

isid/ms/2010/04

June 10, 2010

<http://www.isid.ac.in/~statmath/eprints>

An EM algorithm for estimation of parameters of
bivariate generalised exponential distribution under
random left censoring

ISHA DEWAN
SWAGATA NANDI

Indian Statistical Institute, Delhi Centre
7, SJSS Marg, New Delhi-110 016, India

AN EM ALGORITHM FOR ESTIMATION OF PARAMETERS OF BIVARIATE GENERALISED EXPONENTIAL DISTRIBUTION UNDER RANDOM LEFT CENSORING

ISHA DEWAN AND SWAGATA NANDI

ABSTRACT. We consider the four parameter bivariate generalised exponential distribution proposed by Kundu and Gupta (2009) and propose an EM algorithm to find the maximum likelihood estimators of the four parameters under random left censoring. A numerical experiment is carried out to discuss the properties of the estimators obtained iteratively.

1. INTRODUCTION

Gupta and Kundu (1999) introduced the Generalised Exponential (GE) distribution and studied its probabilistic properties. It is very useful in studying skewed lifetime data and hence provides an alternative to gamma or Weibull distributions. For a recent review, see Gupta and Kundu (2007). This distribution is naturally suitable for modelling left censored data.

Left censored data arises often in medical studies. For example, in the study of epidemics, such as AIDS, the time of onset of infection is typically unknown. What is known is the time at which the patient reports to the doctor. Hence the time from infection to the development of the disease is left censored. Jaqmin-Gadda and Thiebaut (2000) considered the classical mixed effects linear model for the analysis of progression of markers in HIV-infected patients where longitudinal studies of viral load are complicated by left censoring of the measures due to a lower quantification limit. Coburn, McBride and Ziller (2001) carried out a study on patterns of health insurance for rural and urban children and observed that a high proportion of rural children were left censored as they had entered the study without insurance. Bagger (2005) considered an econometric model where the data consisted of comprehensive matched employer employee panel data with detailed information on wages, tenure, experience and a range of other covariates. The data were left censored since some job durations are incomplete at the beginning of the job spells and are observed,

Date: July 25, 2010.

2000 Mathematics Subject Classification. 62N01, 62N02.

Key words and phrases. Bivariate Generalised Exponential Distribution; Random Left Censoring; EM Algorithm; Pseudo Likelihood.

Mitra and Kundu (2007) obtained the maximum likelihood estimators (M.L.E.'s) of the shape and scale parameters of GE distribution under random left censoring. In order to get the M.L.E. of the scale parameter one needs to solve a non linear equation iteratively. After getting the M.L.E. of the scale parameter one can obtain the M.L.E. of the shape parameter in a closed form.

Sarhan and Balakrishnan (2007) defined a new bivariate distribution using the GE distribution. But the marginals are not GE. Kundu and Gupta (2009) introduced a four parameter bivariate generalised exponential (BVGE) distribution with marginals that are GE. They suggested the use of EM algorithm to derive the M.L.E.s of the four parameters. They also provided the observed and the expected Fisher information matrices for the BVGE setup.

We study the M.L.E.s of the parameters of BVGE distribution under random left censoring. In section 2 we restate the joint distribution function and joint density of BVGE distribution and also write the likelihood function under random left censoring. We discuss the EM (conditional) algorithm in section 3 for finding the M.L.E.'s of the parameters iteratively. The findings of the numerical experiments are reported in section 4. One real data set is analysed in section 5 and we conclude in section 6. Details of the likelihood functions are given in Appendix A and the observed Fisher information matrix is in Appendix B.

2. BIVARIATE GENERALISED EXPONENTIAL DISTRIBUTION

Consider a generalised exponential distribution ($GE(\alpha, \lambda)$) with shape parameter $\lambda > 0$ and scale parameter $\alpha > 0$ with the density function and the distribution function, respectively, given by,

$$f_{GE}(x; \alpha, \lambda) = \lambda \alpha e^{-\alpha x} (1 - e^{-\alpha x})^{\lambda-1}, F_{GE}(x; \alpha, \lambda) = (1 - e^{-\alpha x})^\lambda, x > 0. \quad (1)$$

Note that the distribution function of GE random variable is λ th power of the distribution function of an exponential random variable with scale parameter α . If $\lambda \leq 1$, then the density function of GE random variable is strictly decreasing and for $\lambda > 1$, it is a unimodal skewed density function.

Suppose U_1, U_2, U_3 , respectively, are independent $GE(\alpha_1, \lambda)$, $GE(\alpha_2, \lambda)$, $GE(\alpha_3, \lambda)$, random variables where $\alpha_1, \alpha_2, \alpha_3, \lambda > 0$. Let $X_1 = \max(U_1, U_3)$ and $X_2 = \max(U_2, U_3)$. Then (X_1, X_2) has BVGE distribution with shape parameters $\alpha_1, \alpha_2, \alpha_3$ and scale parameter λ and is expressed as $BVGE(\alpha_1, \alpha_2, \alpha_3, \lambda)$. Let $\alpha_{13} = \alpha_1 + \alpha_3$, $\alpha_{23} = \alpha_2 + \alpha_3$ and $\alpha_{123} = \alpha_1 + \alpha_2 + \alpha_3$.

Then, joint distribution function of (X_1, X_2) is given as follows, where $z = \min(x, y)$,

$$\begin{aligned}
 F(x, y) &= F_{GE}(x; \alpha_1, \lambda) F_{GE}(y; \alpha_2, \lambda) F_{GE}(z; \alpha_3, \lambda) \\
 &= \begin{cases} F_{GE}(x; \alpha_{13}, \lambda) F_{GE}(y; \alpha_2, \lambda) & \text{if } x < y \\ F_{GE}(x; \alpha_1, \lambda) F_{GE}(y; \alpha_{23}, \lambda) & \text{if } x > y \\ F_{GE}(x; \alpha_{123}, \lambda) & \text{if } x = y. \end{cases} \quad (2)
 \end{aligned}$$

Then, joint density function of (X_1, X_2) is given as

$$\begin{aligned}
 f(x, y) &= \begin{cases} f_{GE}(x; \alpha_{13}, \lambda) f_{GE}(y; \alpha_2, \lambda) & \text{if } x < y \\ f_{GE}(x; \alpha_1, \lambda) f_{GE}(y; \alpha_{23}, \lambda) & \text{if } x > y \\ \frac{\alpha_3}{\alpha_{123}} f_{GE}(x; \alpha_{123}, \lambda) & \text{if } x = y. \end{cases} \quad (3)
 \end{aligned}$$

The marginals of X_1, X_2 have $GE(\alpha_{13}, \lambda)$ and $GE(\alpha_{23}, \lambda)$, respectively.

Suppose the pair (X_1, X_2) is subject to random left censoring by an independent pair of random variables (Y_1, Y_2) . We observe $(T_1, \delta_1; T_2, \delta_2)$ where

$$\begin{aligned}
 T_1 &= \max(X_1, Y_1) \quad \text{and} \quad \delta_1 = I(X_1 > Y_1), \\
 T_2 &= \max(X_2, Y_2) \quad \text{and} \quad \delta_2 = I(X_2 > Y_2). \quad (4)
 \end{aligned}$$

Therefore, if $X_1 < Y_1$, X_1 is censored. 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 $\int_0^{t_2} f(t_1, y) dy$. 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 $\int_0^{t_1} f(x, t_2) dx$. 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 $(T_{1i}, \delta_{1i}; T_{2i}, \delta_{2i})$, $i = 1, 2, \dots, n$ is given by

$$\begin{aligned}
 L &= L(\alpha_1, \alpha_2, \alpha_3, \lambda; t_{1i}, \delta_{1i}, t_{2i}, \delta_{2i}; i = 1, 2, \dots, n) \\
 &= \prod_{i=1}^n L(t_{1i}, \delta_{1i}; t_{2i}, \delta_{2i}) \\
 &= \prod_{i=1}^n [f(t_{1i} t_{2i})^{\delta_{1i} \delta_{2i}} [\int_0^{t_{2i}} f(t_{1i}, y) dy]^{\delta_{1i} (1 - \delta_{2i})} \\
 &\quad \times [\int_0^{t_{1i}} f(x, t_{2i}) dx]^{(1 - \delta_{1i}) \delta_{2i}} [F(t_{1i} t_{2i})]^{(1 - \delta_{1i})(1 - \delta_{2i})}]. \quad (5)
 \end{aligned}$$

The following quantities are required to express the likelihood explicitly

$$\int_0^y f(x, v)dv = \begin{cases} f_{GE}(x; \alpha_{13}, \lambda)[F_{GE}(y; \alpha_2, \lambda) - F_{GE}(x; \alpha_2, \lambda)] & \text{if } x < y \\ f_{GE}(x; \alpha_1, \lambda)F_{GE}(y; \alpha_{23}, \lambda) & \text{if } x > y \\ f_{GE}(t; \alpha_1, \lambda)F_{GE}(t; \alpha_{23}, \lambda) & \text{if } x = y = t, \end{cases} \quad (6)$$

and

$$\int_0^x f(u, y)du = \begin{cases} F_{GE}(x; \alpha_{13}, \lambda)f_{GE}(y; \alpha_2, \lambda) & \text{if } x < y \\ [F_{GE}(x; \alpha_1, \lambda) - F_{GE}(y; \alpha_1, \lambda)]f_{GE}(y; \alpha_{23}, \lambda) & \text{if } x > y \\ F_{GE}(t; \alpha_{13}, \lambda)f_{GE}(t; \alpha_2, \lambda) & \text{if } x = y = t. \end{cases} \quad (7)$$

Let I_0, I_1, I_2 , denote the following sets

$$I_0 = \{i|t_{1i} = t_{2i} = t_i\}, \quad I_1 = \{i|t_{1i} < t_{2i}\}, \quad I_2 = \{i|t_{1i} > t_{2i}\}. \quad (8)$$

Then the likelihood function can be written as

$$L = \prod_{i \in I_0} L(t_i, \delta_{1i}; t_i, \delta_{2i}) \prod_{i \in I_1} L(t_{1i}, \delta_{1i}; t_{2i}, \delta_{2i}) \prod_{i \in I_2} L(t_{1i}, \delta_{1i}; t_{2i}, \delta_{2i}). \quad (9)$$

Let n_0, n_1, n_2 , respectively, denote the number of elements in the sets I_0, I_1, 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 $(\delta_1, \delta_2) = (i, j)$, $i, j = 0, 1$, $k = 0, 1, 2$.

Using (2)-(7), the contributions to the likelihood on the sets I_0, I_1 and I_2 in terms of the unknown parameters $\alpha_1, \alpha_2, \alpha_3$ and λ are provided in Appendix A as equations (14)-(15). The likelihood function L based on the observed data is the product of these three equations.

3. EM ALGORITHM UNDER RANDOM LEFT CENSORING

As noted by Kundu and Gupta (2009), the M.L.E.'s of the four parameters can not be expressed in a closed form even in the complete sample (all pairs are observable) case when the shape parameter $\lambda = 1$. The same situation holds in the presence of left censored data. We propose the use of ECM (Expectation Conditional Maximisation) algorithm for finding the M.L.E.'s of the unknown parameters. We have earlier observed that X_1 and X_2 can be expressed as functions of independent random variables U_1, U_2 and U_3 . There is no identifiability issue on

I_0 . But on I_1 , we can identify α_{13}, α_2 and λ and on I_2 , we can identify α_{23}, α_1 and λ . In order to estimate α_1, α_2 and α_3 separately, we consider the ‘expectation’ step of the EM algorithm as follows.

Let γ denote the parameter vector $(\alpha_1, \alpha_2, \alpha_3, \lambda)^T$. It is easy to see that

$$P(U_1 < U_3 < U_2) = \frac{\alpha_2 \alpha_3}{\alpha_{123} \alpha_{13}}, \quad P(U_3 < U_1 < U_2) = \frac{\alpha_2 \alpha_1}{\alpha_{123} \alpha_{13}},$$

$$P(U_2 < U_3 < U_1) = \frac{\alpha_1 \alpha_3}{\alpha_{123} \alpha_{23}}, \quad P(U_3 < U_2 < U_1) = \frac{\alpha_1 \alpha_2}{\alpha_{123} \alpha_{23}}.$$

$$\begin{aligned} \mu_1(\gamma) &= P(U_1 < U_3 < U_2 | X_1 < X_2) = \frac{\alpha_3}{\alpha_{13}}, \\ \mu_2(\gamma) &= P(U_3 < U_1 < U_2 | X_1 < X_2) = \frac{\alpha_1}{\alpha_{13}}, \\ \nu_1(\gamma) &= P(U_2 < U_3 < U_1 | X_1 > X_2) = \frac{\alpha_3}{\alpha_{23}}, \\ \nu_2(\gamma) &= P(U_3 < U_2 < U_1 | X_1 > X_2) = \frac{\alpha_2}{\alpha_{23}}. \end{aligned}$$

Each of the sets I_0, I_1, I_2 contributes to the log-likelihood function of the ‘pseudo data’ based on ‘E’ step. In particular, the contribution to the pseudo log-likelihood from I_0 is given as follows;

$$\begin{aligned} &\sum_{i \in I_0} \left[\delta_{1i} \delta_{2i} \{ \log \lambda + \log \alpha_3 + (\alpha_{123} - 1) \log(1 - e^{-\lambda t_i}) - \lambda t_i \} \right. \\ &+ \delta_{1i} (1 - \delta_{2i}) \{ \log \lambda + \log \alpha_1 + (\alpha_{123} - 1) \log(1 - e^{-\lambda t_i}) - \lambda t_i \} \\ &+ (1 - \delta_{1i}) \delta_{2i} \{ \log \lambda + \log \alpha_2 + (\alpha_{123} - 1) \log(1 - e^{-\lambda t_i}) - \lambda t_i \} \\ &\left. + (1 - \delta_{1i}) (1 - \delta_{2i}) \{ \alpha_{123} \log(1 - e^{-\lambda t_i}) \} \right]. \end{aligned}$$

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

$$\begin{aligned}
& \sum_{i \in I_1} \left[\delta_{1i} \delta_{2i} \left\{ \mu_1 \left(2 \log \lambda + \log \alpha_2 + \log \alpha_3 + (\alpha_{13} - 1) \log(1 - e^{-\lambda t_{1i}}) \right. \right. \right. \\
& \quad \left. \left. \left. + (\alpha_2 - 1) \log(1 - e^{-\lambda t_{2i}}) - \lambda(t_{1i} + t_{2i}) \right) + \mu_2 \left(2 \log \lambda + \log \alpha_1 + \log \alpha_2 \right. \right. \right. \\
& \quad \left. \left. \left. + (\alpha_{13} - 1) \log(1 - e^{-\lambda t_{1i}}) + (\alpha_2 - 1) \log(1 - e^{-\lambda t_{2i}}) - \lambda(t_{1i} + t_{2i}) \right) \right\} \right. \\
& \quad \left. + \delta_{1i} (1 - \delta_{2i}) \left\{ \mu_1 \left(\log \lambda + \log \alpha_3 + (\alpha_{13} - 1) \log(1 - e^{-\lambda t_{1i}}) - \lambda t_{1i} \right. \right. \right. \\
& \quad \left. \left. \left. + \log[(1 - e^{-\lambda t_{2i}})^{\alpha_2} - (1 - e^{-\lambda t_{1i}})^{\alpha_2}] \right) + \mu_2 \left(\log \lambda + \log \alpha_1 \right. \right. \right. \\
& \quad \left. \left. \left. + (\alpha_{13} - 1) \log(1 - e^{-\lambda t_{1i}}) - \lambda t_{1i} + \log[(1 - e^{-\lambda t_{2i}})^{\alpha_2} - (1 - e^{-\lambda t_{1i}})^{\alpha_2}] \right) \right\} \right. \\
& \quad \left. + (1 - \delta_{1i}) \delta_{2i} \left\{ \log \lambda + \log \alpha_2 - \lambda t_{2i} + \alpha_{13} \log(1 - e^{-\lambda t_{1i}}) + (\alpha_2 - 1) \log(1 - e^{-\lambda t_{2i}}) \right\} \right. \\
& \quad \left. + (1 - \delta_{1i})(1 - \delta_{2i}) \left\{ \alpha_{13} \log(1 - e^{-\lambda t_{1i}}) + \alpha_2 \log(1 - e^{-\lambda t_{2i}}) \right\} \right].
\end{aligned}$$

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

$$\begin{aligned}
& \sum_{i \in I_2} \left[\delta_{1i} \delta_{2i} \left\{ \nu_1 \left(2 \log \lambda + \log \alpha_1 + \log \alpha_3 + (\alpha_1 - 1) \log(1 - e^{-\lambda t_{1i}}) \right. \right. \right. \\
& \quad \left. \left. \left. + (\alpha_{23} - 1) \log(1 - e^{-\lambda t_{2i}}) - \lambda(t_{1i} + t_{2i}) \right) + \nu_2 \left(2 \log \lambda + \log \alpha_1 + \log \alpha_2 \right. \right. \right. \\
& \quad \left. \left. \left. + (\alpha_1 - 1) \log(1 - e^{-\lambda t_{1i}}) + (\alpha_{23} - 1) \log(1 - e^{-\lambda t_{2i}}) - \lambda(t_{1i} + t_{2i}) \right) \right\} \right. \\
& \quad \left. + \delta_{1i} (1 - \delta_{2i}) \left\{ \log \lambda + \log \alpha_1 - \lambda t_{1i} + (\alpha_1 - 1) \log(1 - e^{-\lambda t_{1i}}) + \alpha_{23} \log(1 - e^{-\lambda t_{2i}}) \right\} \right. \\
& \quad \left. + (1 - \delta_{1i}) \delta_{2i} \left\{ \nu_1 \left(\log \lambda + \log \alpha_3 - \lambda t_{2i} + (\alpha_{23} - 1) \log(1 - e^{-\lambda t_{2i}}) \right. \right. \right. \\
& \quad \left. \left. \left. + \log[(1 - e^{-\lambda t_{1i}})^{\alpha_1} - (1 - e^{-\lambda t_{2i}})^{\alpha_1}] \right) + \nu_2 \left(\log \lambda + \log \alpha_2 - \lambda t_{2i} \right. \right. \right. \\
& \quad \left. \left. \left. + (\alpha_{23} - 1) \log(1 - e^{-\lambda t_{2i}}) + \log[(1 - e^{-\lambda t_{1i}})^{\alpha_1} - (1 - e^{-\lambda t_{2i}})^{\alpha_1}] \right) \right\} \right. \\
& \quad \left. + (1 - \delta_{1i})(1 - \delta_{2i}) \left\{ \alpha_1 \log(1 - e^{-\lambda t_{1i}}) + \alpha_{23} \log(1 - e^{-\lambda t_{2i}}) \right\} \right].
\end{aligned}$$

Let

$$\begin{aligned}
N_\lambda &= n_{11}^0 + n_{10}^0 + n_{01}^0 + 2n_{11}^1 + n_{10}^1 + n_{01}^1 + 2n_{11}^2 + n_{10}^2 + n_{01}^2, \\
C_1(\alpha_1, \alpha_3) &= n_{10}^0 + \mu_2(n_{11}^1 + n_{10}^1) + (n_{11}^2 + n_{10}^2), \\
C_2(\alpha_2, \alpha_3) &= n_{01}^0 + (n_{11}^1 + n_{01}^1) + \nu_2(n_{11}^2 + n_{01}^2), \\
C_3(\alpha_1, \alpha_2, \alpha_3) &= n_{11}^0 + \mu_1(n_{11}^1 + n_{10}^1) + \nu_1(n_{11}^2 + n_{01}^2).
\end{aligned} \tag{10}$$

Using equations (10), the pseudo log-likelihood is given by

$$\begin{aligned}
& N_\lambda \log \lambda + C_1(\alpha_1, \alpha_3) \log \alpha_1 + C_2(\alpha_2, \alpha_3) \log \alpha_2 + C_3(\alpha_1, \alpha_2, \alpha_3) \log \alpha_3 \\
& + \alpha_1 \left[\sum_{i \in I_0} \log(1 - e^{-\lambda t_i}) + \sum_{i \in I_1} \log(1 - e^{-\lambda t_{1i}}) + \sum_{i \in I_2} (1 - \delta_{2i} + \delta_{1i} \delta_{2i}) \log(1 - e^{-\lambda t_{1i}}) \right] \\
& + \alpha_2 \left[\sum_{i \in I_0} \log(1 - e^{-\lambda t_i}) + \sum_{i \in I_1} (1 - \delta_{1i} + \delta_{1i} \delta_{2i}) \log(1 - e^{-\lambda t_{2i}}) + \sum_{i \in I_2} \log(1 - e^{-\lambda t_{2i}}) \right] \\
& + \alpha_3 \left[\sum_{i \in I_0} \log(1 - e^{-\lambda t_i}) + \sum_{i \in I_1} \log(1 - e^{-\lambda t_{1i}}) + \sum_{i \in I_2} \log(1 - e^{-\lambda t_{2i}}) \right] \\
& - \lambda \left[\sum_{i \in I_0} (\delta_{1i} + \delta_{2i} - \delta_{1i} \delta_{2i}) t_i + \sum_{i \in I_1 \cup I_2} (\delta_{1i} t_{1i} + \delta_{2i} t_{2i}) \right] \\
& - \sum_{i \in I_0} (\delta_{1i} + \delta_{2i} - \delta_{1i} \delta_{2i}) \log(1 - e^{-\lambda t_i}) \\
& - \sum_{i \in I_1 \cup I_2} [\delta_{1i} \log(1 - e^{-\lambda t_{1i}}) + \delta_{2i} \log(1 - e^{-\lambda t_{2i}})] \\
& + \sum_{i \in I_1} \delta_{1i} (1 - \delta_{2i}) \log[(1 - e^{-\lambda t_{2i}})^{\alpha_2} - (1 - e^{-\lambda t_{1i}})^{\alpha_2}] \\
& + \sum_{i \in I_2} (1 - \delta_{1i}) \delta_{2i} \log[(1 - e^{-\lambda t_{1i}})^{\alpha_1} - (1 - e^{-\lambda t_{2i}})^{\alpha_1}]. \tag{11}
\end{aligned}$$

In order to implement the ‘M’-step of the EM algorithm, one needs to maximise the pseudo log likelihood equation (11) w.r.t $\alpha_1, \alpha_2, \alpha_3, \lambda$. We denote the first derivatives of the pseudo log-likelihood function as

$$\frac{\partial \log L}{\partial \alpha_i} = g_i, \quad i = 1, 2, 3, \quad \text{and} \quad \frac{\partial \log L}{\partial \lambda} = g_4,$$

where g_i ’s are defined in the Appendix B. Then M.L. equations are given by $g_i = 0, i = 1, \dots, 4$.

We observe that there is no explicit solution of any of the M. L. equations as parameters are interrelated. To maximize the pseudo log-likelihood function we use the method of fixed point iteration and solve for $\alpha_1, \alpha_2, \alpha_3$ and λ as fixed points of

$$\begin{aligned}
g_{\alpha_1}(\alpha_1, \alpha_3, \lambda) &= \alpha_1 & g_{\alpha_2}(\alpha_2, \alpha_3, \lambda) &= \alpha_2, \\
g_{\alpha_3}(\alpha_1, \alpha_2, \alpha_3, \lambda) &= \alpha_3, & g_\lambda(\alpha_1, \alpha_2, \alpha_3, \lambda) &= \lambda,
\end{aligned}$$

respectively. Basically, we solve four separate fixed point iterations given by

$$\begin{aligned}
g_{\alpha_1}(\alpha_1, \alpha_3, \lambda) &= \frac{C_1(\alpha_1, \alpha_3)}{h_1(\alpha_1, \lambda)}, \\
g_{\alpha_2}(\alpha_2, \alpha_3, \lambda) &= \frac{C_2(\alpha_2, \alpha_3)}{h_2(\alpha_2, \lambda)}, \\
g_{\alpha_3}(\alpha_1, \alpha_2, \alpha_3, \lambda) &= \frac{C_3(\alpha_1, \alpha_2, \alpha_3)}{h_3(\lambda)}, \\
g_{\lambda}(\alpha_1, \alpha_2, \alpha_3, \lambda) &= \frac{N_{\lambda}}{h_4(\alpha_1, \alpha_2, \alpha_3, \lambda)},
\end{aligned} \tag{12}$$

where N_{λ} , $C_1(\alpha_1, \alpha_3)$, $C_2(\alpha_2, \alpha_3)$ and $C_3(\alpha_1, \alpha_2, \alpha_3)$ are defined in equations (10) and $h_1(\alpha_1, \lambda)$, $h_2(\alpha_2, \lambda)$, $h_3(\lambda)$ and $h_4(\alpha_1, \alpha_2, \alpha_3, \lambda)$ are given in the Appendix B. Because of the involvements of the other parameters in the iterations, these separate iterative procedures are again repeated for overall maximisation. Kundu and Gupta (2009) used this idea to solve one fixed point equation which has been extended to solving four such equations simultaneously.

We implement the ‘M’ part of the EM algorithm as follows:

Let $\alpha_1^{(0)}$, $\alpha_2^{(0)}$, $\alpha_3^{(0)}$ and $\lambda^{(0)}$ be the initial estimates of α_1 , α_2 , α_3 and λ . We write $\alpha^{(i)}$, $\lambda_0^{(i)}$, $\lambda_1^{(i)}$, and $\lambda_2^{(i)}$ as estimates of α_1 , α_2 , α_3 , and λ at the i^{th} iteration. Given $\alpha^{(i)}$, $\lambda_0^{(i)}$, $\lambda_1^{(i)}$, and $\lambda_2^{(i)}$, the $(i+1)^{\text{th}}$ step estimates are obtained using the following algorithm.

Algorithm:

- (1) Compute $\mu_1, \mu_2, \nu_1, \nu_2$ using $\alpha^{(i)}$, $\lambda_0^{(i)}$, $\lambda_1^{(i)}$, $\lambda_2^{(i)}$.
- (2) Solve $g_{\lambda}(\alpha_1, \alpha_2, \alpha_3, \lambda) = \lambda$ for λ iteratively starting with $\alpha_1^{(i)}$, $\alpha_2^{(i)}$, $\alpha_3^{(i)}$, and $\lambda^{(i)}$; let the estimate be $\lambda_*^{(i+1)}$.
- (3) Similarly for fixed $\alpha_3^{(i)}$ and $\lambda_*^{(i+1)}$, solve $g_{\alpha_1}(\alpha_1, \alpha_3, \lambda) = \alpha_1$ for α_1 and $g_{\alpha_2}(\alpha_2, \alpha_3, \lambda) = \alpha_2$ for α_2 starting from $\alpha_1^{(i)}$ and $\alpha_2^{(i)}$ respectively. Denote the estimates as $\alpha_{1*}^{(i+1)}$ and $\alpha_{2*}^{(i+1)}$.
- (4) For fixed $\alpha_{1*}^{(i+1)}$, $\alpha_{2*}^{(i+1)}$ and $\lambda_*^{(i+1)}$, solve $g_{\alpha_3}(\alpha_1, \alpha_2, \alpha_3, \lambda) = \alpha_3$ for α_3 and $\alpha_3^{(i)}$ as initial estimate; let the solution be $\alpha_{3*}^{(i+1)}$.
- (5) If

$$|\alpha_{1*}^{(i+1)} - \alpha_1^{(i)}| + |\alpha_{2*}^{(i+1)} - \alpha_2^{(i)}| + |\alpha_{3*}^{(i+1)} - \alpha_3^{(i)}| + |\lambda_*^{(i+1)} - \lambda^{(i)}| < \epsilon \tag{13}$$

for a small $\epsilon > 0$, then $(\alpha_{1*}^{(i+1)}, \alpha_{2*}^{(i+1)}, \alpha_{3*}^{(i+1)}, \lambda_*^{(i+1)}) = (\alpha_1^{(i+1)}, \alpha_2^{(i+1)}, \alpha_3^{(i+1)}, \lambda^{(i+1)})$.

- (6) If (13) is not satisfied for a prefixed ϵ , take $(\alpha_{1*}^{(i+1)}, \alpha_{2*}^{(i+1)}, \alpha_{3*}^{(i+1)})$, $\lambda_*^{(i+1)} = (\alpha_1^{(i)}, \alpha_2^{(i)}, \alpha_3^{(i)}, \lambda^{(i)})$ and repeat steps 1-5.

4. NUMERICAL EXPERIMENT

In this section, we discuss results of some numerical experiments to observe how the proposed EM algorithm works in case of finite samples and different set of parameters. We plan our experiment such that certain percentage of data are randomly left censored. We have performed experiments for sample sizes $n = 50, 75$ and 100 . In each case, we replicated the process 2500 times. The uncensored data (X_1, X_2) are first generated as $BVGE(\alpha_1, \alpha_2, \alpha_3, \lambda)$ and we assume that the censoring variables Y_1, Y_2 are jointly distributed as $BVGE(\alpha_1^*, \alpha_2^*, \alpha_3^*, \lambda)$. Note that in this case

$$P[X_1 > Y_1] = \frac{\alpha_1 + \alpha_3}{\alpha_1 + \alpha_3 + \alpha_1^* + \alpha_3^*}, \quad P[X_2 > Y_2] = \frac{\alpha_2 + \alpha_3}{\alpha_2 + \alpha_3 + \alpha_2^* + \alpha_3^*}.$$

We have conducted several experiments by varying one parameter value and keeping fixed, the other parameters. We are reporting the following two cases.

- (1) $\alpha_1 = \alpha_2 = \alpha_3 = 1.0$ and $\lambda = .25, .5, .75$ and 1.0 .
- (2) $\alpha_1 = \alpha_2 = \lambda = 1.0$ and $\alpha_3 = .25, .5, .75$ and 1.0 .

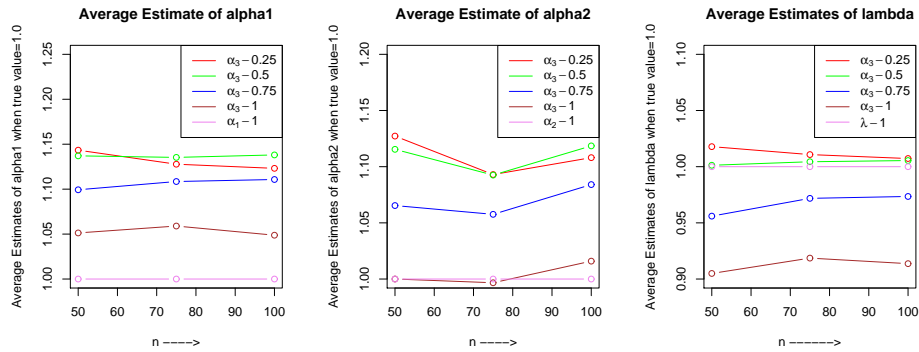
In case of experiment 1, we have used $\alpha_1^* = \alpha_2^* = \alpha_3^* = .1$ and $.18$ which implies that, respectively, 9% and 15% data are censored in X_1 as well as X_2 . In each replication, we first calculated the estimates of the unknown parameters. Then using the observed Fisher information matrix given in Appendix B, we calculated the confidence intervals and checked whether it contains the true value in case of each parameter. Thus we obtained the average estimator (AVEST), mean squared errors (MSE), average lengths of confidence (AVLEN) intervals and coverage probabilities (COVP) by calculating the proportion covering the true values. They are reported in Tables 1 and 2 in case of experiment 1 when $\lambda = .5$ and 1.0 respectively and are presented through Figures 1-3 in case of experiment 2.

Some of the salient features of the numerical experiments for case 1 ($\lambda = .5$ and 1.0) are given below.

- (i) We observe that the average estimators of all the four parameters $\alpha_1, \alpha_2, \alpha_3$ and λ are very close to the true value for both choices of the shape parameter λ . The estimators for α_1 have a positive bias for both cases of censoring and $\lambda = .5$ and 1.0 . However, the estimators for $\alpha_2, \alpha_3, \lambda$ have a negative bias for all cases. The results are similar for sample sizes 50 and 100 .
- (ii) The mean square error of the estimators decreases with increase in sample size. The value of scale parameter λ and the percentage of censoring makes no change in its numerical value.

TABLE 1. Average estimates, mean squared errors, coverage probability and average length of the confidence interval when $\lambda = .5$.

Sample Size \rightarrow		$N = 50$		$N = 100$	
Para- meter	Censoring	AVEST (MSE)	AVLEN (COVP)	AVEST (MSE)	AVLEN (COVP)
α_1	9%	1.05717 (6.228e-2)	.93991 (.9524)	1.05906 (3.727e-2)	.66473 (.9348)
	15%	1.01860 (5.923e-2)	.93750 (.9488)	1.00954 (3.824e-2)	.65088 (.9460)
α_2	9%	.99064 (6.608e-2)	.96248 (.8976)	1.00126 (3.788e-2)	.68412 (.9064)
	15%	.94639 (7.777e-2)	.94564 (.8424)	.96258 (4.676e-2)	.67104 (.8988)
α_3	9%	.93319 (4.673e-2)	.81187 (.8952)	.94331 (2.508e-2)	.57700 (.9172)
	15%	.86888 (5.748e-2)	.79643 (.8420)	.86188 (4.447e-2)	.55107 (.8356)
λ	9%	.45456 (5.105e-3)	.23469 (.9148)	.46115 (4.077e-3)	.15743 (.9392)
	15%	.43103 (8.272e-3)	.22900 (.8340)	.43006 (8.573e-3)	.14984 (.8132)

FIGURE 1. Experiment 4: Average Estimates of α_1 , α_2 and λ for different values of α_3 .

(iii) When $n = 100$. average length of confidence intervals for the three shape parameters is considerably lower and the coverage probabilities are slightly higher compared to the case when $n = 50$.

TABLE 2. Average estimates, mean squared errors, coverage probability and average length of the confidence interval when $\lambda = 1.0$.

Sample Size \rightarrow		$N = 50$		$N = 100$	
Para- meter	Censoring	AVEST (MSE)	AVLEN (COVP)	AVEST (MSE)	AVLEN (COVP)
α_1	9%	1.05575 (6.027e-2)	.93971 (.9500)	1.06290 (3.658e-2)	.67149 (.9384)
	15%	1.01987 (5.771e-2)	.94084 (.9512)	1.01488 (3.591e-2)	.65672 (.9468)
α_2	9%	.98918 (6.619e-2)	.96572 (.8948)	1.00428 (3.699e-2)	.69521 (.9048)
	15%	.94516 (7.651e-2)	.92688 (.8420)	.96759 (4.642e-2)	.67572 (.8976)
α_3	9%	.93718 (4.698e-2)	.82087 (.8952)	.94822 (2.279e-2)	.58481 (.9152)
	15%	.87016 (5.627e-2)	.78999 (.8444)	.86632 (4.184e-2)	.55676 (.8344)
λ	9%	.90957 (2.020e-2)	.47442 (.9148)	.92535 (1.478e-2)	.32379 (.9360)
	15%	.86240 (3.222e-2)	.45517 (.8340)	.86454 (3.145e-2)	.30633 (.8140)

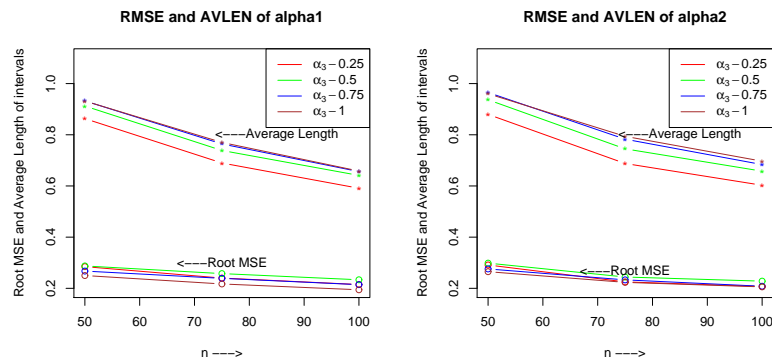


FIGURE 2. Experiment 4: Root MSE and Average Lengths of Intervals of α_1 and α_2 .

The percentage of censoring has no effect on the average length of confidence intervals and coverage probability.

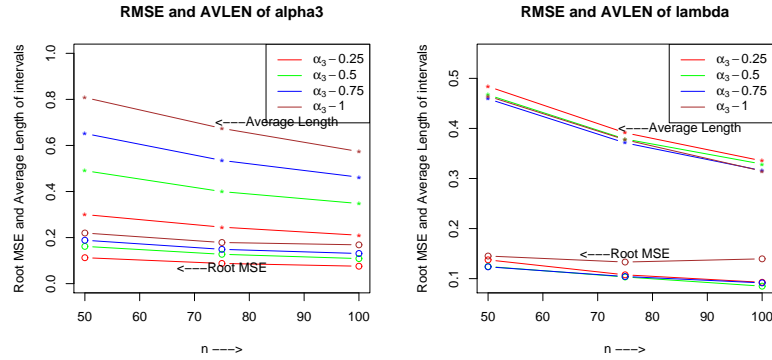


FIGURE 3. Experiment 4: Root MSE and Average Lengths of Intervals of α_3 and λ .

The average length of confidence interval of λ increases with increase in values of λ . The coverage probabilities of confidence intervals for all parameters are not influenced by values of λ .

Similar conclusions hold for the experiment 2 and the results have been shown in Figures 1-3.

We have also considered the cases where $\alpha_2 = \alpha_3 = \lambda = 1.0$ are unchanged and varying $\alpha_1 = .25, .5, .75$ and 1.0. and $\alpha_1 = \alpha_3 = \lambda = 1.0$ are fixed and $\alpha_2 = .25, .5, .75$ and 1.0. The results are on similar lines and are not being reported separately.

5. DATA ANALYSIS

Next we analyse the soccer data for the years 2004-05 and 2005-06 which was studied by Meintanis (2007). Kundu and Gupta (2009) have fitted BVGE distribution and bivariate Marshall-Olkin (1967) distribution to this data. They showed that BVGE gives a better fit. In order to bring out the effect of random left censoring on the M.L.E.'s of the parameters we introduce censoring artificially and then estimate the four unknown parameters using the EM algorithm proposed in Section 3.

The data (X_1, X_2) consists of 37 data points. We assume that the pair (X_1, X_2) has $BVGE(\tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\alpha}_3, \tilde{\lambda})$. The pair of censoring random variables (Z_1, Z_2) has $BVGE(\alpha_1^*, \alpha_2^*, \alpha_3^*, \tilde{\lambda})$. In order to ensure that $P(X_1 < Z_1) = P(X_2 < Z_2) = .1$, we take $(\alpha_1^*, \alpha_2^*, \alpha_3^*) = (.14, .03, .15)$ and $\tilde{\alpha}_1 = 1.445, \tilde{\alpha}_2 = 0.468, \tilde{\alpha}_3 = 1.17$, and $\tilde{\lambda} = 0.039$ are the estimates of $\alpha_1, \alpha_2, \alpha_3, \lambda$, obtained by Kundu and Gupta (2009). In a similar way, $(\alpha_1^*, \alpha_2^*, \alpha_3^*) = (.20, .07, .15)$ ensures that $P(X_1 < Z_1) = P(X_2 < Z_2) = .25$. We have used the proposed EM algorithm to estimate the unknown parameters and the initial estimates used for $\alpha_1, \alpha_2, \alpha_3$, and λ , respectively, were

.75, .75, .75 and .5 in both the cases. The point estimates and the corresponding confidence intervals for 10 % and 25 % censoring are reported in Table 3.

TABLE 3. The point estimates and confidence interval for football data .

Parameter	10% Censoring		25% Censoring	
	Point Est.	Conf. Int.	Point Est.	Conf. Int.
α_1	1.435	(.785, 2.086)	1.565	(.868, 2.262)
α_2	.394	(.1415, .646)	.613	(.261, .966)
α_3	1.094	(.632, 1.557)	1.018	(.569, 1.467)
λ	.0369	(.0287, .0451)	.0384	(.0300, .0467)

6. DISCUSSION

In this paper we have considered the M.L.E.'s of the three shape parameters and the scale parameter of BVGE distribution when both components of the bivariate variable are subject to random left censoring. We use the expectation-conditional maximization algorithm for finding the M.L.E.'s iteratively. The steps involved in the iteration procedure are clearly stated. The results of the numerical experiment indicate that the EM algorithm performs very well for sample sizes 50, 75 and 100 and also for various levels of random censoring that we have studied. The asymptotic confidence intervals give accurate results even for moderate sample sizes and hence can be used for testing purposes.

The proposed tools should prove useful in handling left censored data arising in epidemic models - which are naturally skewed.

APPENDIX A: CONTRIBUTIONS TO THE LIKELIHOOD ON THE SETS I_0 , I_1 AND I_2

The contribution to the likelihood on the set I_0 is

$$\begin{aligned} & \prod_{i \in I_0} L(t_i, \delta_{1i}; \delta_{2i}) \\ = & \prod_{i \in I_0} \left\{ [\lambda \alpha_3 e^{-\lambda t_i} (1 - e^{-\lambda t_i})^{\alpha_{123}-1}]^{\delta_{1i} \delta_{2i}} [\lambda \alpha_1 e^{-\lambda t_i} (1 - e^{-\lambda t_i})^{\alpha_{123}-1}]^{\delta_{1i}(1-\delta_{2i})} \right. \\ & \left. [\lambda \alpha_2 e^{-\lambda t_i} (1 - e^{-\lambda t_i})^{\alpha_{123}-1}]^{(1-\delta_{1i})\delta_{2i}} [(1 - e^{-\lambda t_i})^{\alpha_{123}}]^{(1-\delta_{1i})(1-\delta_{2i})} \right\}. \end{aligned}$$

Similarly, the contribution to the likelihood on the set I_1 is

$$\begin{aligned} & \prod_{i \in I_1} L(t_{1i}, \delta_{1i}; t_{2i}, \delta_{2i}) \\ = & \prod_{i \in I_1} \left\{ [\lambda^2 \alpha_{13} \alpha_2 e^{-\lambda(t_{1i}+t_{2i})} (1 - e^{-\lambda t_{1i}})^{\alpha_{13}-1} (1 - e^{-\lambda t_{2i}})^{\alpha_2-1}]^{\delta_{1i} \delta_{2i}} \right. \\ & [\lambda \alpha_{13} e^{-\lambda t_{1i}} (1 - e^{-\lambda t_{1i}})^{\alpha_{13}-1} \{ (1 - e^{-\lambda t_{2i}})^{\alpha_2} - (1 - e^{-\lambda t_{1i}})^{\alpha_2} \}]^{\delta_{1i}(1-\delta_{2i})} \\ & [\lambda \alpha_2 e^{-\lambda t_{2i}} (1 - e^{-\lambda t_{1i}})^{\alpha_{13}} (1 - e^{-\lambda t_{2i}})^{\alpha_2-1}]^{(1-\delta_{1i})\delta_{2i}} \\ & \left. [(1 - e^{-\lambda t_{1i}})^{\alpha_{13}} (1 - e^{-\lambda t_{2i}})^{\alpha_2}]^{(1-\delta_{1i})(1-\delta_{2i})} \right\}, \end{aligned} \tag{14}$$

and the contribution to the likelihood on the set I_2 is

$$\begin{aligned} & \prod_{i \in I_2} L(t_{1i}, \delta_{1i}; t_{2i}, \delta_{2i}) \\ = & \prod_{i \in I_2} \left\{ [\lambda^2 \alpha_1 \alpha_{23} e^{-\lambda(t_{1i}+t_{2i})} (1 - e^{-\lambda t_{1i}})^{\alpha_1-1} (1 - e^{-\lambda t_{2i}})^{\alpha_{23}-1}]^{\delta_{1i} \delta_{2i}} \right. \\ & [\lambda \alpha_1 e^{-\lambda t_{1i}} (1 - e^{-\lambda t_{1i}})^{\alpha_1-1} (1 - e^{-\lambda t_{2i}})^{\alpha_{23}}]^{\delta_{1i}(1-\delta_{2i})} \\ & [\lambda \alpha_{23} e^{-\lambda t_{2i}} (1 - e^{-\lambda t_{2i}})^{\alpha_{23}-1} \{ (1 - e^{-\lambda t_{1i}})^{\alpha_1} - (1 - e^{-\lambda t_{2i}})^{\alpha_1} \}]^{(1-\delta_{1i})\delta_{2i}} \\ & \left. [(1 - e^{-\lambda t_{1i}})^{\alpha_1} (1 - e^{-\lambda t_{2i}})^{\alpha_{23}}]^{(1-\delta_{1i})(1-\delta_{2i})} \right\}. \end{aligned} \tag{15}$$

APPENDIX B: OBSERVED FISHER INFORMATION MATRIX OF MAXIMUM LIKELIHOOD ESTIMATES

Here we find the observed Fisher information matrix using the procedure described in Louis (1982) for using the EM algorithm to obtain the M.L.E.'s for incomplete data problem. The observed Fisher information matrix helped us to find asymptotic confidence intervals in numerical experiment in section 4. Let $\mathbf{g} = (g_1, g_2, g_3, g_4)^T$ denote the gradient vector and $\mathbf{H} = ((H_{ij}))$, the Hessian matrix of the pseudo log-likelihood function defined in (11). Then, using $C_1(\alpha_1, \alpha_3)$,

$C_2(\alpha_2, \alpha_3)$, $C_3(\alpha_1, \alpha_2, \alpha_3)$ and N_λ , the elements of vector \mathbf{g} are as follows:

$$\begin{aligned} g_1 &= \frac{1}{\alpha_1} C_1(\alpha_1, \alpha_3) + h_1(\alpha_1, \lambda), & g_2 &= \frac{1}{\alpha_2} C_2(\alpha_2, \alpha_3) + h_2(\alpha_2, \lambda), \\ g_3 &= \frac{1}{\alpha_3} C_3(\alpha_1, \alpha_2, \alpha_3) + h_3(\lambda), & g_4 &= \frac{1}{\lambda} N_\lambda - h_4(\alpha_1, \alpha_2, \alpha_3, \lambda). \end{aligned}$$

where

$$\begin{aligned} h_1(\alpha_1, \lambda) &= \left[\sum_{i \in I_0} \log(1 - e^{-\lambda t_i}) + \sum_{i \in I_1} \log(1 - e^{-\lambda t_{1i}}) + \sum_{i \in I_2} (1 - \delta_{2i} + \delta_{1i} \delta_{2i}) \log(1 - e^{-\lambda t_{1i}}) \right] \\ &+ \sum_{i \in I_2} (1 - \delta_{1i}) \delta_{2i} \left[\frac{(1 - e^{-\lambda t_{1i}})^{\alpha_1} \log(1 - e^{-\lambda t_{1i}}) - (1 - e^{-\lambda t_{2i}})^{\alpha_1} \log(1 - e^{-\lambda t_{2i}})}{(1 - e^{-\lambda t_{1i}})^{\alpha_1} - (1 - e^{-\lambda t_{2i}})^{\alpha_1}} \right], \\ h_2(\alpha_2, \lambda) &= \left[\sum_{i \in I_0} \log(1 - e^{-\lambda t_i}) + \sum_{i \in I_1} (1 - \delta_{1i} + \delta_{1i} \delta_{2i}) \log(1 - e^{-\lambda t_{2i}}) + \sum_{i \in I_2} \log(1 - e^{-\lambda t_{2i}}) \right] \\ &+ \sum_{i \in I_1} \delta_{1i} (1 - \delta_{2i}) \left[\frac{(1 - e^{-\lambda t_{2i}})^{\alpha_2} \log(1 - e^{-\lambda t_{2i}}) - (1 - e^{-\lambda t_{1i}})^{\alpha_2} \log(1 - e^{-\lambda t_{1i}})}{(1 - e^{-\lambda t_{2i}})^{\alpha_2} - (1 - e^{-\lambda t_{1i}})^{\alpha_2}} \right], \\ h_3(\lambda) &= \left[\sum_{i \in I_0} \log(1 - e^{-\lambda t_i}) + \sum_{i \in I_1} \log(1 - e^{-\lambda t_{1i}}) + \sum_{i \in I_2} \log(1 - e^{-\lambda t_{2i}}) \right], \\ h_4(\alpha_1, \alpha_2, \alpha_3, \lambda) &= \sum_{i \in I_0} (\delta_{1i} + \delta_{2i} - \delta_{1i} \delta_{2i}) t_i + \sum_{i \in I_1 \cup I_2} (\delta_{1i} t_{1i} + \delta_{2i} t_{2i}) \\ &- \alpha_1 \left[\sum_{i \in I_0} \frac{t_i e^{-\lambda t_i}}{(1 - e^{-\lambda t_i})} + \sum_{i \in I_1} \frac{t_{1i} e^{-\lambda t_{1i}}}{(1 - e^{-\lambda t_{1i}})} + \sum_{i \in I_2} (1 - \delta_{2i} + \delta_{1i} \delta_{2i}) \frac{t_{1i} e^{-\lambda t_{1i}}}{(1 - e^{-\lambda t_{1i}})} \right] \\ &- \alpha_2 \left[\sum_{i \in I_0} \frac{t_i e^{-\lambda t_i}}{(1 - e^{-\lambda t_i})} + \sum_{i \in I_1} (1 - \delta_{1i} + \delta_{1i} \delta_{2i}) \frac{t_{2i} e^{-\lambda t_{2i}}}{(1 - e^{-\lambda t_{2i}})} + \sum_{i \in I_2} \frac{t_{2i} e^{-\lambda t_{2i}}}{(1 - e^{-\lambda t_{2i}})} \right] \\ &- \alpha_3 \left[\sum_{i \in I_0} \frac{t_i e^{-\lambda t_i}}{(1 - e^{-\lambda t_i})} + \sum_{i \in I_1} \frac{t_{1i} e^{-\lambda t_{1i}}}{(1 - e^{-\lambda t_{1i}})} + \sum_{i \in I_2} \frac{t_{2i} e^{-\lambda t_{2i}}}{(1 - e^{-\lambda t_{2i}})} \right] \end{aligned}$$

$$\begin{aligned}
& + \sum_{i \in I_0} (\delta_{1i} + \delta_{2i} - \delta_{1i}\delta_{2i}) \frac{t_i e^{-\lambda t_i}}{(1 - e^{-\lambda t_i})} \\
& + \sum_{i \in I_1 \cup I_2} \delta_{1i} \frac{t_{1i} e^{-\lambda t_{1i}}}{(1 - e^{-\lambda t_{1i}})} + \sum_{i \in I_1 \cup I_2} \delta_{2i} \frac{t_{2i} e^{-\lambda t_{2i}}}{(1 - e^{-\lambda t_{2i}})} \\
& - \alpha_2 \sum_{i \in I_1} \delta_{1i} (1 - \delta_{2i}) \left[\frac{t_{2i} e^{-\lambda t_{2i}} (1 - e^{-\lambda t_{2i}})^{\alpha_2 - 1} - t_{1i} e^{-\lambda t_{1i}} (1 - e^{-\lambda t_{1i}})^{\alpha_2 - 1}}{(1 - e^{-\lambda t_{2i}})^{\alpha_2} - (1 - e^{-\lambda t_{1i}})^{\alpha_2}} \right] \\
& - \alpha_1 \sum_{i \in I_2} \delta_{2i} (1 - \delta_{1i}) \left[\frac{t_{1i} e^{-\lambda t_{1i}} (1 - e^{-\lambda t_{1i}})^{\alpha_1 - 1} - t_{2i} e^{-\lambda t_{2i}} (1 - e^{-\lambda t_{2i}})^{\alpha_1 - 1}}{(1 - e^{-\lambda t_{1i}})^{\alpha_1} - (1 - e^{-\lambda t_{2i}})^{\alpha_1}} \right].
\end{aligned}$$

The Hessian matrix \mathbf{H} is symmetric, so $H_{ij} = H_{ji}, i > j$. Its elements are given below.

$$\begin{aligned}
H_{11} &= -\frac{C_1(\alpha_1, \alpha_3)}{\alpha_1^2} + \sum_{i \in I_2} \delta_{2i}(1 - \delta_{1i}) \times \\
&\quad \left[\frac{(1 - e^{-\lambda t_{1i}})^{\alpha_1} (\log(1 - e^{-\lambda t_{1i}}))^2 - (1 - e^{-\lambda t_{2i}})^{\alpha_1} (\log(1 - e^{-\lambda t_{2i}}))^2}{(1 - e^{-\lambda t_{1i}})^{\alpha_1} - (1 - e^{-\lambda t_{2i}})^{\alpha_1}} \right. \\
&\quad \left. - \frac{\{(1 - e^{-\lambda t_{1i}})^{\alpha_1} \log(1 - e^{-\lambda t_{1i}}) - (1 - e^{-\lambda t_{2i}})^{\alpha_1} \log(1 - e^{-\lambda t_{2i}})\}^2}{\{(1 - e^{-\lambda t_{1i}})^{\alpha_1} - (1 - e^{-\lambda t_{2i}})^{\alpha_1}\}^2} \right], \\
H_{12} &= H_{13} = H_{23} = 0, \\
H_{14} &= \left[\sum_{i \in I_0} \frac{t_i e^{-\lambda t_i}}{(1 - e^{-\lambda t_i})} + \sum_{i \in I_1} \frac{t_{1i} e^{-\lambda t_{1i}}}{(1 - e^{-\lambda t_{1i}})} + \sum_{i \in I_2} (1 - \delta_{2i} + \delta_{1i} \delta_{2i}) \frac{t_{1i} e^{-\lambda t_{1i}}}{(1 - e^{-\lambda t_{1i}})} \right] \\
&\quad + \sum_{i \in I_2} \delta_{2i}(1 - \delta_{1i}) \left[\frac{t_{1i} e^{-\lambda t_{1i}} (1 - e^{-\lambda t_{1i}})^{\alpha_1 - 1} (\alpha_1 \log(1 - e^{-\lambda t_{1i}}) + 1)}{(1 - e^{-\lambda t_{1i}})^{\alpha_1} - (1 - e^{-\lambda t_{2i}})^{\alpha_1}} \right. \\
&\quad - \frac{t_{2i} e^{-\lambda t_{2i}} (1 - e^{-\lambda t_{2i}})^{\alpha_1 - 1} (\alpha_1 \log(1 - e^{-\lambda t_{2i}}) + 1)}{(1 - e^{-\lambda t_{1i}})^{\alpha_1} - (1 - e^{-\lambda t_{2i}})^{\alpha_1}} \\
&\quad \left. - \left\{ \frac{(1 - e^{-\lambda t_{1i}})^{\alpha_1} \log(1 - e^{-\lambda t_{1i}}) - (1 - e^{-\lambda t_{2i}})^{\alpha_1} \log(1 - e^{-\lambda t_{2i}})}{(1 - e^{-\lambda t_{1i}})^{\alpha_1} - (1 - e^{-\lambda t_{2i}})^{\alpha_1}} \right\}^2 \right], \\
H_{22} &= -\frac{C_2(\alpha_2, \alpha_3)}{\alpha_2^2} + \sum_{i \in I_1} \delta_{1i}(1 - \delta_{2i}) \times \\
&\quad \left[\frac{(1 - e^{-\lambda t_{2i}})^{\alpha_2} (\log(1 - e^{-\lambda t_{2i}}))^2 - (1 - e^{-\lambda t_{1i}})^{\alpha_2} (\log(1 - e^{-\lambda t_{1i}}))^2}{(1 - e^{-\lambda t_{2i}})^{\alpha_2} - (1 - e^{-\lambda t_{1i}})^{\alpha_2}} \right. \\
&\quad \left. - \frac{\{(1 - e^{-\lambda t_{2i}})^{\alpha_2} \log(1 - e^{-\lambda t_{2i}}) - (1 - e^{-\lambda t_{1i}})^{\alpha_2} \log(1 - e^{-\lambda t_{1i}})\}^2}{\{(1 - e^{-\lambda t_{2i}})^{\alpha_2} - (1 - e^{-\lambda t_{1i}})^{\alpha_2}\}^2} \right], \\
H_{24} &= \left[\sum_{i \in I_0} \frac{t_i e^{-\lambda t_i}}{(1 - e^{-\lambda t_i})} + \sum_{i \in I_1} (1 - \delta_{1i} + \delta_{1i} \delta_{2i}) \frac{t_{2i} e^{-\lambda t_{2i}}}{(1 - e^{-\lambda t_{2i}})} + \sum_{i \in I_2} \frac{t_{2i} e^{-\lambda t_{2i}}}{(1 - e^{-\lambda t_{2i}})} \right] \\
&\quad + \sum_{i \in I_1} \delta_{1i}(1 - \delta_{2i}) \left[\frac{t_{2i} e^{-\lambda t_{2i}} (1 - e^{-\lambda t_{2i}})^{\alpha_2 - 1} (\alpha_2 \log(1 - e^{-\lambda t_{2i}}) + 1)}{(1 - e^{-\lambda t_{2i}})^{\alpha_2} - (1 - e^{-\lambda t_{1i}})^{\alpha_2}} \right. \\
&\quad - \frac{t_{1i} e^{-\lambda t_{1i}} (1 - e^{-\lambda t_{1i}})^{\alpha_2 - 1} (\alpha_2 \log(1 - e^{-\lambda t_{1i}}) + 1)}{(1 - e^{-\lambda t_{2i}})^{\alpha_2} - (1 - e^{-\lambda t_{1i}})^{\alpha_2}} \\
&\quad \left. - \left\{ \frac{(1 - e^{-\lambda t_{2i}})^{\alpha_2} \log(1 - e^{-\lambda t_{2i}}) - (1 - e^{-\lambda t_{1i}})^{\alpha_2} \log(1 - e^{-\lambda t_{1i}})}{(1 - e^{-\lambda t_{2i}})^{\alpha_2} - (1 - e^{-\lambda t_{1i}})^{\alpha_2}} \right\}^2 \right], \\
H_{33} &= -\frac{C_3(\alpha_1, \alpha_2, \alpha_3)}{\alpha_3^2}, \\
H_{34} &= \left[\sum_{i \in I_0} \frac{t_i e^{-\lambda t_i}}{(1 - e^{-\lambda t_i})} + \sum_{i \in I_1} \frac{t_{1i} e^{-\lambda t_{1i}}}{(1 - e^{-\lambda t_{1i}})} + \sum_{i \in I_2} \frac{t_{2i} e^{-\lambda t_{2i}}}{(1 - e^{-\lambda t_{2i}})} \right], \\
H_{44} &= -\frac{N_\lambda}{\lambda^2} - \alpha_1 \left[\sum_{i \in I_0} \frac{t_i^2 e^{\lambda t_i}}{(1 - e^{\lambda t_i})^2} + \sum_{i \in I_1} \frac{t_{1i}^2 e^{\lambda t_{1i}}}{(1 - e^{\lambda t_{1i}})^2} + \sum_{i \in I_2} (1 - \delta_{2i} + \delta_{1i} \delta_{2i}) \frac{t_{1i}^2 e^{\lambda t_{1i}}}{(1 - e^{\lambda t_{1i}})^2} \right]
\end{aligned}$$

$$\begin{aligned}
& -\alpha_2 \left[\sum_{i \in I_0} \frac{t_i^2 e^{\lambda t_i}}{(1 - e^{\lambda t_i})^2} + \sum_{i \in I_1} (1 - \delta_{1i} + \delta_{1i} \delta_{2i}) \frac{t_{2i}^2 e^{\lambda t_{2i}}}{(1 - e^{\lambda t_{2i}})^2} + \sum_{i \in I_2} \frac{t_{2i}^2 e^{\lambda t_{2i}}}{(1 - e^{\lambda t_{2i}})^2} \right] \\
& -\alpha_3 \left[\sum_{i \in I_0} \frac{t_i^2 e^{\lambda t_i}}{(1 - e^{\lambda t_i})^2} + \sum_{i \in I_1} \frac{t_{1i}^2 e^{\lambda t_{1i}}}{(1 - e^{\lambda t_{1i}})^2} + \sum_{i \in I_2} \frac{t_{2i}^2 e^{\lambda t_{2i}}}{(1 - e^{\lambda t_{2i}})^2} \right] \\
& + \left[\sum_{i \in I_0} ((\delta_{1i} + \delta_{2i} - \delta_{1i} \delta_{2i})) \frac{t_i^2 e^{\lambda t_i}}{(1 - e^{\lambda t_i})^2} + \sum_{i \in I_1 \cup I_2} \frac{\delta_{1i} t_{1i}^2 e^{\lambda t_{1i}}}{(1 - e^{\lambda t_{1i}})^2} + \sum_{i \in I_1 \cup I_2} \frac{\delta_{2i} t_{2i}^2 e^{\lambda t_{2i}}}{(1 - e^{\lambda t_{2i}})^2} \right] \\
& + \sum_{i \in I_1} \delta_{1i} (1 - \delta_{2i}) \left[\frac{\alpha_2 (\alpha_2 - 1) \{ t_{2i}^2 e^{-\lambda t_{2i}} (1 - e^{-\lambda t_{2i}})^{\alpha_2 - 2} - t_{1i}^2 e^{-\lambda t_{1i}} (1 - e^{-\lambda t_{1i}})^{\alpha_2 - 2} \}}{(1 - e^{-\lambda t_{2i}})^{\alpha_2} - (1 - e^{-\lambda t_{1i}})^{\alpha_2}} \right. \\
& \left. - \alpha_2^2 \left\{ \frac{(1 - e^{-\lambda t_{2i}})^{\alpha_2 - 1} t_{2i} e^{-\lambda t_{2i}} - (1 - e^{-\lambda t_{1i}})^{\alpha_2 - 1} t_{1i} e^{-\lambda t_{1i}}}{(1 - e^{-\lambda t_{2i}})^{\alpha_2} - (1 - e^{-\lambda t_{1i}})^{\alpha_2}} \right\}^2 \right], \\
& + \sum_{i \in I_2} \delta_{2i} (1 - \delta_{1i}) \left[\frac{\alpha_1 (\alpha_1 - 1) \{ t_{1i}^2 e^{-\lambda t_{1i}} (1 - e^{-\lambda t_{1i}})^{\alpha_1 - 2} - t_{2i}^2 e^{-\lambda t_{2i}} (1 - e^{-\lambda t_{2i}})^{\alpha_1 - 2} \}}{(1 - e^{-\lambda t_{1i}})^{\alpha_1} - (1 - e^{-\lambda t_{2i}})^{\alpha_1}} \right. \\
& \left. - \alpha_1^2 \left\{ \frac{(1 - e^{-\lambda t_{1i}})^{\alpha_1 - 1} t_{1i} e^{-\lambda t_{1i}} - (1 - e^{-\lambda t_{2i}})^{\alpha_1 - 1} t_{2i} e^{-\lambda t_{2i}}}{(1 - e^{-\lambda t_{1i}})^{\alpha_1} - (1 - e^{-\lambda t_{2i}})^{\alpha_1}} \right\}^2 \right].
\end{aligned}$$

REFERENCES

- [1] Bagger, J. (2005), "Wage growth and turnover in Denmark", University of Aarhus, Denmark.
- [2] Coburn, A. F., McBride, T. and Ziller, E. (2001), "Patterns of Health Insurance Coverage among Rural and Urban Children", Working Paper No. 26, Maine Rural Health Research Center, Edmund S. Muskie School of Public Service, University of Southern Maine, Portland.
- [3] Gupta, R. D. and Kundu, D. (1999), "Generalized exponential distributions", *Australian and New Zealand Journal of Statistics*, **41**, 173-188.
- [4] Gupta, R. D. and Kundu, D. (2007), "Generalized exponential distributions: existing results and some recent developments", *Journal of Statistical Planning Inference*, **137**, 3525-3536.
- [5] Jacqmin-Gadda, H., Thiébaud, R., Chene, G. and Commenges D. (2000), "Analysis of left-censored longitudinal data with application to viral load in HIV infection", *Biostatistics*, **1**, 355-368.
- [6] Kundu, D. and Gupta, R.D. (2009), "Bivariate generalized exponential distribution" *Journal of Multivariate Analysis*, **100**, 581 - 593.
- [7] 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.
- [8] Marshall, A.W. and Olkin, I. (1967). A multivariate exponential distribution. *J. Amer. Statist. Assoc.*, **62**, 30-44.
- [9] Meintanis, S.G. (2007). Test of fit for Marshall-Olkin distributions with applications. *Journal of Statistical Planning and Inference*, **137**, 3954-3963.
- [10] Mitra, S. and Kundu, D. (2008), "Analysis of the left censored data from the generalized exponential distribution", *Journal of Statistical Computation Simulation*, **78**, 669-679.
- [11] Sarhan, A. and Balakrishnan, N. (2007), "A new class of bivariate distribution and its mixture", *Journal of the Multivariate Analysis*, **98**, 1508-1527.

THEORETICAL STATISTICS AND MATHEMATICS UNIT, INDIAN STATISTICAL INSTITUTE, 7, S.J.S. SANSANWAL
MARG, NEW DELHI-110016

E-mail address: isha@isid.ac.in and nandi@isid.ac.in