EM Algorithm for Estimating the Burr XII Parameters in Partially Accelerated Life Tests

Yung-Fu Cheng
•
Research Center for Testing and Assessment
National Academy for Educational Research, Taiwan
yfjeng@gmail.com

Abstract

In this paper, I present maximum likelihood estimation via the expectation-maximization algorithm to estimate the Burr XII parameters and acceleration factor in step-stress partially accelerated life tests under multiple censored data. In addition, the asymptotic variance and covariance matrix of the estimators are derived by using the complete and missing information matrices, and confidence intervals of the parameters are obtained. The simulation results show that the maximum likelihood estimation via the expectation-maximization algorithm performs well in most cases in terms of the absolute relative bias, the root mean square error, and the coverage rate. Furthermore, a numerical example is also given to demonstrate the performance of the proposed method.

Keywords: partially accelerated life test, acceleration factor, Burr XII distribution, maximum likelihood estimation, EM algorithm

I. Introduction

Generally, life testing of products under normal conditions usually requires a long period of time. Long-term testing will increase the test cost and will take a lot of time. Accelerated life test (ALT) is one of the solutions that can avoid above problems. ALT has been successfully applied to obtain information about product life quickly and economically under more severe operating conditions. Stress conditions, such as, cycling rate, load, voltage, pressure, vibration, and temperature are the most common methods in practice. The acceleration factor in ALT is usually assumed to be a known value. On the contrary, the acceleration factor in partial accelerated life testing (PALT) is usually assumed as an unknown value. Constant stress, step stress and progressive stress are three major stress types of PALT. Progressive stress is a more complicated PALT approach among these major stress types. In a constant-stress test, test units are run at some unchanged constant level of stress. In a step-stress test, the level of stress can be changed at a specified time, and this kind of test method is called step-stress partially accelerated life test (SS-PALT).

The Burr XII distribution is widely applied in reliability engineering because of its many advantages. Rodriguez (1977) showed that the area in the ($\sqrt{\beta_1}, \beta_2$) plane corresponding to the Burr XII distribution is wide and it covers various well-known distributions. Zimmer et al. (1998) presented the statistical and probabilistic properties of the Burr XII distribution, and described its connection with other distributions used in reliability analysis. The Burr XII distribution has been applied in reliability analysis widely. Wingo (1993) formatted the MLE to fit the Burr XII distribution through the use of multiple censored data. Ali Mousa (1995) estimated the parameters of the Burr XII distribution with Type II censored data for an ALT model by using the Bayes method. Wang et al. (1996) presented the MLE for obtaining point and interval estimates of the Burr XII parameters. Watkins (1999) developed an algorithm for calculating the MLE of the three-parameter Burr XII distribution. As to the parameter estimation of the Burr XII distribution in SS-PALT, Abd-Elfattah et al. (2008) investigated the maximum likelihood method for the parameters of the Burr XII distribution in SS-PALT under type I censored data. Abdel-Ghaly et al. (2008) considered the estimation problem of the Burr-XII distribution in SS-PALT using censored data. Abdel-Hamid (2009) estimated the parameters of the Burr XII distribution with progressive Type II censoring for a CS-PALT model by using the MLE method. Cheng and Wang (2012) compared the performance of the maximum likelihood estimates of the Burr XII parameters for CS-PALT. So, it has been shown that the Burr XII distribution is a flexible model and is recommended for modeling in the reliability analysis and ALTs.

The MLE via the Newton-Raphson algorithm is very sensitive to its initial parameter estimation value. Other options can be adopted to avoid the above problem, for example, the expectation-maximization (EM) algorithm. EM algorithm is an iterative algorithm approach applied in a variety of incomplete data problems (Dempster et al., 1977). EM algorithm can be used in data sets with missing values, censored and grouped observations, or models with truncated distributions. EM algorithm involves two steps, the E-step and the M-step. In the E-step, the expected values of the complete data sufficient statistics are computed. In the M-step, parameter estimates that maximize the complete data likelihood are solved by using the conditional expected value that computed in the E-step. Both steps of the iterations are repeated until the parameter estimates converge. The development and application of EM algorithms are getting more and more mature. Louis (1982) derived a procedure for extracting the observed information matrix when EM algorithm is used to find maximum likelihood estimates in incomplete data problems. In reliability analysis, EM algorithm has been commonly used. Ng et al. (2002) presented the MLE via EM algorithm to estimate the lognormal and the Weibull parameters with progressively type II censored data. Acusta et al. (2002) proposed an estimator of the probability density function when the data is randomly censored, obtained through an EM algorithm, for solving a maximum likelihood problem. Balakrishnan and Kim (2004) used EM algorithm to find the maximum likelihood estimates under type II right censored samples from a bivariate normal distribution. Park (2005) presented the MLE via EM algorithm to estimate the exponential and lognormal parameters with complex data including: fully-observed, censored, and partially-masked. Cheng and Wang (2012) presented the performance of the maximum likelihood estimates of the Burr XII parameters for CS-PALT by using EM algorithm.

In this paper, I present the performance of the maximum likelihood estimates via EM algorithm for the Burr XII parameters in SS-PALT under multiple censored data in terms of the absolute relative bias, the root mean square error, and the coverage rate. The asymptotic variance and covariance matrix of the estimators are also derived. Then, the confidence intervals of the parameters can be obtained. In addition, an illustrative example is used to demonstrate the proposed method.

II. Model in step-stress PALT under multiple censored data

The probability density function and cumulative distribution function of the two-parameter Burr XII distribution are given by

$$f(t;c,k)=\frac{kct^{c-1}}{\left(1+t^{c}\right)^{k+1}}\qquad,\quad t>0,\ \ c>0,\ \ k>0 \tag{1}$$

$$F(t;c,k)=1-\frac{1}{\left(1+t^{c}\right)^{k}} \quad , \; t>0, \;\; c>0, \;\; k>0 \tag{2}$$

where the parameters *c* and *k* are the shape parameters of the distribution.

In SS-PALT, the test unit is first run at normal condition and if the unit does not fail or be censored before the specified time, τ , the test is switched to a stress condition for testing until the unit fails or be censored. Then, the total lifetime *X* of the unit in SS-PALT is given by

$$X=\begin{cases} T, & T\le\tau \\ \tau+\beta^{-1}\left(T-\tau\right), & T>\tau \end{cases} \tag{3}$$

where *T* is the lifetime of an unit at normal condition, τ is the stress change time and β is the acceleration factor ($\beta>1$). I assume that the lifetime of the test unit follows a two-parameter Burr XII distribution. Therefore, the CDF and PDF of total lifetime *X* of an item are given by

$$F(x;c,k,\beta)=\begin{cases} 0, & x<0 \\ 1-\dfrac{1}{(1+x^{c})^{k}}, & 0<x\le\tau \\ 1-\dfrac{1}{\left\{1+\left[\tau+\beta\left(x-\tau\right)\right]^{c}\right\}^{k}}, & x>\tau \end{cases} \tag{4}$$

where $c > 0$, $k > 0$, $\beta > 1$, and

$$f(x;c,k,\beta)=\begin{cases} 0, & x<0 \\ \dfrac{kcx^{c-1}}{(1+x^{c})^{k+1}}, & 0<x\le\tau \\ \dfrac{\beta kc\left[\tau+\beta\left(x-\tau\right)\right]^{c-1}}{\left\{1+\left[\tau+\beta\left(x-\tau\right)\right]^{c}\right\}^{k+1}}, & x>\tau \end{cases} \tag{5}$$

Suppose that there are n_{1f} failures and n_{1c} units with censoring at normal condition. Also, I assume that there are n_{2f} failures and n_{2c} units with censoring at stress condition. Let $\delta_{i,(1,f)}$, $\delta_{i,(1,c)}$, $\delta_{i,(2,f)}$, $\delta_{i,(2,c)}$ be indicator functions, which (1, *f*) of the indicator function denotes that the sample unit fails before the stress change time, τ , and (1, *c*) denotes that the unit is censored before the time, τ . Also, (2, *f*) denotes that the unit fails after the time, τ . (2, *c*) denotes that the unit is censored after the time, τ . Furthermore, the equations are obtained as follows.

$\sum_{i=1}^{n}\delta_{i,(1,f)}=n_{1f}$, $\sum_{i=1}^{n}\delta_{i,(1,c)}=n_{1c}$, $\sum_{i=1}^{n}\delta_{i,(2,f)}=n_{2f}$, $\sum_{i=1}^{n}\delta_{i,(2,c)}=n_{2c}$, $n_1=n_{1f}+n_{1c}$, and $n_2=n_{2f}+n_{2c}$.

III. Complete-data likelihood function via EM algorithm

Let $\mathbf{y}=\left(\mathbf{y}_1^T,\ldots,\mathbf{y}_n^T\right)^T$ denote the observed data where $\mathbf{y}_i=\left(d_i,\delta_i\right)^T$ and δ_i = 0 (censored) or 1 (failure). As seen in the observations, x_i is censored or uncensored at d_i (i = 1,…,n). Then, the probability density function of the Burr XII distribution, given $x_i > d_i$ is calculated as follows:
Let $a_i=\tau+\beta\left(x_i-\tau\right)$ $A_i=\tau+\beta\left(X_i-\tau\right)$ $D_i=\tau+\beta\left(d_i-\tau\right)$

$$f(x_i\left|x_i > d_i\right.)=\frac{f(x_i)}{1-F(d_i)}=\begin{cases} kc\left(1+d_i^c\right)^k\dfrac{x_i^{c-1}}{\left(1+x_i^c\right)^{k+1}}, & x_i > d_i,\quad d_i<\tau \\ \left(1+D_i^c\right)^k\beta kc\dfrac{a_i^{c-1}}{\left(1+a_i^c\right)^{k+1}}, & x_i > d_i,\quad d_i>\tau \end{cases} \tag{6}$$

the complete data likelihood function of the Burr XII distribution can be expressed as

$$L_c\left(c,k,\beta\right)=\prod_{i=1}^{n}f_c\left(x_i;c,k,\beta\right)=\prod_{i=1}^{n}f(x_i)^{\delta_{i,(1,f)}}f(x_i)^{\delta_{i,(1,c)}}f(x_i)^{\delta_{i,(2,f)}}f(x_i)^{\delta_{i,(2,c)}} \tag{7}$$

the complete data log-likelihood function of the Burr XII distribution is then expressed as

$$\begin{aligned} \log\left[L_c\left(c,k,\beta\right)\right]&=\sum_{i=1}^{n}\log\left[f_c\left(x_i;c,k,\beta\right)\right] \\ &=n\log\left(k\right)+n\log\left(c\right)+n_2\log\left(\beta\right) \\ &\quad+\left(c-1\right)\sum_{i=1}^{n}\delta_{i,(1,f)}\log\left(x_i\right)-\left(k+1\right)\sum_{i=1}^{n}\delta_{i,(1,f)}\log\left(1+x_i^c\right) \\ &\quad+\left(c-1\right)\sum_{i=1}^{n}\delta_{i,(1,c)}\log\left(x_i\right)-\left(k+1\right)\sum_{i=1}^{n}\delta_{i,(1,c)}\log\left(1+x_i^c\right) \\ &\quad+\left(c-1\right)\sum_{i=1}^{n}\delta_{i,(2,f)}\log\left(a_i\right)-\left(k+1\right)\sum_{i=1}^{n}\delta_{i,(2,f)}\log\left(1+a_i^c\right) \\ &\quad+\left(c-1\right)\sum_{i=1}^{n}\delta_{i,(2,c)}\log\left(a_i\right)-\left(k+1\right)\sum_{i=1}^{n}\delta_{i,(2,c)}\log\left(1+a_i^c\right) \end{aligned} \tag{8}$$

then, the Q-function of the Burr XII distribution is obtained as

$$\begin{aligned}
E\left[\log L_c(c,k,\beta)\,|\,\mathbf{y}\right] &= n\log(k)+n\log(c)+n_2\log(\beta) \\
&+(c-1)\sum_{i=1}^{n}\delta_{i,(1,f)}\log(d_i)-(k+1)\sum_{i=1}^{n}\delta_{i,(1,f)}\log\left(1+d_i^c\right) \\
&+(c-1)\sum_{i=1}^{n}\delta_{i,(1,c)}E\left[\log(X_i)\middle|X_i>d_i\right]-(k+1)\sum_{i=1}^{n}\delta_{i,(1,c)}E\left[\log\left(1+X_i^c\right)\middle|X_i>d_i\right] \\
&+(c-1)\sum_{i=1}^{n}\delta_{i,(2,f)}\log(D_i)-(k+1)\sum_{i=1}^{n}\delta_{i,(2,f)}\log\left(1+D_i^c\right) \\
&+(c-1)\sum_{i=1}^{n}\delta_{i,(2,c)}E\left[\log(A_i)\middle|X_i>d_i\right]-(k+1)\sum_{i=1}^{n}\delta_{i,(2,c)}E\left[\log\left(1+A_i^c\right)\middle|X_i>d_i\right]
\end{aligned} \tag{9}$$

For the E-step, $Q\left(\boldsymbol{\psi};\boldsymbol{\psi}_{(m)}\right)$ can be calculated, where $\boldsymbol{\psi}$ denotes the set of parameters, *c*, *k* and *β* and $\boldsymbol{\psi}_{(m)}$ denotes the set of estimates, $c_{(m)}$, $k_{(m)}$ and $\beta_{(m)}$, in *m*-th iteration.

$$\begin{aligned}
Q\left(\boldsymbol{\psi};\boldsymbol{\psi}_{(m)}\right) &= E_{\boldsymbol{\psi}_{(m)}}\left[\log L_c(c,k,\beta)\,|\,\mathbf{y}\right] \\
&= n\log(k)+n\log(c)+n_2\log(\beta) \\
&+(c-1)\sum_{i=1}^{n}\delta_{i,(1,f)}\log(d_i)-(k+1)\sum_{i=1}^{n}\delta_{i,(1,f)}\log\left(1+d_i^c\right) \\
&+(c-1)\sum_{i=1}^{n}\delta_{i,(1,c)}E_{\boldsymbol{\psi}_{(m)}}\left[\log(X_i)\middle|X_i>d_i\right]-(k+1)\sum_{i=1}^{n}\delta_{i,(1,c)}E_{\boldsymbol{\psi}_{(m)}}\left[\log\left(1+X_i^c\right)\middle|X_i>d_i\right] \\
&+(c-1)\sum_{i=1}^{n}\delta_{i,(2,f)}\log(D_i)-(k+1)\sum_{i=1}^{n}\delta_{i,(2,f)}\log\left(1+D_i^c\right) \\
&+(c-1)\sum_{i=1}^{n}\delta_{i,(2,c)}E_{\boldsymbol{\psi}_{(m)}}\left[\log(A_i)\middle|X_i>d_i\right]-(k+1)\sum_{i=1}^{n}\delta_{i,(2,c)}E_{\boldsymbol{\psi}_{(m)}}\left[\log\left(1+A_i^c\right)\middle|X_i>d_i\right]
\end{aligned} \tag{10}$$

For the M-step, $\boldsymbol{\psi}_{(m+1)}$ is the specific value of $\boldsymbol{\psi}\in\Omega$ that maximizes $Q\left(\boldsymbol{\psi};\boldsymbol{\psi}_{(m)}\right)$; that is, $Q\left(\boldsymbol{\psi}_{(m+1)};\boldsymbol{\psi}_{(m)}\right)\geq Q\left(\boldsymbol{\psi};\boldsymbol{\psi}_{(m)}\right)$. The E and M steps repeatedly iterative compute until the estimates of parameters converge to the default value. The above term in equation (10), $E_{\boldsymbol{\psi}(m)}[log(X_i)\,|X_i>d_i]$, can be directly solved by using Monte Carlo method. However the other terms, $E_{\boldsymbol{\psi}_{(m)}}\left[\log\left(1+X_i^c\right)\,|\,X_i>d_i\right]$, $E_{\boldsymbol{\psi}(m)}[log(A_i)\,|X_i>d_i]$ and $E_{\boldsymbol{\psi}(m)}[log(1+A_i^c)\,|X_i>d_i]$ can not be directly solved using Monte Carlo method because the unknown parameter, *c* and *β*, exists within the terms, $log(1+X_i^c)$, $log(A_i)$ and $log(1+A_i^c)$, where $A_i=\tau+\beta(X_i-\tau)$. To decompose these terms, Taylor series expansion can be applied to decompose these terms, $log(1+X_i^c)$, $log(A_i)$ and $log(1+A_i^c)$, and then Monte Carlo method can be applied to compute the integral.

For the Burr XII distribution, the variance-covariance matrix of parameters *c*, *k* and *β* is obtained as

$$\begin{bmatrix} Var(\hat{c}) & Cov(\hat{c},\hat{k}) & Cov(\hat{c},\hat{\beta}) \\ Cov(\hat{c},\hat{k}) & Var(\hat{k}) & Cov(\hat{k},\hat{\beta}) \\ Cov(\hat{c},\hat{\beta}) & Cov(\hat{k},\hat{\beta}) & Var(\hat{\beta}) \end{bmatrix} = (-1)\times \begin{bmatrix} E\left(\frac{\partial^2 \log L}{\partial c^2}\right) & E\left(\frac{\partial^2 \log L}{\partial c \partial k}\right) & E\left(\frac{\partial^2 \log L}{\partial c \partial \beta}\right) \\ E\left(\frac{\partial^2 \log L}{\partial c \partial k}\right) & E\left(\frac{\partial^2 \log L}{\partial k^2}\right) & E\left(\frac{\partial^2 \log L}{\partial k \partial \beta}\right) \\ E\left(\frac{\partial^2 \log L}{\partial c \partial \beta}\right) & E\left(\frac{\partial^2 \log L}{\partial k \partial \beta}\right) & E\left(\frac{\partial^2 \log L}{\partial \beta^2}\right) \end{bmatrix}^{-1} \tag{11}$$

where E symbolizes expectation and L denotes log-likelihood function. The observed information (I_{obs}) can be used to construct the variance-covariance matrix and confidence intervals for c, k and β. Complete (I_{comp}) and missing (I_{miss}) information can be used to calculate the rate of convergence of EM algorithm. Louis (1982) showed that the observed information presents the difference between complete information and missing information within the framework of EM algorithm. The equation is expressed as $I_{obs}\, I_{obs} = I_{comp}$ -I_{miss}. I_{comp} and I_{miss} are obtained in Appendix. Therefore, the variance-covariance matrix of parameters c, k and β can be obtained by inverting the observed information matrix and is given by

$$\begin{bmatrix} Var(\hat{c}) & Cov(\hat{c},\hat{k}) & Cov(\hat{c},\hat{\beta}) \\ Cov(\hat{c},\hat{k}) & Var(\hat{k}) & Cov(\hat{k},\hat{\beta}) \\ Cov(\hat{c},\hat{\beta}) & Cov(\hat{k},\hat{\beta}) & Var(\hat{\beta}) \end{bmatrix} = [I_{comp}(c,k,\beta;\mathbf{y}) - I_{miss}(c,k,\beta;\mathbf{y})]^{-1} \tag{12}$$

Thus, an approximate (1-α)100% confidence intervals for c, k and β are obtained as

$$\hat{c} \pm z_{\frac{\alpha}{2}}\sqrt{var(\hat{c})}\ ,\hat{k} \pm z_{\frac{\alpha}{2}}\sqrt{var(\hat{k})} \text{ and } \hat{\beta} \pm z_{\frac{\alpha}{2}}\sqrt{var(\hat{\beta})} \tag{13}$$

where $z_{\frac{\alpha}{2}}$ is a standard normal variate.

IV. Observed-data likelihood function via BFGS algorithm

The MLE based on observed-data likelihood function of the Burr XII distribution with multiple censored data in a SS-PALT is given by

$$L = \prod_{i=1}^{n} f\left(x_{i;1,f}\right)\left[1-F\left(x_{i;1,c}\right)\right] f\left(x_{i;2,f}\right)\left[1-F\left(x_{i;2,c}\right)\right] \tag{14}$$

The log-likelihood function is obtained as

$$\begin{aligned} \log L = {} & n_{1f}\log(c) + n_{1f}\log(k) + (c-1)\sum_{i=1}^{n}\log\left(x_{i;1,f}\right) - (k+1)\sum_{i=1}^{n}\log\left(1+x_{i;1,f}^{c}\right) - k\sum_{i=1}^{n}\log\left(1+x_{i;1,c}^{c}\right) \\ & + n_{2f}\log(\beta) + n_{2f}\log(c) + n_{2f}\log(k) + (c-1)\sum_{i=1}^{n}\log\left(a_{i;2,f}\right) - (k+1)\sum_{i=1}^{n}\log\left(1+a_{i;2,c}^{c}\right) - k\sum_{i=1}^{n}\log\left(1+a_{i;2,c}^{c}\right), \end{aligned} \tag{15}$$

where $a_{i;2,f} = \tau + \beta\left(x_{i;2,f} - \tau\right)$ and $a_{i;2,c} = \tau + \beta\left(x_{i;2,c} - \tau\right)$.

The estimates of c, k, and β are obtained by setting the first partial derivatives of the log-likelihood to zero with respect to *c*, *k* , and *β*, respectively. The simultaneous equations are given as follows:

$$\partial \log L / \partial c = n_{1f} c^{-1} + \sum_{i=1}^{n} \log\left(x_{i;1,f}\right) - (k+1) \sum_{i=1}^{n} \log\left(x_{i;1,f}\right) x_{i;1,f}^{c} \left(1 + x_{i;1,f}^{c}\right)^{-1}$$
$$- k \sum_{i=1}^{n} \log\left(x_{i;1,c}\right) x_{i;1,c}^{c} \left(1 + x_{i;1,c}^{c}\right)^{-1} + n_{2f} c^{-1} + \sum_{i=1}^{n} \log\left(a_{i;2,f}\right) \tag{16}$$
$$- (k+1) \sum_{i=1}^{n} \log\left(a_{i;2,f}\right) a_{i;2,f}^{c} \left(1 + a_{i;2,f}^{c}\right)^{-1} - k \sum_{i=1}^{n} \log\left(a_{i;2,c}\right) a_{i;2,c}^{c} \left(1 + a_{i;2,c}^{c}\right)^{-1} = 0,$$

$$\partial \log L / \partial k = n_{1f} k^{-1} - \sum_{i=1}^{n} \log\left(1 + x_{i;1,f}^{c}\right) - \sum_{i=1}^{n} \log\left(1 + x_{i;1,c}^{c}\right)$$
$$+ n_{2f} k^{-1} - \sum_{i=1}^{n} \log\left(1 + a_{i;2,f}^{c}\right) - \sum_{i=1}^{n} \log\left(1 + a_{i;2,c}^{c}\right) = 0, \tag{17}$$

$$\partial \log L / \partial \beta = n_{2f} \beta^{-1} + (c-1) \sum_{i=1}^{n} \left(x_{i;2,f} - \tau\right) a_{i;2,f}^{-1} - (k+1) \sum_{i=1}^{n} c a_{i}^{c-1} \left(x_{i;2,f} - \tau\right) \left(1 + a_{i;2,f}^{c}\right)^{-1}$$
$$- k \sum_{i=1}^{n} c a_{i;2,c}^{c-1} \left(x_{i;2,c} - \tau\right) \left(1 + a_{i;2,c}^{c}\right)^{-1} = 0. \tag{18}$$

BFGS algorithm is then applied for solving these simultaneous equations to obtain the estimated values of c , k , and *β*. The initial estimates of the parameters are chosen using pseudo complete estimates which the samples are completely treated as failures. The asymptotic variance-covariance matrix of *c*, *k* , and *β* is established as

$$Var\left(\hat{\boldsymbol{\psi}}\right) = I_{obs}^{-1}\left(\boldsymbol{\psi}; \mathbf{x}\right) = \left[-\partial^{2} \log L\left(\boldsymbol{\psi}\right) / \partial \boldsymbol{\psi} \partial \boldsymbol{\psi}^{T}\right]^{-1}, \tag{19}$$

where $\boldsymbol{\psi}$ denotes the set of *c*, *k*, and *β*. Thus, the approximate (1-α)100% confidence intervals for *c*, *k*, and *β* are obtained as

$$\hat{c} \pm z_{\alpha/2} \sqrt{var(\hat{c})}\,,\ \hat{k} \pm z_{\alpha/2} \sqrt{var(\hat{k})} \text{ and } \hat{\beta} \pm z_{\alpha/2} \sqrt{var(\hat{\beta})}\,, \tag{20}$$

where $z_{\alpha/2}$ is the $100(1-\alpha/2)$ percentile of the standard normal distribution.

V. Simulation study

The method in Wang, Cheng and Lu (2012) was used for generating multiple censored samples. Censored samples were randomly generated from the Burr XII distribution with specified values of *c*, *k* and *β*. The simulation included the following conditions: sample sizes *n* = 100, 200; the stress change time, *τ* = 0.5, 1.5; censoring level *CL* = 0.2. Here we considered (*c*, *k*, *β*) = (1, 0.5, 1.25), (1, 0.5, 2), (1, 1, 1.25), (1, 1, 2), (2, 0.5, 1.25), (2, 0.5, 2),(2, 1, 1.25), (2, 1, 2), (2, 2, 1.25) and (2, 2, 2) as true parameter values. For each data set, 1000 replications are simulated. To assess the performance of the MLE via EM algorithm, I consider three major measures including the absolute relative bias (ARB), the root mean squared error (RMSE) , and the coverage rate (CR).

They are defined as follows:

1) $\text{ARB}(\hat{c}) = N^{-1}\sum_{i=1}^{N}\left|\left(\hat{c}_i - c\right)/c\right|$, $\text{ARB}(\hat{k}) = N^{-1}\sum_{i=1}^{N}\left|\left(\hat{k}_i - k\right)/k\right|$ and $\text{ARB}(\hat{\beta}) = N^{-1}\sum_{i=1}^{N}\left|\left(\hat{\beta}_i - \beta\right)/\beta\right|$,
2) $RMSE(\hat{c}) = N^{-1}\sum_{i=1}^{N}(\hat{c}_i - c)^2$, $RMSE(\hat{k}) = N^{-1}\sum_{i=1}^{N}(\hat{k}_i - k)^2$ and $RMSE(\hat{\beta}) = N^{-1}\sum_{i=1}^{N}(\hat{\beta}_i - \beta)^2$,
3) The coverage rate at the 95% confidence intervals for *c*, *k* and β is based on *N* simulations, where $\overline{c} = N^{-1}\sum_{i=1}^{N}\hat{c}_i$, $\overline{k} = N^{-1}\sum_{i=1}^{N}\hat{k}_i$, $\overline{\beta} = N^{-1}\sum_{i=1}^{N}\hat{\beta}_i$, and *N* = 1,000.

The simulation results for the multiple censored with *CL*=0.2 for sample sizes 100 and 200 are presented in Tables 1-2. The following conclusions were observed.

1) For the sample size of 100 in Table 1, EM algorithm provides lower levels of ARB and RMSE for parameters *c*, *k*, and β than BFGS algorithm does in most scenarios. EM algorithm estimates perform better than BFGS algorithm does, the proportion accounting for 68.3% (41 cases/60 cases) for ARB and 71.7% (43 cases/60 cases) for RMSE. This indicates that EM algorithm performs better than BFGS algorithm does in this simulation study.
2) For the sample size of 100 in Table 1, the 95% C.I. is calculated for parameters *c*, *k*, and β. In most scenarios, EM algorithm provides higher levels of CR for parameters *c*, *k*, and β than BFGS algorithm does. EM algorithm estimates perform better than BFGS algorithm does, the proportion accounting for 100% (60 cases/60 cases). The average values of CR are 95.6% for EM algorithm and 72.0% for BFGS algorithm. This indicates that EM algorithm performs better than BFGS algorithm does in this simulation study.
3) For the sample size of 200 in Table 2, the results are similar with those for the sample size of 100. EM algorithm estimates perform better than BFGS algorithm does, the proportion accounting for 58.3% (35 cases/60 cases) for ARB and 65.0% (39 cases/60 cases) for RMSE. EM algorithm estimates perform better than BFGS algorithm does, the proportion accounting for 73.3% (44 cases/60 cases) for CR. The average values of CR are 93.9% for EM algorithm and 88.6% for BFGS algorithm.
4) With the sample size of complete data increasing from 100 to 200, EM algorithm and BFGS algorithm estimates for parameters *c*, *k*, and β are more accurate and have fewer errors, and lower ARB and RMSE.

Table 1: *ARB, RMSE and CR of the estimates with n = 100.*

k	*c*	β	τ	*Parameters*	BFGS algorithm			EM algorithm		
					ARB	RMSE	CR (%)	ARB	RMSE	CR (%)
				k	0.1480	0.1728	62.4	0.1461	0.1677	90.7
1	0.5	1.25	0.5	*c*	0.1084	0.0692	89.2	0.1010	0.0652	99.7
				β	0.3148	0.4633	44.9	0.2644	0.4025	87.7
				k	0.1479	0.1717	63.0	0.1347	0.1601	91.4
1	0.5	2	0.5	*c*	0.1036	0.0661	90.6	0.1015	0.0652	99.2
				β	0.2947	0.6911	49.2	0.2544	0.6264	89.2
				k	0.1288	0.1505	74.4	0.1265	0.1487	97.4
1	1	1.25	0.5	*c*	0.1046	0.1384	90.2	0.1021	0.1349	99.5
				β	0.2272	0.3456	64.7	0.2045	0.3117	95.8
				k	0.1296	0.1543	72.1	0.1150	0.1407	98.3
1	1	2	0.5	*c*	0.0981	0.1280	90.9	0.1028	0.1315	99.7
				β	0.2235	0.5401	67.1	0.2299	0.5481	94.2
				k	0.1431	0.3294	64.2	0.1505	0.3404	92.0
2	0.5	1.25	0.5	*c*	0.0871	0.0558	93.4	0.0831	0.0523	99.6
				β	0.2383	0.3631	59.0	0.2080	0.3209	96.3
				k	0.1369	0.3176	67.4	0.1217	0.2849	95.3

2	0.5	2	0.5	*c*	0.0856	0.0549	93.7	0.0903	0.0578	99.3
				β	0.2435	0.5978	57.7	0.2433	0.5905	94.5
				k	0.1423	0.3341	65.8	0.1340	0.3145	96.9
2	1	1.25	0.5	*c*	0.0921	0.1204	90.6	0.0879	0.1130	99.8
				β	0.2343	0.3538	64.9	0.2020	0.3101	98.3
				k	0.1425	0.3354	64.6	0.1128	0.2747	98.1
2	1	2	0.5	*c*	0.0963	0.1259	88.0	0.1069	0.1385	99.2
				β	0.2459	0.5962	64.0	0.2257	0.5507	94.6
				k	0.1827	0.4300	55.7	0.1325	0.3338	98.2
2	2	1.25	0.5	*c*	0.0849	0.2105	92.2	0.1016	0.2526	99.8
				β	0.2377	0.4180	79.7	0.1904	0.3279	99.4
				k	0.1807	0.4256	54.8	0.1610	0.4157	99.6
2	2	2	0.5	*c*	0.0805	0.2029	94.2	0.1295	0.3319	99.6
				β	0.2329	0.6224	77.9	0.2243	0.5268	95.3
				k	0.1605	0.0935	60.0	0.1508	0.0891	91.4
0.5	1	1.25	1.5	*c*	0.1313	0.1734	82.5	0.1339	0.1712	98.1
				β	0.2792	0.4255	55.4	0.2310	0.3533	91.5
				k	0.1573	0.0927	62.9	0.1461	0.0874	91.9
0.5	1	2	1.5	*c*	0.1227	0.1635	85.9	0.1234	0.1628	99.0
				β	0.2592	0.6214	59.5	0.2373	0.5795	91.2
				k	0.1226	0.0731	77.3	0.1672	0.0981	90.4
0.5	2	1.25	1.5	*c*	0.0987	0.2407	89.2	0.1331	0.3307	99.5
				β	0.2117	0.3364	74.0	0.2007	0.3091	97.9
				k	0.1217	0.0723	78.3	0.1489	0.0896	92.6
0.5	2	2	1.5	*c*	0.0948	0.2315	90.1	0.1333	0.3554	99.5
				β	0.2045	0.5196	75.5	0.2186	0.5354	95.1
				k	0.1592	0.1822	56.1	0.1542	0.1767	87.5
1	0.5	1.25	1.5	*c*	0.1066	0.0682	89.9	0.1047	0.0674	98.9
				β	0.3348	0.5185	42.8	0.2703	0.4126	88.3
				k	0.1481	0.1713	62.5	0.1411	0.1641	90.3
1	0.5	2	1.5	*c*	0.1019	0.0650	91.4	0.1069	0.0675	99.1
				β	0.3152	0.7462	43.4	0.2740	0.6668	88.9
				k	0.1511	0.1756	58.7	0.1517	0.1774	89.0
1	1	1.25	1.5	*c*	0.0927	0.1221	92.7	0.0927	0.1242	99.0
				β	0.2241	0.3389	65.0	0.2157	0.3251	96.6
				k	0.1499	0.1738	61.3	0.1370	0.1601	92.4
1	1	2	1.5	*c*	0.0971	0.1256	91.6	0.1046	0.1330	99.5
				β	0.2300	0.5519	61.8	0.2238	0.5417	94.7
				k	0.1721	0.1883	52.9	0.1706	0.1876	89.0
1	2	1.25	1.5	*c*	0.0879	0.2233	91.4	0.0852	0.2125	99.1
				β	0.2461	0.4011	69.6	0.2206	0.3670	99.1
				k	0.1700	0.1872	52.4	0.1534	0.1738	89.8
1	2	2	1.5	*c*	0.0882	0.2284	92.9	0.0894	0.2292	99.1
				β	0.2402	0.6245	66.4	0.2250	0.6059	98.5

Table 2: *ARB, RMSE and CR of the estimates with n = 200.*

k	*c*	*β*	*τ*	*Parameters*	BFGS algorithm			EM algorithm		
					ARB	RMSE	CR (%)	ARB	RMSE	CR (%)
				k	0.1508	0.1660	78.2	0.1473	0.1616	81.2
1	0.5	1.25	0.5	*c*	0.0711	0.0455	99.2	0.0743	0.0478	98.8
				β	0.2466	0.3757	85.2	0.1990	0.3043	93.0
				k	0.1484	0.1650	76.4	0.1334	0.1496	85.9
1	0.5	2	0.5	*c*	0.0693	0.0444	99.0	0.0760	0.0479	99.2
				β	0.2342	0.5641	85.8	0.2042	0.5030	92.1
				k	0.1408	0.1581	100.0	0.1352	0.1515	89.0
1	1	1.25	0.5	*c*	0.0753	0.0971	98.6	0.0759	0.0974	99.1
				β	0.1647	0.2510	97.9	0.1561	0.2368	98.2
				k	0.1369	0.1563	99.2	0.1206	0.1410	91.7
1	1	2	0.5	*c*	0.0716	0.0922	99.0	0.0799	0.1033	99.1
				β	0.1650	0.4029	98.6	0.1721	0.4149	95.9
				k	0.1426	0.3155	66.7	0.1351	0.2970	81.8

2	0.5	1.25	0.5	*c*	0.0630	0.0396	97.9	0.0612	0.0379	99.0
				β	0.1960	0.2975	74.1	0.1781	0.2736	97.2
				k	0.1441	0.3176	63.5	0.1220	0.2738	88.4
2	0.5	2	0.5	*c*	0.0612	0.0394	97.9	0.0707	0.0445	98.8
				β	0.1968	0.4826	74.4	0.2002	0.4909	94.5
				k	0.1448	0.3255	70.4	0.1244	0.2817	91.8
2	1	1.25	0.5	*c*	0.0640	0.0830	97.2	0.0641	0.0817	99.4
				β	0.1880	0.2884	89.4	0.1548	0.2302	99.3
				k	0.1445	0.3240	72.2	0.1029	0.2425	97.4
2	1	2	0.5	*c*	0.0650	0.0827	97.8	0.0761	0.0958	98.3
				β	0.1819	0.4532	91.1	0.1733	0.4092	96.5
				k	0.1556	0.3650	97.0	0.1130	0.2798	97.6
2	2	1.25	0.5	*c*	0.0637	0.1588	99.9	0.0790	0.1994	99.2
				β	0.1594	0.2723	99.4	0.1446	0.2255	99.3
				k	0.1534	0.3619	96.9	0.1172	0.2882	99.3
2	2	2	0.5	*c*	0.0651	0.1621	99.7	0.1089	0.2649	98.6
				β	0.1632	0.4386	99.6	0.1880	0.4362	96.8
				k	0.1656	0.0921	84.4	0.1525	0.0857	82.7
0.5	1	1.25	1.5	*c*	0.0931	0.1204	98.9	0.0908	0.1147	98.5
				β	0.2053	0.3156	94.9	0.1715	0.2636	94.0
				k	0.1599	0.0896	82.9	0.1475	0.0842	85.5
0.5	1	2	1.5	*c*	0.0845	0.1135	98.6	0.0966	0.1262	99.3
				β	0.2012	0.4938	94.0	0.1805	0.4461	92.5
				k	0.1351	0.0761	90.2	0.1570	0.0884	83.5
0.5	2	1.25	1.5	*c*	0.0704	0.1716	99.2	0.0823	0.2116	99.6
				β	0.1619	0.2558	95.6	0.1430	0.2146	98.2
				k	0.1337	0.0752	92.4	0.1446	0.0827	87.1
0.5	2	2	1.5	*c*	0.0698	0.1717	99.2	0.0892	0.2341	99.2
				β	0.1427	0.3658	94.2	0.1450	0.3584	96.6
				k	0.1569	0.1722	66.2	0.1466	0.1617	75.5
1	0.5	1.25	1.5	*c*	0.0695	0.0449	98.5	0.0719	0.0448	99.5
				β	0.2454	0.3697	71.9	0.1939	0.3028	95.6
				k	0.1616	0.1754	61.3	0.1431	0.1575	83.0
1	0.5	2	1.5	*c*	0.0686	0.0438	99.0	0.0752	0.0475	99.0
				β	0.2301	0.5617	74.0	0.1996	0.4927	92.8
				k	0.1542	0.1683	67.5	0.1509	0.1663	75.6
1	1	1.25	1.5	*c*	0.0677	0.0871	98.6	0.0688	0.0875	99.1
				β	0.1779	0.2689	83.7	0.1609	0.2451	98.2
				k	0.1545	0.1697	66.5	0.1367	0.1544	80.2
1	1	2	1.5	*c*	0.0666	0.0842	99.1	0.0752	0.0935	99.6
				β	0.1714	0.4153	85.4	0.1796	0.4242	97.3
				k	0.1436	0.1596	73.3	0.1347	0.1509	80.4
1	2	1.25	1.5	*c*	0.0603	0.1521	99.6	0.0616	0.1566	99.8
				β	0.1864	0.2935	85.0	0.1612	0.2483	98.4
				k	0.1395	0.1547	75.4	0.1225	0.1384	88.5
1	2	2	1.5	*c*	0.0641	0.1601	98.6	0.0706	0.1787	99.5
				β	0.1742	0.4363	86.2	0.1722	0.4117	97.5

VI. Illustrative example

To illustrate the proposed MLEs via EM algorithm for the Burr XII distribution in SS-PALT, one data set from a light-emitting diode (LED) life test was used. The life test data with 1,000 hours of unit are as follows:

0.02*, 0.03*, 0.08*, 0.11*, 0.13*, 0.14*, 0.15*, 0.19*, 0.21*, 0.25*, 0.25*, 0.27, 0.28*, 0.31, 0.33, 0.35, 0.37*, 0.42, 0.43*, 0.44*, 0.46, 0.46, 0.49, 0.51, 0.51, 0.55*, 0.56, 0.58, 0.58*, 0.59, 0.59*, 0.6, 0.71, 0.71*, 0.73, 0.73, 0.73, 0.78, 0.79*, 0.81, 0.84, 0.87, 0.89, 0.9, 0.92, 0.92, 0.95, 1.01, 1.02, 1.06, 1.07, 1.08, 1.24, 1.24*, 1.25, 1.26, 1.31, 1.5*, 1.51*, 1.52*, 1.53*, 1.54, 1.55*, 1.56, 1.57*, 1.64, 1.64*, 1.65*, 1.67, 1.69, 1.7*, 1.83, 1.91, 2.03, 2.1*, 2.36, 2.78, 4.67

There are 78 samples with stress change time, τ = 1.5 and censoring level CL = 0.4. The samples of failure and censoring in the two phases of SS-PALT, respectively, are 36 failures in phase 1, 21 censoring in phase 1, 11 failures in phase 2 and 10 censoring in phase 2. The symbol "*" denotes multiple censored values. The histogram of the samples is illustrated in Figure 1 and the plot of the probability density function is illustrated in Figure 2. The initial estimates for the parameters were chosen by using pseudo complete estimates. Here, the pseudo complete estimates are computed from the samples which are completely treated as failures. Using the MLE with EM algorithm, the estimates are converged to 2.538 for c, 0.776 for k and 1.795 for β. The information matrices based on EM algorithm are obtained as

$$I_{comp} = \begin{bmatrix} 17.1297 & 25.1871 & 3.9065 \\ 25.1871 & 129.4696 & 10.3786 \\ 3.9065 & 10.3786 & 3.5688 \end{bmatrix}$$

$$I_{miss} = \begin{bmatrix} 6.0889 & 14.6162 & 1.6656 \\ 14.6162 & 51.4559 & 4.4966 \\ 1.6656 & 4.4966 & 1.4097 \end{bmatrix}$$

$$I_{obs} = I_{comp} - I_{miss} = \begin{bmatrix} 11.0408 & 10.5709 & 2.2409 \\ 10.5709 & 78.0137 & 5.8820 \\ 2.2409 & 5.8820 & 2.1591 \end{bmatrix}$$

Then, the asymptotic variance-covariance matrix based on EM algorithm can be obtained as

$$I_{obs}^{-1} = \begin{bmatrix} 0.1191 & -0.0086 & -0.1003 \\ -0.0086 & 0.0168 & -0.0367 \\ -0.1003 & -0.0367 & 0.6673 \end{bmatrix}$$

Then, the 95% confidence intervals, (1.862, 3.214) for c, (0.521, 1.031) for k and (0.194, 3.396) for β are obtained. The rates of convergence of c, k and β computed by $J(\hat{\psi}) = I_{miss}(\hat{\psi}) / I_{comp}(\hat{\psi})$ are 0.355 for c, 0.397 for k and 0.395 for β, respectively.

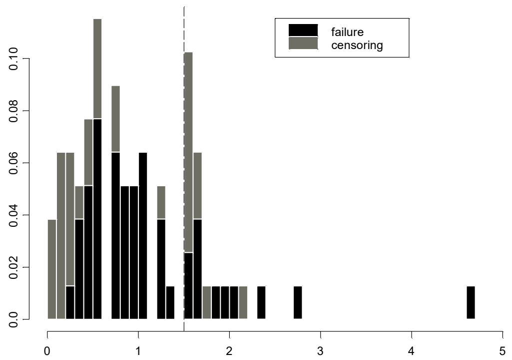

Figure 1: *Histogram of the samples*

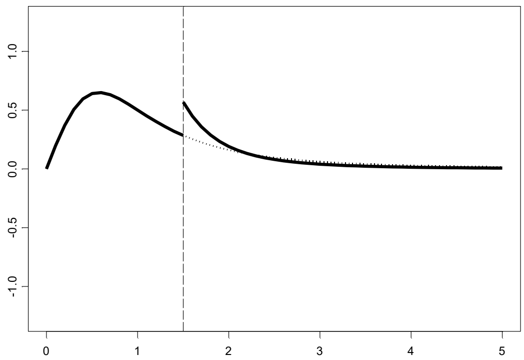

Figure 2: *Probability density plot*

VII. Conclusion

The lifetime of products under normal conditions usually requires a long period of time, which makes the test costly. Accelerated life test is used to obtain information about the lifetime of products quickly and economically under more severe operation conditions. In this paper, I present maximum likelihood estimation via EM algorithm to estimate the Burr XII parameters and acceleration factor in SS-PALT under multiple censored data. Simulation results show that the MLE via EM algorithm perform well in most cases in terms of the absolute relative bias, the root mean square, and the coverage rate. The simulation results and a real data analysis show the MLE via EM algorithm is a better alternative for estimating the Burr XII parameter in SS-PALT with multiple censored data.

Appendix:

The second partials of the complete data log-likelihood function for calculating elements of the complete information matrix are calculated. Then, the expected values of the second partials of the complete data log-likelihood function are obtained as

$$E_{\Psi}\left[\frac{\partial^2 \log L_c}{\partial c^2}|\mathbf{y}\right]$$

$$=\frac{-n}{c^2}-(k+1)\sum_{i=1}^{n}\delta_{i,(1,f)}\frac{d_i^c\log(d_i)^2}{\left(1+d_i^c\right)^2}-(k+1)\sum_{i=1}^{n}\delta_{i,(1,c)}E_{\Psi}\left[\frac{X_i^c\log(X_i)^2}{\left(1+X_i^c\right)^2}\mid X_i>d_i\right]$$

$$-(k+1)\sum_{i=1}^{n}\delta_{i,(2,f)}\frac{D_i^c\log(D_i)^2}{\left(1+D_i^c\right)^2}-(k+1)\sum_{i=1}^{n}\delta_{i,(2,c)}E_{\Psi}\left[\frac{A_i^c\log(A_i)^2}{\left(1+A_i^c\right)^2}\mid X_i>d_i\right]$$

$$E_{\Psi}\left[\frac{\partial^2 \log L_c}{\partial k^2}|\mathbf{y}\right]=-\frac{n}{k^2}$$

$$E_{\Psi}\left[\frac{\partial^2 \log L_c}{\partial \beta^2}|\mathbf{y}\right]=\frac{-n_2}{\beta^2}-(c-1)\sum_{i=1}^{n}\delta_{i,(2,f)}\frac{(d_i-\tau)^2}{D_i^2}$$

$$-(k+1)c(c-1)\sum_{i=1}^{n}\delta_{i,(2,f)}\frac{(d_i-\tau)^2D_i^{c-2}}{1+D_i^c}+(k+1)c^2\sum_{i=1}^{n}\delta_{i,(2,f)}\frac{(d_i-\tau)^2D_i^{2c-2}}{\left(1+D_i^c\right)^2}$$

$$-(c-1)\sum_{i=1}^{n}\delta_{i,(2,c)}E_{\Psi}\left[\frac{(X_i-\tau)^2}{A_i^2}|X_i>d_i\right]$$

$$-(k+1)c(c-1)\sum_{i=1}^{n}\delta_{i,(2,c)}E_{\Psi}\left[\frac{(X_i-\tau)^2A_i^{c-2}}{1+A_i^c}|X_i>d_i\right]$$

$$+(k+1)c^2\sum_{i=1}^{n}\delta_{i,(2,c)}E_{\Psi}\left[\frac{(X_i-\tau)^2A_i^{2c-2}}{\left(1+A_i^c\right)^2}|X_i>d_i\right]$$

$$E_{\psi}\left[\frac{\partial^2 \log L_c}{\partial c \partial k}|\mathbf{y}\right] = -\sum_{i=1}^{n}\delta_{i,(1,f)}\frac{d_i^c \log(d_i)}{1+d_i^c} - \sum_{i=1}^{n}\delta_{i,(1,c)}E_{\psi}\left[\frac{X_i^c \log(X_i)}{1+X_i^c} \mid X_i > d_i\right]$$
$$-\sum_{i=1}^{n}\delta_{i,(2,f)}\frac{D_i^c \log(D_i)}{1+D_i^c} - \sum_{i=1}^{n}\delta_{i,(2,c)}E_{\psi}\left[\frac{A_i^c \log(A_i)}{1+A_i^c} \mid X_i > d_i\right]$$

$$E_{\psi}\left[\frac{\partial^2 \log L_c}{\partial c \partial \beta}|\mathbf{y}\right] = \sum_{i=1}^{n}\delta_{i,(2,f)}\frac{d_i-\tau}{D_i} - (k+1)c\sum_{i=1}^{n}\delta_{i,(2,f)}\frac{(d_i-\tau)D_i^{c-1}\log(D_i)}{1+D_i^c}$$
$$-(k+1)\sum_{i=1}^{n}\delta_{i,(2,f)}\frac{(d_i-\tau)D_i^{c-1}}{1+D_i^c} + (k+1)c\sum_{i=1}^{n}\delta_{i,(2,f)}\frac{(d_i-\tau)D_i^{2c-1}\log(D_i)}{\left(1+D_i^c\right)^2}$$
$$+\sum_{i=1}^{n}\delta_{i,(2,c)}E_{\psi}\left[\frac{X_i-\tau}{A_i} \mid X_i > d_i\right] - (k+1)c\sum_{i=1}^{n}\delta_{i,(2,c)}E_{\psi}\left[\frac{(X_i-\tau)A_i^{c-1}\log(A_i)}{1+A_i^c} \mid X_i > d_i\right]$$
$$-(k+1)\sum_{i=1}^{n}\delta_{i,(2,c)}E_{\psi}\left[\frac{(X_i-\tau)A_i^{c-1}}{1+A_i^c} \mid X_i > d_i\right]$$
$$+(k+1)c\sum_{i=1}^{n}\delta_{i,(2,c)}E_{\psi}\left[\frac{(X_i-\tau)A_i^{2c-1}\log(A_i)}{\left(1+A_i^c\right)^2} \mid X_i > d_i\right]$$

$$E_{\psi}\left[\frac{\partial^2 \log L_c}{\partial k \partial \beta}|\mathbf{y}\right] = -c\sum_{i=1}^{n}\delta_{i,(2,f)}\frac{(d_i-\tau)D_i^{c-1}}{1+D_i^c} - c\sum_{i=1}^{n}\delta_{i,(2,c)}E_{\psi}\left[\frac{(X_i-\tau)A_i^{c-1}}{1+A_i^c} \mid X_i > d_i\right]$$

The expected values of the second partials of the complete data log-likelihood function can also be computed by using Monte Carlo integral. Then, the complete information becomes

$$I_{comp}(\boldsymbol{\psi};\mathbf{y}) = E_{\psi}\left\{I_{comp}(\boldsymbol{\psi};\mathbf{x})|\mathbf{y}\right\}$$
$$=(-1)\times\begin{bmatrix} E_{\psi}\left\{\frac{\partial^2 \log L_c}{\partial c^2}|\mathbf{y}\right\} & E_{\psi}\left\{\frac{\partial^2 \log L_c}{\partial ck}|\mathbf{y}\right\} & E_{\psi}\left\{\frac{\partial^2 \log L_c}{\partial c\beta}|\mathbf{y}\right\} \\ E_{\psi}\left\{\frac{\partial^2 \log L_c}{\partial ck}|\mathbf{y}\right\} & E_{\psi}\left\{\frac{\partial^2 \log L_c}{\partial k^2}|\mathbf{y}\right\} & E_{\psi}\left\{\frac{\partial^2 \log L_c}{\partial k\beta}|\mathbf{y}\right\} \\ E_{\psi}\left\{\frac{\partial^2 \log L_c}{\partial c\beta}|\mathbf{y}\right\} & E_{\psi}\left\{\frac{\partial^2 \log L_c}{\partial k\beta}|\mathbf{y}\right\} & E_{\psi}\left\{\frac{\partial^2 \log L_c}{\partial \beta^2}|\mathbf{y}\right\} \end{bmatrix}$$

Now, the missing information matrix by using the likelihood function of X given Y can be derived and is given as follows

$$k\left(\mathbf{x}|\mathbf{y};\boldsymbol{\psi}\right) = \prod_{i=1}^{n} f(x_i|x_i > d_i)^{\delta_{i,(1,c)}} f(x_i|x_i > d_i)^{\delta_{i,(2,c)}}$$

Then, the log-likelihood function of X given Y is expressed as

$$\log k\left(\mathbf{x}|\mathbf{y};\boldsymbol{\psi}\right)$$
$$=\sum_{i=1}^{n}\delta_{i,(1,c)}\left[\log(k)+\log(c)+k\log\left(1+d_i^c\right)+(c-1)\log(x_i)-(k+1)\log\left(1+x_i^c\right)\right]$$
$$+\sum_{i=1}^{n}\delta_{i,(2,c)}\left[\log(\beta)+\log(k)+\log(c)+k\log\left(1+D_i^c\right)+(c-1)\log(a_i)-(k+1)\log\left(1+a_i^c\right)\right]$$

The second partials of the log-likelihood functions for calculating elements of missing information matrix can be calculated. The expected values of the second partials of the log-likelihood function of X given Y are calculated as

$$E_{\boldsymbol{\psi}}\left[\frac{\partial^2\log k\left(\mathbf{x}|\mathbf{y};\boldsymbol{\psi}\right)}{\partial c^2}|\mathbf{y}\right]$$
$$=-\frac{n_{1c}}{c^2}+k\sum_{i=1}^{n}\delta_{i,(1,c)}\frac{d_i^c\log(d_i)^2}{\left(1+d_i^c\right)^2}-(k+1)\sum_{i=1}^{n}\delta_{i,(1,c)}E_{\boldsymbol{\psi}}\left[\frac{X_i^c\log(X_i)^2}{\left(1+X_i^c\right)^2}|X_i>d_i\right]$$
$$-\frac{n_{2c}}{c^2}+k\sum_{i=1}^{n}\delta_{i,(2,c)}\frac{D_i^c\log(D_i)^2}{\left(1+D_i^c\right)^2}-(k+1)\sum_{i=1}^{n}\delta_{i,(2,c)}E_{\boldsymbol{\psi}}\left[\frac{A_i^c\log(A_i)^2}{\left(1+A_i^c\right)^2}|X_i>d_i\right]$$
$$E_{\boldsymbol{\psi}}\left[\frac{\partial^2\log k\left(\mathbf{x}|\mathbf{y};\boldsymbol{\psi}\right)}{\partial k^2}|\mathbf{y}\right]=-\frac{n_{1c}+n_{2c}}{k^2}$$
$$E_{\boldsymbol{\psi}}\left[\frac{\partial^2\log k\left(\mathbf{x}|\mathbf{y};\boldsymbol{\psi}\right)}{\partial \beta^2}|\mathbf{y}\right]$$
$$=\frac{-n_{2c}}{\beta^2}+kc(c-1)\sum_{i=1}^{n}\delta_{i,(2,c)}\frac{(d_i-\tau)^2D_i^{c-2}}{\left(1+D_i^c\right)}-kc^2\sum_{i=1}^{n}\delta_{i,(2,c)}\frac{(d_i-\tau)^2D_i^{2c-2}}{\left(1+D_i^c\right)^2}$$
$$-(c-1)\sum_{i=1}^{n}\delta_{i,(2,c)}E_{\boldsymbol{\psi}}\left[\frac{(X_i-\tau)^2}{A_i^2}|X_i>d_i\right]$$
$$-(k+1)c(c-1)\sum_{i=1}^{n}\delta_{i,(2,c)}E_{\boldsymbol{\psi}}\left[\frac{(X_i-\tau)^2A_i^{c-2}}{1+A_i^c}|X_i>d_i\right]$$
$$+(k+1)c^2\sum_{i=1}^{n}\delta_{i,(2,c)}E_{\boldsymbol{\psi}}\left[\frac{(X_i-\tau)^2A_i^{2c-2}}{\left(1+A_i^c\right)^2}|X_i>d_i\right]$$
$$E_{\boldsymbol{\psi}}\left[\frac{\partial^2\log k\left(\mathbf{x}|\mathbf{y};\boldsymbol{\psi}\right)}{\partial c\partial k}|\mathbf{y}\right]=\sum_{i=1}^{n}\delta_{i,(1,c)}\frac{d_i^c\log(d_i)}{1+d_i^c}-\sum_{i=1}^{n}\delta_{i,(1,c)}E_{\boldsymbol{\psi}}\left[\frac{X_i^c\log(X_i)}{1+X_i^c}|X_i>d_i\right]$$
$$+\sum_{i=1}^{n}\delta_{i,(2,c)}\frac{D_i^c\log(D_i)}{1+D_i^c}-\sum_{i=1}^{n}\delta_{i,(2,c)}E_{\boldsymbol{\psi}}\left[\frac{A_i^c\log(A_i)}{1+A_i^c}|X_i>d_i\right]$$

$$E_{\boldsymbol{\psi}}\left[\frac{\partial^2 \log k(\mathbf{x}|\mathbf{y};\boldsymbol{\psi})}{\partial c \partial \beta}|\mathbf{y}\right] = kc\sum_{i=1}^{n}\delta_{i,(2,c)}\frac{D_i^{c-1}(d_i-\tau)\log(D_i)}{1+D_i^c} + k\sum_{i=1}^{n}\delta_{i,(2,c)}\frac{(d_i-\tau)D_i^{c-1}}{1+D_i^c}$$

$$-kc\sum_{i=1}^{n}\delta_{i,(2,c)}\frac{(d_i-\tau)D_i^{2c-1}\log(D_i)}{(1+D_i^c)^2} + \sum_{i=1}^{n}\delta_{i,(2,c)}E_{\boldsymbol{\psi}}\left[\frac{X_i-\tau}{A_i}|X_i > d_i\right]$$

$$-(k+1)c\sum_{i=1}^{n}\delta_{i,(2,c)}E_{\boldsymbol{\psi}}\left[\frac{A_i^{c-1}(X_i-\tau)\log(A_i)}{1+A_i^c}|X_i > d_i\right]$$

$$-(k+1)\sum_{i=1}^{n}\delta_{i,(2,c)}E_{\boldsymbol{\psi}}\left[\frac{(X_i-\tau)A_i^{c-1}}{1+A_i^c}|X_i > d_i\right]$$

$$+(k+1)c\sum_{i=1}^{n}\delta_{i,(2,c)}E_{\boldsymbol{\psi}}\left[\frac{(X_i-\tau)A_i^{2c-1}\log(A_i)}{(1+A_i^c)^2}|X_i > d_i\right]$$

$$E_{\boldsymbol{\psi}}\left[\frac{\partial^2 \log k(\mathbf{x}|\mathbf{y};\boldsymbol{\psi})}{\partial k \partial \beta}|\mathbf{y}\right]$$

$$= c\sum_{i=1}^{n}\delta_{i,(2,c)}\left[\frac{(d_i-\tau)D_i^{c-1}}{1+D_i^c}\right] - c\sum_{i=1}^{n}\delta_{i,(2,c)}E_{\boldsymbol{\psi}}\left[\frac{(X_i-\tau)A_i^{c-1}}{1+A_i^c}|X_i > d_i\right]$$

The expected values of the second partials of the log-likelihood functions can also be computed by using Monte Carlo integral. Thus, the missing information matrix can be computed from equations (22-26) and is expressed as following:

$$I_{miss}(\boldsymbol{\psi};\mathbf{y}) = E_{\boldsymbol{\psi}}\{I_{miss}(\boldsymbol{\psi};\mathbf{x})|\mathbf{y}\}$$

$$=(-1)\times\begin{bmatrix} E_{\boldsymbol{\psi}}\left\{\frac{\partial^2 \log k(\mathbf{x}|\mathbf{y};\boldsymbol{\psi})}{\partial c^2}|\mathbf{y}\right\} & E_{\boldsymbol{\psi}}\left\{\frac{\partial^2 \log k(\mathbf{x}|\mathbf{y};\boldsymbol{\psi})}{\partial ck}|\mathbf{y}\right\} & E_{\boldsymbol{\psi}}\left\{\frac{\partial^2 \log k(\mathbf{x}|\mathbf{y};\boldsymbol{\psi})}{\partial c\beta}|\mathbf{y}\right\} \\ E_{\boldsymbol{\psi}}\left\{\frac{\partial^2 \log k(\mathbf{x}|\mathbf{y};\boldsymbol{\psi})}{\partial ck}|\mathbf{y}\right\} & E_{\boldsymbol{\psi}}\left\{\frac{\partial^2 \log k(\mathbf{x}|\mathbf{y};\boldsymbol{\psi})}{\partial k^2}|\mathbf{y}\right\} & E_{\boldsymbol{\psi}}\left\{\frac{\partial^2 \log k(\mathbf{x}|\mathbf{y};\boldsymbol{\psi})}{\partial k\beta}|\mathbf{y}\right\} \\ E_{\boldsymbol{\psi}}\left\{\frac{\partial^2 \log k(\mathbf{x}|\mathbf{y};\boldsymbol{\psi})}{\partial c\beta}|\mathbf{y}\right\} & E_{\boldsymbol{\psi}}\left\{\frac{\partial^2 \log k(\mathbf{x}|\mathbf{y};\boldsymbol{\psi})}{\partial k\beta}|\mathbf{y}\right\} & E_{\boldsymbol{\psi}}\left\{\frac{\partial^2 \log k(\mathbf{x}|\mathbf{y};\boldsymbol{\psi})}{\partial \beta^2}|\mathbf{y}\right\} \end{bmatrix}$$

References

[1] Abd-Elfattah, A. M., Hassan, A. S. and Nassr, S. G. (2008). Estimation in step-stress partially accelerated life tests for the Burr type XII distribution using type I censoring. *Statistical Methodology*, 5:502–514.

[2] Abdel-Ghaly, A. A., Amin, Z. H. and Omar, D. A. (2008). Estimation of the Burr-XII distribution for partially accelerated life tests using censored data. *Model Assisted Statistics and Applications*, 3:317–334.

[3] Abdel-Hamid, A. H. (2009). Constant-partially accelerated life tests for Burr type-XII distribution with progressive type-II censoring. *Computational Statistics & Data Analysis*, 53:2511–2523.

[4] Acusta, A., Eggermont, P. and Lariccia, V. (2002). An EM algorithm for density estimation with randomly censored data. *Journal of Statistical Computation and Simulation*, 73:223–232.

[5] Ali Mousa, M. A. M. (1995). Empirical bayes estimators for the Burr type XII accelerated life testing model based on type-2 censored data. *Journal of Statistical Computation and Simulation*, 52:95–103.

[6] Balakrishnan, N. and Kim, J. A. (2004). EM algorithm for Type-II right censored bivariate normal data. In Nikulin MS, Balakrishnan N, Mesbah M, Limnios N. (Eds.), *Advances in Parametric and Semi-Parametric Inference with Applications in Reliability, Survival Analysis and Quality of Life.* Boston: Birkhauser, 177–210.

[7] Cheng, Y.-F. and Wang, F.-K. (2012). Estimating the Burr XII parameters in constant-stress partially accelerated life tests under multiple censored data. *Communications in Statistics - Simulation and Computation*, 41:1711–1727.

[8] Dempster, A. P., Laird, N. M and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39:1–38.

[9] Louis, T. A. (1982). Finding the observed information matrix when using the EM Algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 44:226–233.

[10] Ng, H. K. T., Chan, P. S. and Balakrishnan, N. (2002). Estimation of parameters from progressively censored data using EM algorithm. *Computational Statistics and Data Analysis*, 39:371–386.

[11] Park, C. (2005). Parameter estimation of incomplete data in competing risks using the EM algorithm. *IEEE Transactions on Reliability*, 54:282–290.

[12] Rodriguez, R. N. (1977). A guide to the Burr type XII distributions. *Biometrika*, 64:129–134.

[13] Wang, F.-K., Cheng, Y.-F. and Lu, W.-L. (2012). Partially accelerated life tests for the Weibull distribution under multiply censored data. *Communications in Statistics - Simulation and Computation*, 41:1667–1678.

[14] Wang, F. K., Keats, J. B. and Zimmer, W. J. (1996). Maximum likelihood estimation of the Burr XII parameters with censored and uncensored data. *Microelectronics and Reliability*, 36:359–362.

[15] Watkins, A. J. (1999). An algorithm for maximum likelihood estimation in the three parameter Burr XII distribution. *Computational Statistics and Data Analysis*, 32:19–27.

[16] Wingo, D. (1993). Maximum likelihood methods for fitting the burr type XII distribution to multiple (progressively) censored life test data. *Metrika*, 40:203–210.

[17] Zimmer, W. J., Keats, J. B. and Wang, F. K. (1998). The Burr XII distribution in reliability analysis. *Journal of Quality Technology*, 30:386–394.

Received: September 11, 2020
Accepted: November 25, 2020