675	1.5897	1.4861
		MSE	0.0802	0.0819	0.0839	0.0658
		LCI	0.12	1.35	1.13	1.08
		CP	10	87	80	92
		$EMMLE$	0.6605	1.2946	1.3004	1.3681
		$EMMSE$	0.0099	0.0776	0.0744	0.0432
		$EMLCI$	0.16	0.77	0.72	0.77
		$EMCP$	66	89	82	93
250	$(25, 25)$					
		MLE	0.3445	1.5658	1.5577	1.4933
		MSE	0.0654	0.0341	0.0326	0.0232
		LCI	0.07	0.64	0.58	0.59
		CP	13	83	90	93
		$EMMLE$	0.6543	1.3423	1.3455	1.3863
		$EMMSE$	0.0052	0.0376	0.0373	0.0222
		$EMLCI$	0.11	0.48	0.46	0.46
		$EMCP$	68	86	92	97
	$(40, 40)$					
		MLE	0.3174	1.5727	1.5678	1.4913
		MSE	0.0799	0.0370	0.0360	0.0249
		LCI	0.08	0.83	0.68	0.67
		CP	12	88	82	92
		$EMMLE$	0.6584	1.3171	1.3183	1.3563
		$EMMSE$	0.0059	0.0478	0.0474	0.0310
		$EMLCI$	0.13	0.49	0.47	$0.48S$
		$EMCP$	67	90	84	93

k_2, it is more satisfactory than the maximum likelihood estimates in the first method.
It can be seen by looking at all the cases; Increasing the sample size reduces the mean square error and length of confidence interval, because large amounts of n lead to better inference. Also, in all cases, the mean square error using the EM algorithm is less than the direct method, which states that the EM algorithm is more accurate. It is also observed; the confidence intervals created using the Louis (1982) method include the actual value of the parameters. Changing the values of k_3 as a correlation parameter changes the results. As k_3 increases, the mean square error and length of the confidence interval of the parameters k_1, k_2 and k_3 increase, which is expressed by increasing k_3 the correlation of the variables X_1 and X_2 increases, so the accuracy of the results decreases. The results state; the mean square error, the length of confidence interval, and the coverage probability increase with a percentage of censorship from 25 to 40 % increased; because the lower percentage of censorship lead to more observations and better inference.

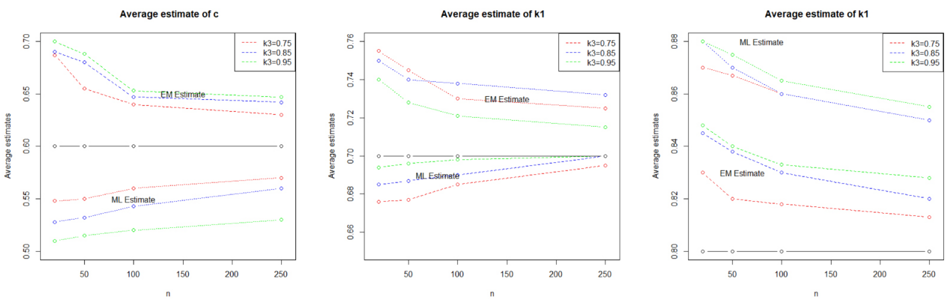

Figure 1. Average estimates c, k_1 and k_2

6 Data Analysis

The data set that analyzed in this section, are from the American Football (National Football League) matches played on three consecutive weekends in 1986. It has been originally published in 'Washington Post' and it is also available in Csorgo and Welsh (1989). Kundu and Gupta (2010), Jamalizadeh and Kundu (2013), and Balakrishna and Shiji (2014) analyzed this data. In this bivariate data set, X_1 represents the 'game time' to the first

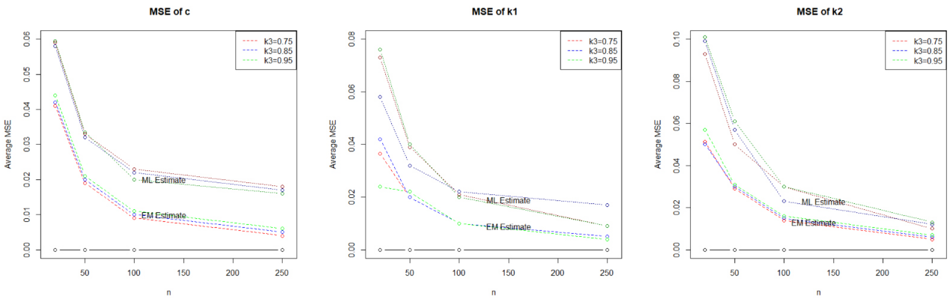

Figure 2. Mean squared errors c, k_1 and k_2

points scored by kicking the ball between goal posts, and X_2 represents the 'game time' to the first points scored by moving the ball into the end zone. The variables X_1 and X_2 have the following structure: (i)$X_1 < X_2$ means that the first score is a field goal, (ii) $X_1 = X_2$ means the first score is a converted touchdown, (iii) $X_1 > X_2$ means the first score is an unconverted touchdown or safety. In this case the ties are exact because no 'game time' elapses between a touchdown and a point-after conversion attempt. It should be noted that the possible scoring times are restricted by the duration of the game but it has been ignored similarly as in Csorgo and Welsh (1989).
Before analyzing the data using proposed EM algorithm and censoring scheme, we fit the Burr III distribution to X_1, X_2 and $\max(X_1, X_2)$, separately. The MLEs of the parameters of the Burr III distribution for X_1, X_2 and $\max(X_1, X_2)$ are (1.090, 5.152), (0.959, 5.244) and (0.952, 5.239), respectively. The Kolmogorov - Smirnov distances between the fitted distribution and the empirical distribution function and the corresponding p values (in brackets) for X_1, X_2 and $\max(X_1, X_2)$ are 0.186 (0.11), 0.196 (0.09) and 0.192 (0.10), respectively. Based on the p values Burr III distribution can be used for analyzing X_1, X_2 and $\max(X_1, X_2)$.
Now we fit the $MOBBIII$ model under the assumptions that all the four parameters are unknown. To start the EM algorithm we need some initial guesses of the unknown parameters. For c, we suggest to take the average values of 1.090, 0.959 and 0.952, i.e. 0.996. Assuming the initial guess of c as 0.996, solving three equations in three unknowns for k's, we get the initial guess values of k_1, k_2 and k_3 as 1.079, 2.257 and 3.731, respectively.

The data (X_1, X_2) consists of 42 data points. We assume that the pair (X_1, X_2) has MOBBIII$(\widehat{c}, \widehat{k_1}, \widehat{k_2}, \widehat{k_3})$. The pair of censoring random vari-

Table 7. Point estimate, Standard Erorr and confidence interval for real data set.

(S_1, S_2)	AVE	c	k_1	k_2	k_3
$(15, 15)$	$MLPE$	0.5824	0.9423	2.1526	3.3574
	$MLSE$	0.0427	0.1152	0.3745	0.4852
	$MLCI$	$(0.38, 0.68)$	$(0.63, 1.12)$	$(1.35, 2.95)$	$(2.65, 4.05)$
	$EMPE$	0.8127	0.9785	1.8427	3.0485
	$EMSE$	0.0374	0.0896	0.2349	0.3964
	$EMLCI$	$(0.69, 0.91)$	$(0.45, 1.49)$	$(1.35, 2.2.35)$	$(2.34, 3.74)$
$(30, 30)$	$MLPE$	0.4702	0.6601	2.0202	3.5145
	$MLSE$	0.0601	0.1776	0.5196	0.6569
	$MLCI$	$(0.35, 0.59)$	$(0.31, 1.02)$	$(1.00, 3.04)$	$(2.22, 4.80)$
	$EMPE$	0.8270	0.7227	1.4048	2.9941
	$EMSE$	0.0764	0.3417	0.3342	0.4979
	$EMLCI$	$(0.67, 0.97)$	$(0.05, 1.39)$	$(0.75, 2.06)$	$(2.02, 3.97)$

ables (Y_1, Y_2) has MOBBIII $(c^*, k_1^*, k_2^*, k_3^*)$. In order to ensure that $P(X_1 < Y_1) = P(X_2 < Y_2) = 0.15$, we take $(k_1^*, k_2^*, k_3^*) = (0.34, 0.52, 0.50)$ and $(\widehat{c}, \widehat{k_1}, \widehat{k_2}, \widehat{k_3}) = (0.80, 1.61, 2.27, 3.18)$ which is the estimates of c, k_1, k_2 and k_3 for complete data. In a similar way, $(k_1^*, k_2^*, k_3^*) = (1.04, 1.33, 1.00)$ ensures that $P(X_1 < Y_1) = P(X_2 < Y_2) = 0.30$.

We have used the proposed ECM algorithm to estimate the unknown parameters and the initial estimates used for c, k_1, k_2 and k_3, in both the cases. The point estimates and the corresponding confidence intervals for 15 % and 30 % censoring are reported in Table 7.

Testing whether the two marginals have the same distributions or not, can be carried out as follows;

$$H_0 : k_3 = 0 \qquad\qquad H_1 : k_3 \neq 0$$

Under the assumption H_0, the likelihood function of observations is written as follows;

$$\begin{aligned}\ell_{H_0}(c,k_1,k_2) = &(2n_1+2n_2)\log c + n_2\log k_1 + n_1\log k_2\\ &-(c+1)\big[\sum_{i\in I_1\cup I_2}\log x_{1i} + \sum_{i\in I_1\cup I_2}\log x_{2i}\big]\\ &-(k_1+1)\sum_{i\in I_2\cup I_2}\log(1+x_{1i}^{-c}) - (k_2+1)\sum_{i\in I_1\cup I_2}\log(1+x_{2i}^{-c}),\end{aligned}$$

we need to maximize the pseudo-log-likelihood equation w.r.t. c, k_1, k_2, and k_3;

$$\begin{aligned}\frac{\partial\ell_{H_0}}{\partial c} =&\frac{(2n_1+2n_2)}{c} + \sum_{i\in I_1\cup I_2}\log x_{1i} + \sum_{i\in I_1\cup I_2}\log x_{2i}\\ &-(k_1+1)\sum_{i\in I_1\cup I_2}\frac{\ln x_{1i}x_{1i}^{-c}}{1+x_{1i}^{-c}} - (k_2+1)\sum_{i\in I_1\cup I_2}\frac{\ln x_{2i}x_{2i}^{-c}}{1+x_{2i}^{-c}},\\ \frac{\partial\ell_{H_0}}{\partial k_1} =&\frac{n_2}{k_1} - \sum_{i\in I_1\cup I_2}\log(1+x_{1i}^{-c}),\\ \frac{\partial\ell_{H_0}}{\partial k_2} =&\frac{n_1}{k_2} - \sum_{i\in I_1\cup I_2}\log(1+x_{2i}^{-c}).\end{aligned}$$

Using Theorem 3 of Self and Liang (1987), it follows that;

$$2\big(\ell(\hat c,\hat k_1,\hat k_2,\hat k_3) - \ell_{H_0}(\hat c,\hat k_1,\hat k_2,0)\big) \to \frac{1}{2} + \frac{1}{2}\chi_1^2.$$

The value of the test statistic is 18.85. and the $p = 0.00$, hence we reject the null hypothesis. This observation also shows that the proposed MOBBIII may be used for analyzing this bivariate data set.

7 Conclusions

In this paper we have considered the MLEs of the four parameters of Marshal-Olkin Bivariate Burr III distribution when both components of the bivariate variable are subject to random censoring. The maximum likelihood estimator of the unknown parameters is obtained using the direct method and Expectation Conditional Maximization algorithm. We have looked at the pseudo-likelihood with information on ordering of U_1, U_2 and U_3 missing and treated this problem as missing value problem. The existence and uniqueness of the MLEs is studied graphically. The simulation results indicate that ECM algorithm performs very well for different sample sizes and also for various levels of random censoring that we have studied. We have also constructed the asymptotic confidence intervals using the idea of Louis (1982) and observed that the asymptotic confidence intervals give accurate results and hence can be used for testing purposes. In data analysis section, for obtaining MLEs of unknown parameters we needed to obtain the MLEs of compelete data by using complex calculations, only the resuls are reported in this paper.

References

Abdel-Ghaly, A.A., Al-Dayian, G.R., and Al-Kashkari, F.H. (1997). The Use of Burr Type XII Distribution on Software Reliability Growth Modelling. *Microelectronics Reliability*, **37**, 305-313.

Al-Khedhairi, A., and El-Gohary, A. (2008) A New Class of Bivariate Gompertz Distributions. *Internatinal Journal of Mathematics Analysis*, **2**, 235-253.

Azizi, A., and Sayyareh, A. (2019). Estimating the Parameters of the Marshal-Olkin Bivariate Burr Type III Distribution by EM Algorithm. *Journal of The Iranian Statistical Society*, **18**, 133-155.

Azizi, A., and Sayyareh, A. (2019). Inference about the Bivariate New Extended Weibull Distribution Based on Complete and Censored Data. *Communications in Statistics - Simulation and Computation*, https://doi.org/10.1080/03610918.2019.1658779

Asimit, A., Furman, V.E., and Vernic, R. (2016). Statistical Inference for a New Class of Multivariate Pareto Distributions. *Communications in Statistics Simulation and Computation*, **45**, 456-471.

Bagger, J. (2005). Wage Growth and Turnover in Denmark. *University of Aarhus*, Denmark.

Balakrishnan, N. (1989). Approximate MLE of the Scale Parameter of the Rayleigh Distribution with Censoring. *IEEE Transactions on Reliability*, **38**, 355-357.

Balakrishna, N., and Shiji, K. (2014). On a Class of Bivariate Exponential Distributions. *Statistics and Probability Letters*, **85**, 153-160.

Balakrishnan, N., and Varadan, J. (1991). Approximate MLEs for the Location and Scale Parameters of the Extreme Value Distribution with Censoring. *IEEE Transactions on Reliability*, **40**, 146-151.

Burr, I.W. (1942). Cumulative Frequency Distributions. *Annals of Mathematical Statistics*, **13**, 215-232.

Chernobai, A.S., Fabozzi, F.J., and Rachev, S.T. (2007). Operational Risk: A Guide to Basel II Capital Requirements. *Models and Analysis, John Wiley & Sons*, New York, USA.

CsöorgöCsörgö, S. and Welsh, A. (1989). Testing for Exponential and Marshall-Olkin Distributions. *Journal of Statistical Planning and Inference*, **23**, 287-300.

Dagum, C.A. (1977). New Model of Personal Income Distribution: Specification and Estimation. *Applied Economics*, **30**, 413-437.

Dey, A.K., Paul, B., and Kundu, D. (2018). An EM Algorithm for Absolute Continuous Bivariate Pareto Distribution. arXiv:1608.02199v4 [stat.CO] 18 Mar 2018.

El-Sherpieny, E.A., Ibrahim, S.A., and Bedar, R.E. (2013). A New Bivariate Generalized Gompertz Distribution. *Asian Journal of Applied Sciences*, **1-4**, 2321-0893.

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*, **53**, 956-965.

Kundu, D., and Gupta, R.D. (2009). Bivariate Generalized Exponential Distribution. *Journal of Multivariate Analysis*, **100**, 581-593.

Kundu, D., and Gupta, R.D. (2010). Modified Sarhan-Balakrishnan Singular Bivariate Distribution. *Journal of Statistical Planning and Inference*, **140**, 526-538.

Gove, J.H., Ducey, M.J., Leak, W.B., and Zhang, L. (2008). Rotated Sigmoid Structures in Managed Uneven-aged Northern Hardwork Stands: a Look at the Burr Type III Distribution. *Foresty*, **81**, 161-176.

Jamalizadeh, A., and Kundu, D. (2013). Weighted Marshall-Olkin Bivariate Exponential Distribution. *Statistics*, **47**, 917-928.

Houggard, P., Harvald, B., and Holm, N.V. (1992). Measuring the Similarities Between the Lifetimes of Adult Danish Twins Born Between 1881–1930. *Journal of the American Statistical Association*, **87**, 17–24.

Lin, D.Y., Sun, W., and Ying, Z. (1999). Nonparametric Estimation of the Gap Time Distribution for Serial Events with Censored Data. *Biometrika*, **86**, 59–70.

Lindsay, S.R., Wood, G.R., and Woollons, R.C. (1996). Modeling the Diameter Distribution of Forest Stands Using the Burr Distribution. *Journal of Applied Statistics*, **23**, 609-619.

Lee, K.R., Kapadia, C.H., and Dwight, B.B. (1980). On Estimating the Scale Parameter of the Rayleigh Distribution from Doubly Censored Samples. *Statistical Hefte*, **21**, 14-21.

Louis, T.A. (1982). Finding the Observed Information Matrix when Using the EM Algorithm. *Journal of the Royal Statistical Society, Series B*, **44**, 226-233.

Mielke, P.W. (1973). Another Family of Distributions for Describing and Analyzing Precipitation Data. *Journal of Applied Meteorology and Climatology*, **12**, 275-280.

Mitra, S., and Kundu, D. (2008). Analysis of the Left Censored Data from the Generalized Exponential Distribution. *Journal of Statistical Computation and Simulation*, **78**, 669-679.

Mirhosseini, S.M., Amini, M., Kundu, D., and Dolati, A. (2015). On a New Absolutely Continuous Bivariate Generalized Exponential Distribution. *Statistical Methods and Applications*, **24**, 61-83.

Nadarajah, S., and Kotz, S. (2006). Q Exponential is a Burr Distribution. *Physics Letters A*, **359**, 451-456.

Nadarajah, S., and Kotz, S. (2007). On the Alternative to the Weibull Function. *Engineering Fracture Mechanics* , **74**, 577-579.

Nandi, S., and Dewan, I. (2010). An EM Algorithm for Estimating the Parameters of Bivariate Weibull Distribution under Random Censoring. *Computational Statistics and Data Analysis*, **54**, 1559-1569.

Sarhan, A., and Balakrishnan, N. (2007). A New Class of Bivariate Distribution and its Mixture. *Journal of the Multivariate Analysis*, **98**, 1508-1527.

Sherrick, B.J., Garcia, P., and Tirupattur, V. (1996). Recovering Probabilistic Information from Option Markets: Tests of Distributional Assumptions. *Journal of Futures Markets* , **16**, 545-560.

Tejeda, H.A., and Goodwin, B.K. (2008). Modelling Crop Price Through a Burr Distribution and Analysis of Correlation Between Crop Prices and Yields Using Copula Method. *Paper Presented at the Annual Meeting of the Agriculture and Applied Economics Association*, Orlanndo, FL, USA.

Wingo, D.R. (1983). Maximum Likelihood Methods for Fitting the Burr Type XII Distribution of Life Test Data. *Journal of Biomedical Science*, **25**, 77-84.

Wingo, D.R. (1993). Maximum Likelihood Methods for Fitting the Burr Type XII Distribution to Multiply (Progressively) Censored Life Test Data. *Metrika*, **40**, 203-210.

Apendix A: Fisher Information Matrix

The observed Fisher information matrix elements for the censored data for the $MOBBIII$ distribution is as follows;

$$
\begin{aligned}
I_{11} &= -\frac{\partial^2 \ell}{\partial c^2} = \frac{M_0}{c^2} + B_1,\\
I_{22} &= -\frac{\partial^2 \ell}{\partial k_1^2} = \frac{M_1}{k_1^2} + \frac{M_{13}}{(k_1+k_3)^2} - B_2,\\
I_{33} &= -\frac{\partial^2 \ell}{\partial k_2^2} = \frac{M_2}{k_2^2} + \frac{M_{23}}{(k_2+k_3)^2} - B_3,\\
I_{44} &= -\frac{\partial^2 \ell}{\partial k_3^2} = \frac{n_{11}^0}{k_3^2} + \frac{M_{13}}{(k_1+k_3)^2} + \frac{M_{23}}{(k_2+k_3)^2},\\
I_{12} &= -\frac{\partial^2 \ell}{\partial c\,\partial k_1} = -\sum_{i\in I_0}\frac{\ln t_i t_i^{-c}}{1+t_i^{-c}} - \sum_{i\in I_1}\frac{\ln t_{1i} t_{1i}^{-c}}{1+t_{1i}^{-c}} - \sum_{i\in I_2}(1-\delta_{1i}+\delta_{1i}\delta_{2i})\frac{\ln t_{1i} t_{1i}^{-c}}{1+t_{1i}^{-c}}\\
&\quad - \sum_{i\in I_2}(1-\delta_{1i})\delta_{2i}\Big\{\frac{\ln t_{1i}t_{1i}^{-c}(1+t_{1i}^{-c})^{-k_1-1} - k_1 \ln t_{1i}t_{1i}^{-c}\ln(1+t_{1i}^{-c})(1+t_{1i}^{-c})^{-k_1-1}}{(1+t_{1i}^{-c})^{-k_1}-(1+t_{2i}^{-c})^{-k_1}}\Big\}\\
&\quad + \sum_{i\in I_2}(1-\delta_{1i})\delta_{2i}\Big\{\frac{\ln t_{2i}t_{2i}^{-c}(1+t_{2i}^{-c})^{-k_1-1} - k_1 \ln t_{2i}t_{2i}^{-c}\ln(1+t_{2i}^{-c})(1+t_{2i}^{-c})^{-k_1-1}}{(1+t_{1i}^{-c})^{-k_1}-(1+t_{2i}^{-c})^{-k_1}}\Big\}\\
&\quad - \sum_{i\in I_2}(1-\delta_{1i})\delta_{2i}\Big\{\frac{k_1\ln t_{1i}t_{1i}^{-c}(1+t_{1i}^{-c})^{-k_1-1} - k_1 \ln t_{2i}t_{2i}^{-c}(1+t_{2i}^{-c})^{-k_1-1}}{[(1+t_{1i}^{-c})^{-k_1}-(1+t_{2i}^{-c})^{-k_1}]}\Big\}\\
&\quad \times\Big\{\frac{\ln(1+t_{1i}^{-c})(1+t_{1i}^{-c})^{-k_1} - \ln(1+t_{2i}^{-c})(1+t_{2i}^{-c})^{-k_1}}{[(1+t_{1i}^{-c})^{-k_1}-(1+t_{2i}^{-c})^{-k_1}]}\Big\}\\
I_{13} &= -\frac{\partial^2 \ell}{\partial c\,\partial k_2} = -\sum_{i\in I_0}\frac{\ln t_i t_i^{-c}}{1+t_i^{-c}} - \sum_{i\in I_2}\frac{\ln t_{2i} t_{2i}^{-c}}{1+t_{2i}^{-c}} - \sum_{i\in I_1}(1-\delta_{2i}+\delta_{1i}\delta_{2i})\frac{\ln t_{2i} t_{2i}^{-c}}{1+t_{2i}^{-c}}\\
&\quad - \sum_{i\in I_1}\delta_{1i}(1-\delta_{2i})\Big\{\frac{\ln t_{2i}t_{2i}^{-c}(1+t_{2i}^{-c})^{-k_2-1} - k_2 \ln t_{2i}t_{2i}^{-c}\ln(1+t_{2i}^{-c})(1+t_{2i}^{-c})^{-k_2-1}}{(1+t_{2i}^{-c})^{-k_2}-(1+t_{1i}^{-c})^{-k_2}}\Big\}\\
&\quad + \sum_{i\in I_1}\delta_{1i}(1-\delta_{2i})\Big\{\frac{\ln t_{1i}t_{1i}^{-c}(1+t_{1i}^{-c})^{-k_2-1} - k_2 \ln t_{1i}t_{1i}^{-c}\ln(1+t_{1i}^{-c})(1+t_{1i}^{-c})^{-k_2-1}}{(1+t_{2i}^{-c})^{-k_2}-(1+t_{1i}^{-c})^{-k_2}}\Big\}\\
&\quad - \sum_{i\in I_1}\delta_{1i}(1-\delta_{2i})\Big\{\frac{k_2\ln t_{2i}t_{2i}^{-c}(1+t_{2i}^{-c})^{-k_2-1} - k_2 \ln t_{1i}t_{1i}^{-c}(1+t_{1i}^{-c})^{-k_2-1}}{[(1+t_{2i}^{-c})^{-k_2}-(1+t_{1i}^{-c})^{-k_2}]}\Big\}\\
&\quad \times\Big\{\frac{\ln(1+t_{2i}^{-c})(1+t_{2i}^{-c})^{-k_2} - \ln(1+t_{1i}^{-c})(1+t_{1i}^{-c})^{-k_2}}{[(1+t_{2i}^{-c})^{-k_2}-(1+t_{1i}^{-c})^{-k_2}]}\Big\}\\
I_{14} &= -\frac{\partial^2 \ell}{\partial k_1\,\partial k_2} = -\sum_{i\in I_0}\frac{\ln t_i t_i^{-c}}{1+t_i^{-c}} - \sum_{i\in I_1}\frac{\ln t_{1i} t_{1i}^{-c}}{1+t_{1i}^{-c}} - \sum_{i\in I_2}\frac{\ln t_{2i} t_{2i}^{-c}}{1+t_{2i}^{-c}}\\
I_{23} &= -\frac{\partial^2 \ell}{\partial k_1\,\partial k_2} = 0, \qquad I_{23} = -\frac{\partial^2 \ell}{\partial k_2\,\partial k_3} = 0, \qquad I_{34} = -\frac{\partial^2 \ell}{\partial k_2\,\partial k_3} = 0.
\end{aligned}
$$

Where B_1 B_2 and B_3 are defined as follows

$$B_1=(k_1+k_2+k_3)\sum_{i\in I_0}\frac{\ln^2 t_i t_i^{-c}}{1+t_i^{-c}}+(k_1+k_2+k_3)\sum_{i\in I_0}\left\{\frac{\ln t_i t_i^{-c}}{1+t_i^{-c}}\right\}^2$$
$$+(k_1+k_3)\sum_{i\in I_1}\frac{\ln^2 t_{1i} t_{1i}^{-c}}{1+t_{1i}^{-c}}+(k_1+k_3)\sum_{i\in I_1}\left\{\frac{\ln t_{1i} t_{1i}^{-c}}{1+t_{1i}^{-c}}\right\}^2$$
$$+(k_2+k_3)\sum_{i\in I_2}\frac{\ln^2 t_{2i} t_{2i}^{-c}}{1+t_{2i}^{-c}}+(k_2+k_3)\sum_{i\in I_2}\left\{\frac{\ln t_{2i} t_{2i}^{-c}}{1+t_{2i}^{-c}}\right\}^2$$
$$+k_2\sum_{i\in I_1}(1-\delta_{1i}+\delta_{1i}\delta_{2i})\frac{\ln^2 t_{2i} t_{2i}^{-c}}{1+t_{2i}^{-c}}+k_2\sum_{i\in I_1}(1-\delta_{1i}+\delta_{1i}\delta_{2i})\left\{\frac{\ln t_{2i} t_{2i}^{-c}}{1+t_{2i}^{-c}}\right\}^2$$
$$+k_1\sum_{i\in I_2}(1-\delta_{2i}+\delta_{1i}\delta_{2i})\frac{\ln^2 t_{1i} t_{1i}^{-c}}{1+t_{1i}^{-c}}+k_1\sum_{i\in I_2}(1-\delta_{2i}+\delta_{1i}\delta_{2i})\left\{\frac{\ln t_{1i} t_{1i}^{-c}}{1+t_{1i}^{-c}}\right\}^2$$
$$+\sum_{i\in I_0}(\delta_{1i}+\delta_{2i}-\delta_{1i}\delta_{2i})\frac{\ln^2 t_i t_i^{-c}}{1+t_i^{-c}}+\sum_{i\in I_0}(\delta_{1i}+\delta_{2i}-\delta_{1i}\delta_{2i})\left\{\frac{\ln t_i t_i^{-c}}{1+t_i^{-c}}\right\}^2$$
$$+\sum_{i\in I_1\cup I_2}[\frac{\delta_{1i}\ln^2 t_{1i} t_{1i}^{-c}}{1+t_{1i}^{-c}}]+\sum_{i\in I_1\cup I_2}[\frac{\delta_{1i}\ln t_{1i} t_{1i}^{-c}}{1+t_{1i}^{-c}}]^2$$
$$+\sum_{i\in I_1\cup I_2}[\frac{\delta_{2i}\ln^2 t_{2i} t_{2i}^{-c}}{1+t_{2i}^{-c}}]+\sum_{i\in I_1\cup I_2}[\frac{\delta_{2i}\ln t_{2i} t_{2i}^{-c}}{1+t_{2i}^{-c}}]^2$$
$$+\sum_{i\in I_1}\delta_{1i}(1-\delta_{2i})\left\{\frac{k_2\ln^2 t_{2i}t_{2i}^{-c}(1+t_{2i}^{-c})^{-k_2-1}-k_2(k_2+1)(\ln t_{2i}t_{2i}^{-c})^2(1+t_{2i}^{-c})^{-k_2-2}}{(1+t_{2i}^{-c})^{-k_2}-(1+t_{1i}^{-c})^{-k_2}}\right\}$$
$$-\sum_{i\in I_1}\delta_{1i}(1-\delta_{2i})\left\{\frac{k_2\ln^2 t_{1i}t_{1i}^{-c}(1+t_{1i}^{-c})^{-k_2-1}-k_2(k_2+1)(\ln t_{1i}t_{1i}^{-c})^2(1+t_{1i}^{-c})^{-k_2-2}}{(1+t_{2i}^{-c})^{-k_2}-(1+t_{1i}^{-c})^{-k_2}}\right\}$$
$$+\sum_{i\in I_1}\delta_{1i}(1-\delta_{2i})\left\{\frac{k_2\ln t_{2i}t_{2i}^{-c}(1+t_{2i}^{-c})^{-k_2-1}-k_2\ln t_{1i}t_{1i}^{-c}(1+t_{1i}^{-c})^{-k_2-1}}{(1+t_{2i}^{-c})^{-k_2}-(1+t_{1i}^{-c})^{-k_2}}\right\}^2$$
$$+\sum_{i\in I_2}(1-\delta_{1i})\delta_{2i}\left\{\frac{k_1\ln^2 t_{1i}t_{1i}^{-c}(1+t_{1i}^{-c})^{-k_1-1}-k_1(k_1+1)(\ln t_{1i}t_{1i}^{-c})^2(1+t_{1i}^{-c})^{-k_1-2}}{(1+t_{1i}^{-c})^{-k_1}-(1+t_{2i}^{-c})^{-k_1}}\right\}$$
$$-\sum_{i\in I_2}(1-\delta_{1i})\delta_{2i}\left\{\frac{k_1\ln^2 t_{2i}t_{2i}^{-c}(1+t_{2i}^{-c})^{-k_1-1}-k_1(k_1+1)(\ln t_{2i}t_{2i}^{-c})^2(1+t_{2i}^{-c})^{-k_1-2}}{(1+t_{1i}^{-c})^{-k_1}-(1+t_{2i}^{-c})^{-k_1}}\right\}$$
$$+\sum_{i\in I_2}(1-\delta_{1i})\delta_{2i}\left\{\frac{k_1\ln t_{1i}t_{1i}^{-c}(1+t_{1i}^{-c})^{-k_1-1}-k_1\ln t_{2i}t_{2i}^{-c}(1+t_{2i}^{-c})^{-k_1-1}}{(1+t_{1i}^{-c})^{-k_1}-(1+t_{2i}^{-c})^{-k_1}}\right\}^2$$

$$B_2=\sum_{i\in I_2}(1-\delta_{1i})\delta_{2i}\left\{\frac{\ln^2(1+t_{1i}^{-c})(1+t_{1i}^{-c})^{-k_1}-\ln^2(1+t_{2i}^{-c})(1+t_{2i}^{-c})^{-k_1}}{(1+t_{1i}^{-c})^{-k_1}-(1+t_{2i}^{-c})^{-k_1}}\right\}$$
$$-\sum_{i\in I_2}(1-\delta_{1i})\delta_{2i}\left\{\frac{\ln(1+t_{1i}^{-c})(1+t_{1i}^{-c})^{-k_1}-\ln(1+t_{2i}^{-c})(1+t_{2i}^{-c})^{-k_1}}{(1+t_{1i}^{-c})^{-k_1}-(1+t_{2i}^{-c})^{-k_1}}\right\}^2$$

$$B_3 = \sum_{i \in I_1} \delta_{1i}(1-\delta_{2i})\Big\{\frac{\ln^2(1+t_{2i}^{-c})(1+t_{2i}^{-c})^{-k_2} - \ln^2(1+t_{1i}^{-c})(1+t_{1i}^{-c})^{-k_2}}{(1+t_{2i}^{-c})^{-k_2} - (1+t_{1i}^{-c})^{-k_2}}\Big\}$$
$$- \sum_{i \in I_1} \delta_{1i}(1-\delta_{2i})\Big\{\frac{\ln(1+t_{2i}^{-c})(1+t_{2i}^{-c})^{-k_2} - \ln(1+t_{1i}^{-c})(1+t_{1i}^{-c})^{-k_2}}{(1+t_{2i}^{-c})^{-k_2} - (1+t_{1i}^{-c})^{-k_2}}\Big\}^2 \tag{9}$$

In this section, the Fisher information matrix is observed using the Louis method (1982) obtained. This method is used when the EM algorithm is used to obtain the maximum likelihood estimates for censored data. The observed Fisher information matrix to obtain the variance-covariance matrix the estimates obtained from the EM algorithm are inverted. Suppose S is the first derivative vector and H is the Hessein matrix. The Fisher information matrix will be calculated as $H - SS^T$. Vector Components $S = (S_1, S_2, S_3, S_4)^T$ using N_0, N_1, N_2 and N_3 are as follows;

$$S_1 = \frac{N_0}{c} + h_1(c, k_1, k_2, k_3), \qquad S_2 = \frac{N_1}{k_1} - h_2(c, k_1)$$
$$S_3 = \frac{N_2}{k_2} - h_3(c, k_2), \qquad S_4 = \frac{N_3}{k_3} - h_4(c)$$

where

$$h_1(c, k_1, k_2, k_3) = (k_1+k_2+k_3)\sum_{i \in I_0} \frac{\ln t_i t_i^{-c}}{1+t_i^{-c}} + (k_1+k_3)\sum_{i \in I_1} \frac{\ln t_{1i} t_{1i}^{-c}}{1+t_{1i}^{-c}}$$
$$+(k_2+k_3)\sum_{i \in I_2} \frac{\ln t_{2i} t_{2i}^{-c}}{1+t_{2i}^{-c}} + k_2 \sum_{i \in I_1}(1-\delta_{1i}+\delta_{1i}\delta_{2i})\frac{\ln t_{2i} t_{2i}^{-c}}{1+t_{2i}^{-c}}$$
$$+k_1 \sum_{i \in I_2}(1-\delta_{2i}+\delta_{1i}\delta_{2i})\frac{\ln t_{1i} t_{1i}^{-c}}{1+t_{1i}^{-c}} - \sum_{i \in I_0}(\delta_{1i}+\delta_{2i}-\delta_{1i}\delta_{2i})\log t_i$$
$$- \sum_{i \in I_1 \cup I_2}[\delta_{1i}\log t_{1i} + \delta_{2i}\log t_{2i}] + \sum_{i \in I_0}(\delta_{1i}+\delta_{2i}-\delta_{1i}\delta_{2i})\frac{\ln t_i t_i^{-c}}{1+t_i^{-c}}$$
$$+ \sum_{i \in I_1 \cup I_2}\Big[\frac{\delta_{1i}\ln t_{1i} t_{1i}^{-c}}{1+t_{1i}^{-c}} + \frac{\delta_{2i}\ln t_{2i} t_{2i}^{-c}}{1+t_{2i}^{-c}}\Big]$$
$$+ \sum_{i \in I_1} \delta_{1i}(1-\delta_{2i})\Big\{\frac{k_2 \ln t_{2i} t_{2i}^{-c}(1+t_{2i}^{-c})^{-k_2-1} - k_2 \ln t_{1i} t_{1i}^{-c}(1+t_{1i}^{-c})^{-k_2-1}}{(1+t_{2i}^{-c})^{-k_2} - (1+t_{1i}^{-c})^{-k_2}}\Big\}$$
$$+ \sum_{i \in I_2} (1-\delta_{1i})\delta_{2i}\Big\{\frac{k_1 \ln t_{1i} t_{1i}^{-c}(1+t_{1i}^{-c})^{-k_1-1} - k_1 \ln t_{2i} t_{2i}^{-c}(1+t_{2i}^{-c})^{-k_1-1}}{(1+t_{1i}^{-c})^{-k_1} - (1+t_{2i}^{-c})^{-k_1}}\Big\}$$

$$h_2(c,k_1) = \big[\sum_{i\in I_0} \log(1+t_i^{-c}) + \sum_{i\in I_1} \log(1+t_{1i}^{-c})$$
$$+ \sum_{i\in I_2} (1-\delta_{2i}+\delta_{1i}\delta_{2i}) \log(1+t_{1i}^{-c})\big]$$
$$+ \sum_{i\in I_2} (1-\delta_{1i})\delta_{2i} \big\{ \frac{\ln(1+t_{1i}^{-c})(1+t_{1i}^{-c})^{-k_1} - \ln(1+t_{2i}^{-c})(1+t_{2i}^{-c})^{-k_1}}{(1+t_{1i}^{-c})^{-k_1} - (1+t_{2i}^{-c})^{-k_1}} \big\},$$

$$h_3(c,k_2) = \big[\sum_{i\in I_0} \log(1+t_i^{-c}) + \sum_{i\in I_2} \log(1+t_{2i}^{-c})$$
$$+ \sum_{i\in I_1} (1-\delta_{1i}+\delta_{1i}\delta_{2i}) \log(1+t_{2i}^{-c})\big]$$
$$+ \sum_{i\in I_1} \delta_{1i}(1-\delta_{2i}) \big\{ \frac{\ln(1+t_{2i}^{-c})(1+t_{2i}^{-c})^{-k_2} - \ln(1+t_{1i}^{-c})(1+t_{1i}^{-c})^{-k_2}}{(1+t_{2i}^{-c})^{-k_2} - (1+t_{1i}^{-c})^{-k_2}} \big\},$$
$$h_4(c) = \big[\sum_{i\in I_0} \log(1+t_i^{-c}) + \sum_{i\in I_1} \log(1+t_{1i}^{-c}) + \sum_{i\in I_2} \log(1+t_{2i}^{-c})\big].$$

Hessian Matrix

$$H_{11} = -\frac{N_0}{c^2} - B_1, \qquad H_{12} = -I_{12}, \qquad H_{13} = -I_{13}$$
$$H_{22} = -\frac{N_1}{k_1^2} + B_2, \qquad H_{14} = -I_{14}, \qquad H_{23} = 0,$$
$$H_{33} = -\frac{N_2}{k_2^2} + B_3, \qquad H_{24} = 0, \qquad H_{44} = -\frac{N_3}{k_3^2}$$

A. Azizi
Department of Statistics,
Razi University,
Kermanshah, Iran.
email: *aziziafsaneh425@yahoo.com*

A. Sayyareh
Department of Computer Science and Statistics,
K.N. Toosi University of Technology,
Tehran, Iran.
email: *asayyareh@kntu.ac.ir*