Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/060106

Bayesian Estimation of the Parameters of Exponentiated Exponential Distribution under Progressive Interval Type-I Censoring Scheme with Binomial Removals

*Arun Kaushik**

Department of Statistics and DST-CIMS, Banaras Hindu University, Varanasi-221 005, India

Received: 6 Oct. 2016, Revised: 15 Jan. 2017, Accepted: 17 Jan. 2017
Published online: 1 Mar. 2017

Abstract: In this article, we consider the problem of point and interval estimation of the parameters of exponentiated exponential distribution (EED) under progressive type-I interval (PTII) censoring scheme with random removals. Maximum likelihood, expectation maximization and Bayesian procedures have been developed for the estimation of parameters of the EED, based on a PTII censored sample. Twp real examples has been considered to illustrate the applicability of the proposed methodology for the considered censoring scheme. Further, we have compared the performances of the proposed estimators under PTII censoring with complete sample case.

Keywords: Statistical Computing, Bayesian, Maximum likelihood, Simulation, Progressive Interval Censoring

1 Introduction

Recently, exponentiated exponential distribution has gained popularity in the statistical literature because of its simplicity and various shape of probability density function(pdf). [11] introduced exponentiated exponential distribution (EED) as an alternative to Gamma and Weibull distribution. [15], [17] and many others have further studied it. [11] noted that, in many situations, the two-parameter exponentiated exponential distribution provides a better fit than the two-parameter Weibull distribution. It is worthwhile to noted here that EE distribution is a special case of a distribution that was used by [10].

The probability density function of EE distribution is given below,

$$f(x|\alpha,\theta) = \alpha\theta e^{-\theta x}(1-e^{-\theta x})^{\alpha-1}; \qquad x \geq 0, \alpha, \theta > 0, \tag{1}$$

where α is the shape parameter and λ is the scale parameter of considered distribution. Its cumulative distribution and survival functions function are given by,

$$F(x|\alpha,\theta) = (1-e^{-\theta x})^{\alpha} \tag{2}$$

and

$$S(x|\alpha,\theta) = 1-(1-e^{-\theta x})^{\alpha}; \qquad x \geq 0, \alpha, \theta > 0, \tag{3}$$

respectively. [12] studied different methods of point estimation for EED parameters which include maximum likelihood estimation, method of moment estimation and probability plot method of estimation based on complete sample. [18] discussed the parameter estimation and reliability characteristic of EED under Bayesian paradigm. It is worthwhile to mention here that a very little attention has been paid to the inferences based on censored sample from EED under Bayesian paradigm, although censoring is quite common in various clinical and life testing experiments.

Situations do arise when the units under study are lost or removed from the experiments while they are still alive i.e., we get censored data in such cases. If the censoring is time dependent, it is called Type-I censoring. On the other hand, if it is unit dependent, it is called Type-II censoring. Depending on the need and practical considerations, various modified forms of censoring schemes have been discussed in the literature. [1] proposed a combination of interval Type-I censoring

* Corresponding author e-mail: arundevkauhsik@gmail.com

and progressive censoring called as progressive Type-I interval (PTII) censoring which naturally arises in most of the clinical experiments. To have a clear visualization of this censoring scheme, let us consider an experiment with n bladder cancer patients for whom remission times are to be recorded. The patients are called for regular check up at scheduled times and those who turn up are checked. At the first visit, scheduled at time T_1, only $n-R_1$ patients out of the total n patients report i.e. R_1 patients leave the experiment during the time interval $(0,T_1]$. Experimenter examines these $n-R_1$ patients and finds that cancer has reoccurred in D_1 patients. It may be noted here that the exact time of recurrence for these D_1 patients are not known to the experimenter; he only has the information about the number of recurrences during the time period between start of the experiment and first visit. At second visit, scheduled at time T_2, $n-R_1-D_1-R_2$ out of the remaining $n-R_1-D_1$ patients in the experiment after their first visit report i.e. R_2 patients leave the experiment at this stage (during the time interval $(T_1,T_2]$). Experimenter examines these patients and finds that cancer has reoccurred in D_2 patients out of remaining $n-R_1-D_1-R_2$ patients, and in this way the experiment continues till the m^{th} visit. At this stage (m^{th} visit) all the remaining $R_m = n-D_1-D_2\cdots-D_m-R_1-R_2\cdots R_{m-1}$ units are removed i.e. the experiment is terminated at this stage. Recently [7] proposed the methodology of the estimation of parameters involve in EED under PTII censored case under the assumption that the proportions (p_i) of the patients leaving the experiment during $(T_{i-1},T_i]$ is known in advance i.e. they prefixed the proportion $p_1, p_2, \ldots, p_m$ and considered that at i^{th} stage, $\lfloor n_i * p_i \rfloor$ patients shall leave the experiment. Here, $\lfloor n_i * p_i \rfloor$ denotes the largest integer less than or equal to $n_i * p_i$. The author's claim that exactly $\lfloor n_i * p_i \rfloor$ patients out of $\lfloor n_i \rfloor$ will drop out of the experiment at i^{th} stage (visit) , seems unrealistic and hypothetical. In fact, the number of patients dropping out from the clinical trial at any stage is beyond the control of the experimenter and can not be predetermined. It seems more logical and natural to consider these as random variables subjected to the risk of drooping at i^{th} stage as p_i. Perhaps, keeping a similar thought in mind, [21] discussed progressive censoring scheme with binomial removal. [2] and [20] have used PTII censoring scheme with binomial removals assuming that the exact value of the life times of the units are observable. In their studies, they have assumed that the number of removals $R_i^{,}s$ at i^{th} stage ($i = 1, 2, \ldots, m$) is random and follow the binomial distribution with probability p_i. Thus, R_1 (at 1^{st} stage) may be considered to follow *Binomial* $(n,\ p_1)$ distribution and R_2 (at 2^{nd} stage) follows *Binomial* $(n-D_1-R_1,\ p_2)$. In general, the number of units dropping at i^{th} stage, R_i follows the binomial distribution with parameter $(n-\sum_{l=1}^{i} D_l + R_l,\ p_i)$ for $i = 1, 2, 3, \cdots, m-1$. In this paper, we will consider PTII censored data with binomial removals and develop estimators for the shape and scale parameter under the situation that the exact value of the life times of the units are not observable, only the number of observations lying in the specified interval of times are known. For parameter estimation problem, we have considered the most popular loss function, namely the squared error loss function (SELF) which can be easily justified on the grounds of minimum variance unbiased estimation (see [5]). We will compare the performance of the proposed estimators of the parameters obtained under above stated censoring scheme with the estimates under complete sample case.

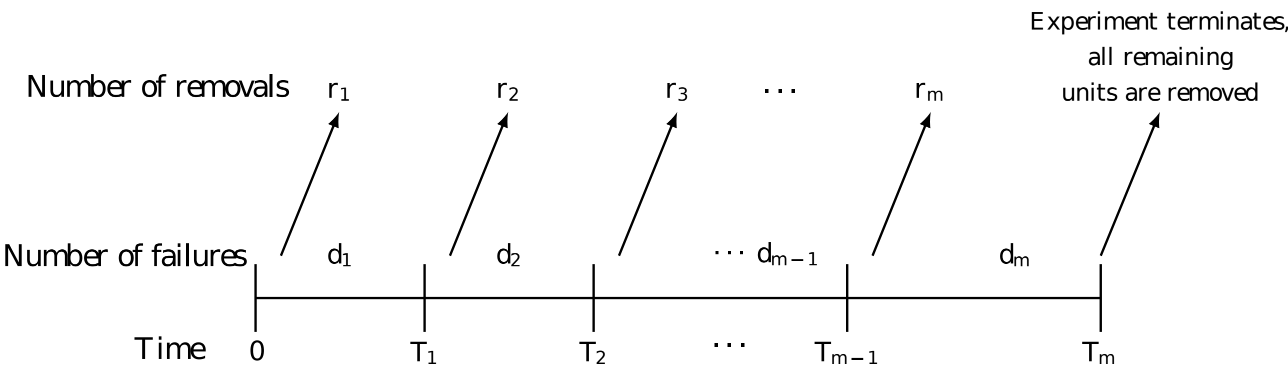
Rest of the paper comprises of the following sections. In Section 2, classical and Bayes procedures for the estimation of the model parameters based on PTII with binomial removal samples have been developed. Two real data set, first related to the survival time of patients with plasma cell myeloma and second regarding the number of million revolutions before failure of groove ball bearings have been considered for the illustration of the proposed methodology in Section 3. Comparison of the estimators based on simulation study has been provided in Section 4. Finally, conclusions have been summarized in Section 5.

2 Parameter Estimation

2.1 Maximum Likelihood Estimation

In this section, we provide the MLEs of α and θ; the parameters of the lifetime distribution given in equation (1). Let us consider that n units are put on test initially at time $T_0 = 0$ and we record the number of droppings and number of failures during pre-specified times intervals $(T_{i-1},T_i]$ $(i = 1, 2, \ldots, m)$ amongst the available units; i.e. we get the data of consisting number of failures $D = (d_1, d_2, \ldots, d_m)$ and number of droppings $R = (r_1, r_2, \ldots, r_m)$ during the time interval ($(0,T_1]$, $(T_1, T2]$, ..., $(T_{m-1},T_m]$) through the censoring scheme described in the previous section. It may be noted here that the

individual units dropping from the test at i^{th} stage (during the time interval $(T_{i-1},T_i]$), $i = 1, 2, \cdots (m-1)$ are random and independent of each other with certain probability of removal, say p_i, for $i = 1, 2, \cdots (m-1)$. Therefore, the number R_i of the unit dropping at i^{th}; $i = 1, 2, \cdots (m-1)$ stage follows binomial distribution with parameters $\left(n-m-\sum_{l=1}^{i-1} r_i, p_i\right)$



1

Fig. 1: Progressive Type-I Interval Censoring Scheme

i.e.

$$P(r_1|d_1,p_1) = \binom{n-d_1}{r_1} p_1^{r_1}(1-p_1)^{n-d_1-r_1}$$

$$P(r_2|r_1,d_2,p_2) = \binom{n-d_1-d_2-r_1}{r_2} p_2^{r_2}(1-p_2)^{n-d_1-d_2-r_1-r_2}$$

and in general

$$P(r_i|r_{i-1},d_i,p_i) = \binom{n-\sum\limits_{j=1}^{i} d_j - \sum\limits_{j=1}^{i-1} r_j}{r_i} \quad p_i^{r_i}(1-p_i)^{n-\sum\limits_{j=1}^{i} d_j-\sum\limits_{j=1}^{i} r_j}$$

Now the complete likelihood for the observed data can easily be written as

$$\begin{aligned} L(\alpha,\theta|R,D,T) \propto & \prod_{i=1}^{m-1} [F(T_i)-F(T_{i-1})]^{d_i} \times [1-F(T_i)]^{r_i} \times P(r_i|r_{i-1},d_i,p_i) \\ & \times [F(T_m)-F(T_{m-1})]^{d_m}[1-F(T_m)]^{r_m} \\ = & \prod_{i=1}^{m}\left[\left(1-e^{-\theta T_i}\right)^{\alpha}-\left(1-e^{-\theta T_{i-1}}\right)^{\alpha}\right]^{d_i}\left[1-\left(1-e^{-\theta T_i}\right)^{\alpha}\right]^{r_i} \\ & \times \binom{n-\sum\limits_{j=1}^{i} d_j - \sum\limits_{j=1}^{i-1} r_j}{r_i} \quad p^{r_i}(1-p)^{n-\sum\limits_{j=1}^{i} d_j-\sum\limits_{j=1}^{i} r_j}. \end{aligned} \tag{4}$$

Above expression bifurcates as

$$L(\alpha,\theta|R,D,T) \propto L_1(\alpha,\theta|R,D,T)L_2(P|R,D,T), \tag{5}$$

where

$$L_1(\alpha,\theta|R,D,T) = \prod_{i=1}^{m}\left[\left(1-e^{-\theta T_i}\right)^{\alpha}-\left(1-e^{-\theta T_{i-1}}\right)^{\alpha}\right]^{d_i}\left[1-\left(1-e^{-\theta T_i}\right)^{\alpha}\right]^{r_i} \tag{6}$$

Note that L_2 is free from α and θ. Thus, to compute ML estimate of α and θ, we require only L_1. The corresponding log likelihood function can be written as

$$\begin{aligned} logL_1(T,\alpha,\theta) = C & \sum_{i=1}^{m} d_i \ln\left(\left(1-e^{-\theta T_i}\right)^{\alpha}-\left(1-e^{-\theta T_{i-1}}\right)^{\alpha}\right) \\ & + r_i \ln\left(1-\left(1-e^{-\theta T_i}\right)^{\alpha}\right). \end{aligned} \tag{7}$$

Hence, the likelihood equations can be obtained as;

$$\frac{dL}{d\alpha} = \sum_{i=1}^{m} d_i \frac{\left[\left(1-e^{-\theta T_i}\right)^{\alpha} ln(1-e^{-\theta T_i}) - \left(1-e^{-\theta T_{i-1}}\right)^{\alpha} ln(1-e^{-\theta T_{i-1}})\right]}{\left[(1-e^{-\theta T_i})^{\alpha} - \left(1-e^{-\theta T_{i-1}}\right)^{\alpha}\right]}$$

$$- r_i \frac{\left(1-e^{-\theta T_i}\right)^{\alpha} ln(e^{-\theta T_i})}{\left[1-(1-e^{-\theta T_i})^{\alpha}\right]} = 0 \quad (8)$$

$$\frac{dL}{d\theta} = \sum_{i=1}^{m} d_i \frac{\left[\left(1-e^{-\theta T_i}\right)^{\alpha-1} e^{-\theta T_i}\alpha T_i - \left(1-e^{-\theta T_{i-1}}\right)^{\alpha} e^{-\theta T_{i-1}}\alpha T_{i-1}\right]}{\left[(1-e^{-\theta T_i})^{\alpha} - \left(1-e^{-\theta T_{i-1}}\right)^{\alpha}\right]}$$

$$- r_i \frac{\left(1-e^{-\theta T_i}\right)^{\alpha-1} e^{-\theta T_i}\alpha T_i}{\left[1-(1-e^{-\theta T_i})^{\alpha}\right]} = 0 \quad (9)$$

The MLEs of α and θ can be obtained by solving (8) and (9) simultaneously. But it may be noted here that explicit solutions cannot be obtained from the above equations. Thus, we propose the use of a suitable numerical technique to solve these two non-linear equations. One may use Newton-Raphson or simulated Annealing or their variants to solve these equations. This can be routinely done using R, Matlab, Mathcad or other packages. We have also obtained the observed information matrix,

$$I(\alpha,\theta|data) = \begin{bmatrix} -L_{\alpha\alpha} & -L_{\alpha\theta} \\ -L_{\theta\alpha} & -L_{\theta\theta} \end{bmatrix}, \quad (10)$$

where, all the second partial derivatives of the log-likelihood function $L_{\alpha\alpha}$, $L_{\alpha\theta}$ and $L_{\theta\theta}$ are provided in the Appendix-A. Based on it, the asymptotic confidence (AC) interval and standard errors of the parameter estimates can be obtained in the usual way. While using the standard Newton-Raphson algorithm (the details are provided in the simulation section) to compute the MLEs for the parameters, it is observed that the iterations converge approximately $85\% - 90\%$ of the times. In order to have a high convergence rate, we used the EM algorithm also, and it is noted that it has a convergence rate more than 99%.

2.2 EM method of estimation

The expectation maximization (EM) algorithm is broadly applicable approach of the iterative computation of maximum likelihood estimates and it is useful in a variety of incomplete data problems where algorithms such as the Newton-Raphson method, often, turn out to be more complicated and has less convergence rate. Each iteration of the EM algorithm, consist of two steps, namely the expectation step (E-step) and the maximization step (M-step). Therefore, the algorithm is called as EM algorithm. The details about the EM algorithm can be found in [8]. For the present study, the EM algorithm for finding the MLEs of parameters in the two-parameter EE distribution is developed as follows.

We assume that $\tau_{i,j},\ j=1,2,\ldots,d_i$, denote the exact survival times of those d_i units which fail within subinterval $(T_{i-1},T_i]$ and $\tau^*_{i,j},\ j=1,2,\ldots,r_i$ would have been the exact survival times for those r_i units which drop within time interval $(T_{i-1},T_i]$. The log likelihood, $ln(L^C)$, for the lifetimes of n items following the EE distribution can be written as:

$$ln(L^C) \propto \sum_{i=1}^{m}\left[\sum_{j=1}^{d_i}\ln(f(\tau_{ij})) + \sum_{j=1}^{r_i}\ln\left(f(\tau^*_{ij})\right)\right]$$

$$= [\ln(\alpha)+\ln(\theta)]\left(\sum_{i=1}^{m}(d_i+r_i)\right) - \theta\sum_{i=1}^{m}\left(\sum_{j=1}^{d_i}\tau_{i_j} + \sum_{j=1}^{r_i}\tau^*_{ij}\right)$$

$$+(\alpha-1)\sum_{i=1}^{m}\left[\sum_{j=1}^{d_i}\ln\left(1-e^{-\theta\tau_{ij}}\right) + \sum_{j=1}^{r_i}\ln\left(1-e^{-\theta\tau^*_{ij}}\right)\right] \quad (11)$$

Equating the partial derivatives of the log likelihood given in equation (11) with respect to α and θ, the corresponding likelihood equations can be obtained as follows:

$$\frac{-n}{\alpha} = \sum_{i=1}^{m}\left[\sum_{j=1}^{d_i}\ln\left(1-e^{-\theta\tau_{ij}}\right) + \sum_{j=1}^{r_i}\ln\left(1-e^{-\theta\tau^*_{ij}}\right)\right] \quad (12)$$

and

$$\frac{n}{\theta}=\sum_{i=1}^{m}\left[\sum_{j=1}^{d_i}\tau_{ij}+\sum_{j=1}^{r_i}\tau_{ij}^{*}\right]-(\alpha-1)\sum_{i=1}^{m}\left[\sum_{j=1}^{d_i}\frac{\tau_{ij}}{e^{\theta\,\tau_{ij}}-1}+\sum_{j=1}^{r_i}\frac{\tau_{ij}^{*}}{e^{\theta\,\tau_{ij}^{*}}-1}\right]. \tag{13}$$

It is evident that the lifetimes of the d_i failures in the i^{th} interval $(T_{i-1},T_i]$ are independent and all follow a doubly truncated EE distribution with left truncation point at T_{i-1}. Further right truncation point at T_i and the lifetimes of the r_i censored items in the i^{th} interval $(T_{i-1},T_i]$ are independent and follow a left truncated EE distribution at T_i-1.

The EM algorithm proceeds in this case through the following iterative process:

①Choose, starting values of α and θ , say $\hat{\alpha}_{(0)}$ and $\hat{\theta}_{(0)}$ and set $k=0$.
②At the $k+1th$ iteration,

✓The E-step considers expectations of the terms involving τ_{ij} and τ_{ij}^{*} in equations (12) and (13), which finally reduces to the following:

$$\frac{n}{\alpha^{(k)}}=-\sum_{i=1}^{m}\left(d_iE_{2i}+r_iE_{4i}\right) \tag{14}$$

$$\frac{n}{\theta^{(k)}}=\sum_{i=1}^{m}\left(d_iE_{1i}+r_iE_{3i}\right)-(\alpha-1)\sum_{i=1}^{m}\left(d_iE_{5i}+r_iE_{6i}\right), \tag{15}$$

where, $E_{1i},E_{2i},\ldots,E_{6i}$ are given in Appendix-B

✓The M-step requires solving of (14) and (15) to obtain the next values, $\alpha^{(k+1)}$ and $\theta^{(k+1)}$, of α and θ, respectively. The values thus obtained are as follows:

$$\hat{\alpha}^{(k+1)}=-\frac{n}{\sum_{i=1}^{m}\left(d_iE_{2i}+r_iE_{4i}\right)} \tag{16}$$

$$\hat{\theta}^{(k+1)}=\frac{n}{\sum_{i=1}^{m}\left(d_iE_{1i}+r_iE_{3i}\right)-\left(\hat{\alpha}^{(k+1)}-1\right)\left(d_iE_{5i}+r_iE_{6i}\right)} \tag{17}$$

③If the convergence occurs then the current $\hat{\alpha}^{(k)}$ and $\hat{\theta}^{(k)}$ are the approximate maximum likelihood estimates α of θ via EM algorithm; otherwise, set $k=k+1$ and go to Step ②

It can be easily seen that the EM algorithm has no complicated likelihood equations involved for finding the solutions, rather directly provide the maximum likelihood estimates of α and θ. Therefore, the algorithm can be efficiently implemented through a computing program. We have developed the program in R language and used it for the evaluation of the estimates.

2.3 Bayesian Estimation

In this section, we provide the Bayesian inferences for α and θ , when we have the progressive type-I interval censored data as explained in figure 1. We have also obtained the highest posterior density (HPD) intervals for both the parameters. Before proceeding further, we make selection for the prior distributions of the parameters. Following [4], it is assumed that both α and θ are independent gamma variates, having pdfs

$$g_1(\alpha)=\frac{\lambda_1^{\nu_1}}{\Gamma(\nu_1)}e^{-(\lambda_1\alpha)}\alpha^{(\nu_1-1)}\quad;\qquad 0<\alpha<\infty,\ \lambda_1>0,\nu_1>0 \tag{18}$$

and

$$g_2(\theta)=\frac{\lambda_2^{\nu_2}}{\Gamma(\nu_2)}e^{-(\lambda_2\theta)}\theta^{(\nu_2-1)}\quad;\qquad 0<\theta<\infty,\ \lambda_2>0,\nu_2>0, \tag{19}$$

Here, all the hyper parameters λ_1, ν_1, λ_2 and ν_2 are assumed to be known and can be evaluated following the method as suggested by [19]. We compute the Bayes estimate of the unknown parameters under squared error loss function. Using the priors given in (18) and (19) and the likelihood function (4), the joint posterior density of α and θ for the given data can be written as

$$\begin{aligned}\pi(\alpha,\theta|R,D,T)&=\frac{L(\alpha,\theta|R,D,T)\,g_1(\alpha)\,g_2(\theta)}{\int_0^{\infty}\int_0^{\infty}L(\alpha,\theta|R,D,T)\,g_1(\alpha)\,g_2(\theta)\,d\alpha\,d\theta}\\&=\frac{J}{\iint_0^{\infty}J\,d\alpha\,d\theta},\end{aligned} \tag{20}$$

where,

$$J = J(\alpha,\theta) = e^{-(\lambda_1\alpha+\lambda_2\theta)}\alpha^{(\nu_1-1)}\theta^{(\nu_2-1)}\prod_{i=1}^{k}\left[(1-e^{-\theta T_i})^{\alpha}-(1-e^{-\theta T_{i-1}})^{\alpha}\right]^{d_i} \times\left[1-(1-e^{-\theta T_i})^{\alpha}\right]^{r_i}.$$

Let $h(\cdot)$ be a function of α and θ. Then, the Bayes estimate of $h(\cdot)$ under squared error loss function is given by

$$\hat{h}_B(\alpha,\theta) = E_\pi(h(\alpha,\theta)) = \frac{\iint_0^\infty h(\alpha,\theta)\,J\,d\alpha\,d\theta}{\iint_0^\infty J\,d\alpha\,d\theta}. \tag{21}$$

It is clear from the expression (20) that there is no closed form for the estimators, so we suggest MCMC procedure to compute the Bayes estimates. After getting MCMC samples from posterior distribution, we can find the Bayes estimate for the parameters in the following way

$$[E(\Theta|data)] = \left[\frac{1}{N-N_0}\sum_{i=N_0+1}^{N}\Theta_i\right],$$

where N_0 is burn-in period of Markov chain. For computation of the highest posterior density (HPD) interval of Θ, order the MCMC sample of Θ as $\Theta_{(1)},\Theta_{(2)},\Theta_{(3)},\cdots,\Theta_{(N)}$. Then construct all the 100(1-γ)% credible intervals of Θ say $(\Theta_{(1)},\Theta_{(N\lfloor 1-\gamma\rfloor+1)})$, $(\Theta_{(2)},\Theta_{(N\lfloor 1-\gamma\rfloor+2)})\cdots$,$(\Theta_{(\lfloor N\gamma\rfloor)},\Theta_{(N)})$. Finally, the HPD credible interval of α and β is that interval which has the shortest length.

In order to obtain the MCMC samples from the joint posterior density of α and θ, we have used the Metropolis-Hastings (M-H) algorithm. We have considered a bivariate normal distribution as the proposal density i.e. $N_2(\mu,\Sigma)$ where Σ is the variance-covariance matrix. It may be noted here that if we generate observation from bivariate normal distribution, we may get negative values also which are not possible as the parameters under consideration are positive valued. Therefore, we take absolute value of generated observation. Following this, the Metropolis-Hastings algorithm associated with the target density $\pi(\cdot)$ and the proposal density $N_2(\mu,\Sigma)$ produces a Markov chain Θ^i through the following steps.

①Set initial values $\Theta^0 = [\alpha^0,\theta^0]'$.
②Generate new candidate parameter values $\Theta^* = [\alpha^*,\theta^*]'$ from $N_2(\mu,\Sigma)$.
③Calculate the ratio

$$\rho\left(\Theta^*,\Theta^{i-1}\right) = \min\left\{\frac{\pi(\Theta^*)}{\pi(\Theta^{i-1})},1\right\}.$$

④Accept candidate Θ^* as

$$\Theta^i = \begin{cases}\Theta^* \text{ with probability } \rho\left(\Theta^*,\Theta^{i-1}\right)\\ \Theta^{i-1} \text{ with probability } 1-\rho\left(\Theta^*,\Theta^{i-1}\right)\end{cases}$$

In using the above, algorithm, the problem arises how to choose the initial guess. Here, we propose the use of MLEs of (α,θ), obtained by using the method described in sub section 2.1, as initial value for MCMC process. The choice of covariance matrix Σ is also an important issue, see [14] for details. One choice for Σ would be the asymptotic variance-covariance matrix $I^{-1}(\hat{\alpha},\hat{\theta})$. While generating M-H samples by taking $\Sigma = I^{-1}(\hat{\alpha},\hat{\theta})$, we noted that the acceptance rate for such a choice of Σ is about 15%. By acceptance rate, we mean the proportion of times a new set of values is generated at the iteration stages. It is well known that if the acceptance rate is low, a good strategy is to run a small pilot run using diagonal Σ as a rough estimate of the correlation structure for the target posterior distribution and then re-run the algorithm using the corresponding estimated variance-covariance matrix; for more detail see [9, pp. 334-335]. Therefore, we have also used the later described strategy for the calculations in the following sections.

3 Illustration

In this section, we illustrate our proposed methodology with the real examples. The first data set considered by us, represent the survival times for patients with plasma cell myeloma, already reported in [6]. Data contains the response time to therapy of 112 patients with plasma cell myeloma (a tumor of the bone marrow composed of cells normally found

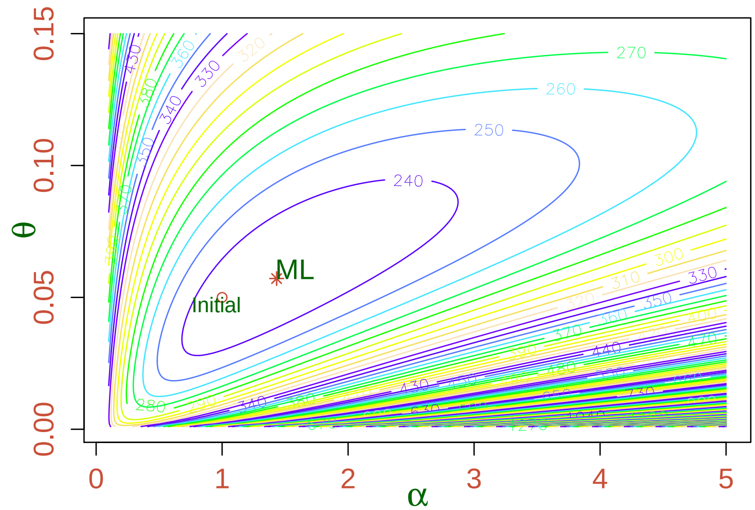


Fig. 2: Contour plot for plasma cell myeloma data

in bone marrow) treated at the National Cancer Institute, Bethesda, Maryland. Figure 2 represents the contour plot of negative log-likelihood for the considered data set. The ellipses are obtained by joining those points which are having equal values of the negative log-likelihood. Every inner ellipse has smaller value than that of outer ellipse. Thus, the inner most ellipse has the minimum value. In other words, the minimum of minus log-likelihood will (maximum of the likelihood) correspond to the inner most ellipse. We used an arbitrary point (1,0.05) from this inner most ellipse, as initial guess. The ML estimates for the data set is then calculated, using the procedure explained in sub-section 2.1. Finally, these are obtained as $\hat{\alpha}_{ML} = 1.4325$, $\hat{\theta}_{ML} = 0.0571$. Similarly, 95% asymptotic confidence intervals for α is obtained as (0.9706, 1.8944) and for β as (0.0420, 0.0727). The same initial guess (1,0.05) have been used for EM algorithm also and using the procedure explained in sub-section 2.2, we found the estimates $\hat{\alpha}_{EM} = 1.4247$, $\hat{\theta}_{EM} = 0.05783$.

To compute Bayes estimates for considered data set, we used MCMC technique discussed in section 2.3. Following [16], we have run three MCMC chains with initial values selected as MLE, MLE - (asymptotic standard deviation) and MLE + (asymptotic standard deviation), respectively. Figure 3 shows the iterations and density plot of samples generated from the posterior distribution using MCMC technique. From this figure, we see that all the three chains have converged and are well mixed. It is, further, noted that the posterior of α is approximately symmetric but posterior of θ is right skewed. Utilizing these MCMC samples, we computed Bayes estimates, following the method discussed in section 2.3, and got $\hat{\alpha}_B = 1.4301$, $\hat{\theta}_B = 0.0581$ under non-informative independent priors. The 95% highest posterior density (HPD) interval estimates for α is obtained as (1.0001,1.6109) and for θ as (0.0424, 0.0719).

The second data set, considered here, arose in the tests on endurance of deep groove ball bearings. This data contains the number of million revolutions before failure for each of the 23 ball bearings in the life test and has been reported by [13, pp.228]. The data points are exact observations. For the illustration of our methodology, we have generated censored data for prefixed number of inspections by specifying the inspection times and dropping probabilities.

We fixed the experimentation time as 140 unit of time and decided to have 7 inspections during this period. We have considered four different inspection plans. The first plan consists of equally spaced inspection time i.e. at 20, 40, ..., 140 units of time. The next inspection plan is designed under the motivation that if probability of failure is high during some time interval, an early inspection should be scheduled. Thus, the second inspection plan is based on such a notion. The third inspection plan is designed on the basis of estimated cdf; although such a plan is not feasible in practice but we have included it for theoretical interest. First, we calculate $u = F(140, \alpha_{ML}, \theta_{ML})$, then inspection times is obtained as $T_1 = F^{-1}(u/7, \alpha_{ML}, \theta_{ML})$, $T_2 = F^{-1}(2u/7, \alpha_{ML}, \theta_{ML}), \ldots, T_6 = F^{-1}(6u/7, \alpha_{ML}, \theta_{ML})$ and $T_7 = 140$. The fourth inspection plan is chosen so as to have approximately equal probability of failure in each interval of inspection and are approximated to the nearest multiple of 10. The dropping schemes are selected in the following manner: First scheme considers the risk of dropping at all the intermediate stages to be zero i.e. $p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = p_7 = 0, p_8 = 1$. In the second scheme, risk at all stages is equal but not zero i.e. $p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = p_7 = 0.2, p_8 = 1$. Third scheme is constructed so that the risk of dropping is low in the earlier stages and high in latter stages. Contrary to it, in the fourth scheme, risk of dropping is high in earlier stages and low in the latter stages. Lastly, we consider the case, when risk is high at first stage but no risk at all other stages. These inspection schemes and dropping schemes are summarizes in table 1b

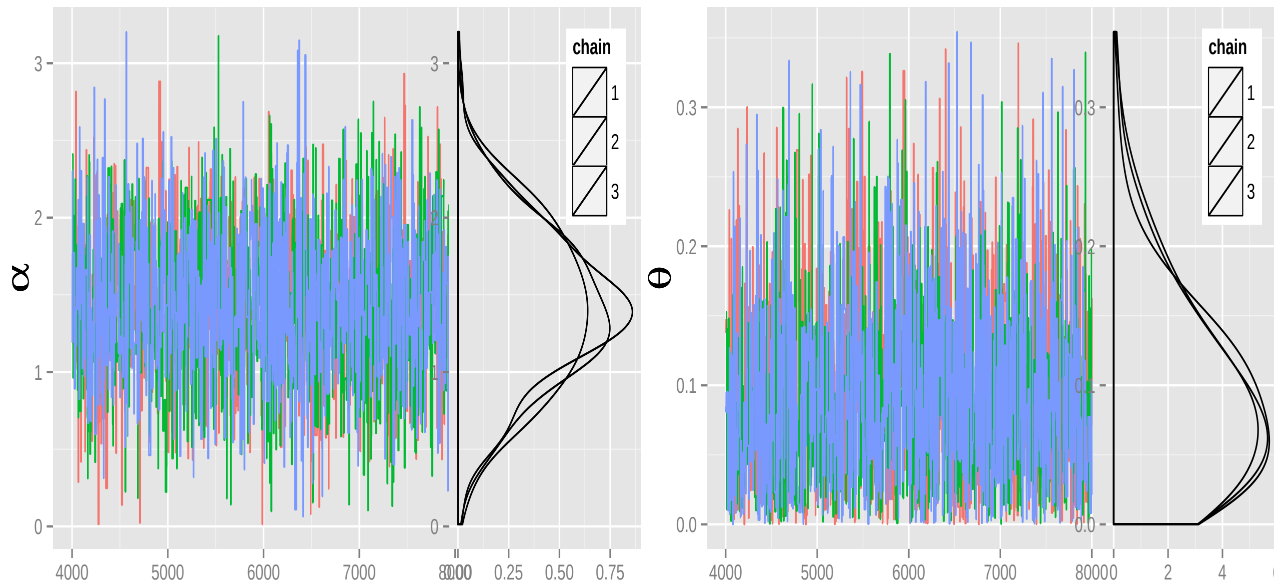


Fig. 3: Iteration and density plot of MCMC samples for plasma cell myeloma data

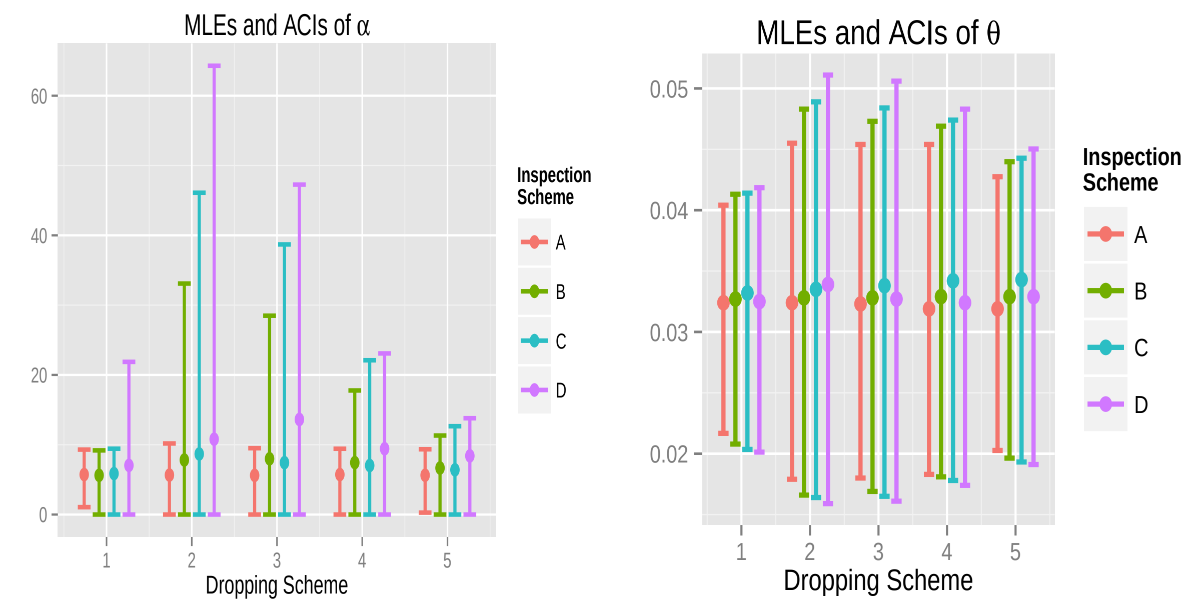
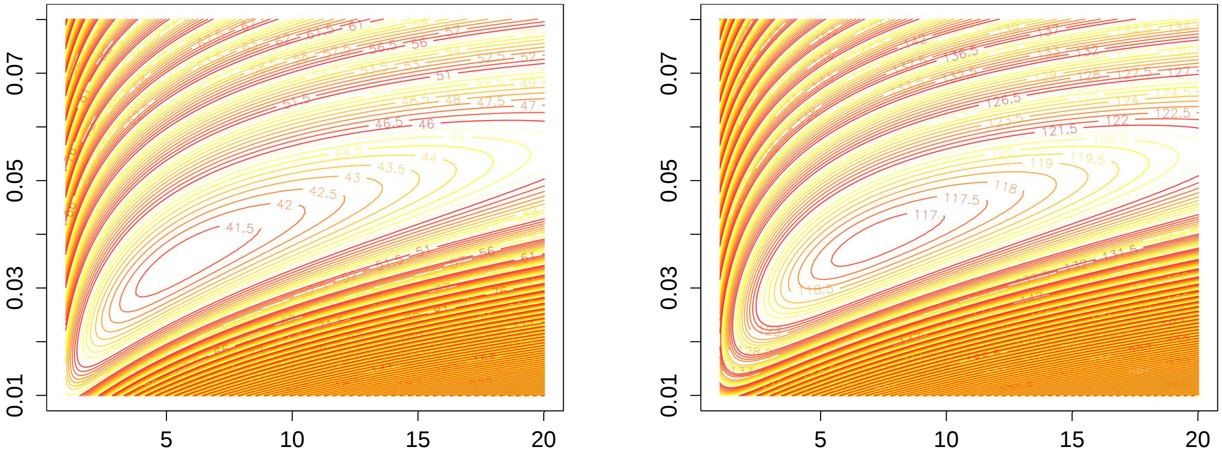


Fig. 4: MLEs and Interval estimates for model parameters under different choice of censoring scheme

and table 1a, respectively. Under the dropping scheme 1 and inspection scheme A, we obtained the number of failures at seven stages as 1, 2, 8, 4, 3, 2, 2 respectively and zero droppings at all the stages. Figure 5 represents the contour plots for this generated artificial censored sample as well as for complete data set and from this we have chosen the initial guess for the computation of ML estimates. Following the same procedure, as followed in previous example, we calculated the ML estimates, EM estimates, Interval estimates and Bayes estimates for the data sets as mentioned above.



(a) Censored **(b)** Complete

Fig. 5: Contour plot for censored data with Inspection scheme A when dropping scheme 1 and complete data set

This result is summarized in the first row of table 2. The last row of the table provides the ML and Bayes estimates with corresponding interval estimates for complete data set.

It may be worthwhile to mention here that the number of droppings are random and we are generating the progressive type-I interval censored data from the complete sample data, therefore we can study the average performance of the estimators. For this purpose, we generated 2000 censored data sets of R_i's for given p_i's and accordingly the d_i's from considered complete data set. Figure 4 shows the average ML estimates with corresponding average asymptotic confidence interval for both the parameters. Table 2 provides the average ML and Bayes estimates, along with AC and HPD interval estimates of the parameters based on the generated censored data sets. It may be seen from the table that the width of the interval estimates under dropping scheme 1 when risk of droppings at all stages is zero, is least of all the estimators under other schemes. It may further be seen that width of the interval estimates under the dropping scheme 2 is more than others. Further, under the 4^{th} scheme the interval width is lesser than those under 3rd scheme. While studying the effect of inspection time on the performance of the estimators, we noted that the average estimate under inspection scheme A and dropping scheme 1 is close to the estimate obtained under complete sample case. For other inspection and dropping schemes the average estimates are larger than that obtained under complete sample case. Similarly, the average width of the interval estimates under scheme A is least among all considered inspections schemes. The width of the interval estimates under scheme B is more than that of under scheme A but less than under scheme C. The width of the interval estimates under scheme D is largest. It is also noted that as proportion of droppings increases, width of the interval estimates increase.

Table 1: Censoring scheme

(a) Dropping Scheme

Name	Dropping Probabilities
1	0*7,1
2	0.2*7,1
3	0.2*3,0.1*4,1
4	0.1*4,0.2*3,1
5	0.25,0*6,1

(b) Inspection Scheme

Name	Inspection time
A	20 40 60 80 100 120 140
C	33.00 45.60 51.96 67.80 68.88 98.64 140
B	36.96 46.85 55.88 65.55 77.31 94.54 140
D	40 50 65 70 100 120 140

Table 2: Average ML, EM and Bayes estimates of α and θ along with their AC and HPD intervals for different dropping and inspection scheme related to the ball bearing data set

Inspection Scheme	Dropping scheme	$\hat{\alpha}_{ML}$	AC Interval of α		$\hat{\alpha}_{EM}$	$\hat{\alpha}_B$	HPD Interval of α		$\hat{\theta}_{ML}$	AC Interval of θ		$\hat{\theta}_{EM}$	$\hat{\theta}_B$	HPD Interval of θ	
			Lower	Upper			Lower	Upper		Lower	Upper			Lower	Upper
A	1	5.7256	1.0579	9.3032	5.7265	5.7256	1.2590	8.2030	0.0324	0.0197	0.0449	0.0328	0.0324	0.0248	0.0411
A	2	5.6571	0.0000	10.1801	5.6573	5.6574	0.3648	9.0799	0.0324	0.0179	0.0455	0.0328	0.0328	0.0240	0.0415
A	3	5.6128	0.0000	9.5055	5.6137	5.6140	0.0673	8.4045	0.0323	0.0180	0.0454	0.0327	0.0323	0.0241	0.0415
A	4	5.7293	0.0000	9.4265	5.7297	5.7299	0.0791	8.3262	0.0319	0.0183	0.0454	0.0323	0.0322	0.0242	0.0413
A	5	5.6273	0.2748	9.3497	5.6282	5.6278	0.4651	8.2486	0.0319	0.0193	0.0450	0.0321	0.0323	0.0245	0.0412
B	1	5.6181	0.0000	9.1908	5.6190	5.6190	0.0000	8.0897	0.0327	0.0189	0.0459	0.0328	0.0325	0.0250	0.0415
B	2	7.8160	0.0000	33.0908	7.8169	7.8162	0.0000	31.9897	0.0328	0.0166	0.0483	0.0328	0.0330	0.0231	0.0429
B	3	7.9877	0.0000	28.4876	7.9883	7.9890	0.0000	27.3872	0.0328	0.0169	0.0473	0.0333	0.0327	0.0235	0.0428
B	4	7.4394	0.0000	17.7743	7.4401	7.4404	0.0000	16.6734	0.0329	0.0181	0.0469	0.0330	0.0325	0.0241	0.0424
B	5	6.6682	0.0000	11.3201	6.6683	6.6687	0.0000	9.0188	0.0329	0.0187	0.0463	0.0335	0.0326	0.0245	0.0422
C	1	5.8680	0.0000	9.4336	5.8687	5.8693	0.0000	8.3327	0.0332	0.0185	0.0460	0.0337	0.0332	0.0249	0.0420
C	2	8.6780	0.0000	46.1005	8.6781	8.6786	0.0000	45.0000	0.0335	0.0164	0.0489	0.0340	0.0336	0.0229	0.0439
C	3	7.4335	0.0000	38.6954	7.4337	7.4343	0.0000	37.5953	0.0338	0.0165	0.0484	0.0343	0.0329	0.0230	0.0435
C	4	7.0232	0.0000	22.1020	7.0233	7.0242	0.0000	21.0012	0.0342	0.0178	0.0474	0.0347	0.0337	0.0239	0.0427
C	5	6.4041	0.0000	12.6524	6.4044	6.4051	0.0000	9.5510	0.0343	0.0184	0.0466	0.0345	0.0338	0.0244	0.0425
D	1	7.0402	0.0000	21.8707	7.0411	7.0415	0.0000	20.7702	0.0325	0.0183	0.0465	0.0327	0.0326	0.0247	0.0423
D	2	10.7939	0.0000	64.2913	10.7939	10.7942	0.0000	63.1908	0.0339	0.0159	0.0511	0.0342	0.0335	0.0225	0.0443
D	3	13.6243	0.0000	47.2660	13.6250	13.6248	0.0000	46.1648	0.0327	0.0161	0.0506	0.0328	0.0328	0.0229	0.0437
D	4	9.4270	0.0000	23.0812	9.4277	9.4276	0.0000	19.9800	0.0324	0.0174	0.0483	0.0325	0.0326	0.0238	0.0431
D	5	8.4050	0.0000	13.7972	8.4050	8.4063	0.0000	12.6967	0.0329	0.0182	0.0474	0.0331	0.0334	0.0243	0.0426
Complete		5.2525	1.2716	9.2933	—	5.2428	1.572	9.3819	0.0322	0.0198	0.0449	—	0.0319	0.0256	0.0427

Table 3: Simulated bias (MSE) of estimates of parameters, reliability and hazard rate for fixed $\alpha = 2.5$, $\theta = 2$ and inspection time 0.2(0.2)1.6.

n	Dropping Scheme	α_{ML}	θ_{ML}	α_{EM}	θ_{EM}	α_B	θ_B	$S_{ML}(t=1)$ [a]	$H_{ML}(t=1)$ [b]
20	0 [c]	0.5720(1.1609)	0.2027(0.5126)	–	–	0.4853(0.7594)	0.0697(0.2792)	0.0056(0.0069)	0.3079(0.1871)
	1	0.5936(1.3329)	0.1844(0.5697)	0.5936(1.3329)	0.1844(0.5697)	0.5866(0.9275)	0.0389(0.3441)	0.0057(0.0069)	0.3090(0.1758)
	2	2.0742(4.2195)	0.3974(0.8522)	2.0742(4.2195)	0.3974(0.8522)	2.0920(3.7513)	0.2711(0.6740)	0.0065(0.0089)	0.4004(0.3915)
	3	1.7535(3.1504)	0.3918(0.8760)	1.7535(3.1504)	0.3918(0.8761)	1.6910(2.6922)	0.2595(0.5982)	0.0059(0.0077)	0.3097(0.3841)
	4	1.7493(3.1037)	0.3863(0.8205)	1.7493(3.1037)	0.3863(0.8205)	1.4366(2.6659)	0.1936(0.5589)	0.0058(0.0073)	0.3094(0.3770)
	5	1.0187(1.8215)	0.2976(0.7256)	1.0187(1.8215)	0.2976(0.7256)	0.8828(1.4053)	0.1024(0.4667)	0.0057(0.0072)	0.3093(0.2853)
30	0	0.3437(0.803)	0.1146(0.4054)	–	–	0.0995(0.5193)	0.0155(0.2524)	0.0048(0.0047)	0.3076(0.1023)
	1	0.343(0.8694)	0.1005(0.4128)	0.343(0.8694)	0.1005(0.4128)	0.1555(0.6105)	0.0093(0.2257)	0.0048(0.0046)	0.3083(0.1024)
	2	0.569(1.3157)	0.2074(0.6575)	0.569(1.3157)	0.2074(0.6575)	0.3678(1.0634)	0.0814(0.483)	0.0059(0.0059)	0.4005(0.1033)
	3	0.5598(1.2536)	0.1693(0.6293)	0.5194(1.2536)	0.1693(0.6293)	0.3014(1.0399)	0.0551(0.4667)	0.0054(0.0054)	0.3091(0.1028)
	4	0.5194(1.2116)	0.1617(0.5704)	0.5598(1.2116)	0.1617(0.5704)	0.3984(0.9592)	0.0304(0.4672)	0.0052(0.0053)	0.3091(0.1026)
	5	0.4323(1.0941)	0.1377(0.5208)	0.4323(1.094)	0.1377(0.5208)	0.2674(0.8894)	0.0013(0.3746)	0.0050(0.0050)	0.3084(0.1025)
40	0	0.2522(0.7113)	0.0954(0.3489)	–	–	0.091(0.5347)	0.0044(0.244)	0.0043(0.0034)	0.3073(0.0962)
	1	0.2533(0.8041)	0.0976(0.3893)	0.2533(0.8041)	0.0776(0.3893)	0.0881(0.673)	0.0194(0.2539)	0.0044(0.0035)	0.3076(0.0962)
	2	0.4759(1.0462)	0.1162(0.7111)	0.4138(1.0462)	0.1556(0.7112)	0.2162(0.9159)	0.076(0.6126)	0.0056(0.0047)	0.3998(0.0973)
	3	0.4138(1.0445)	0.1556(0.7052)	0.4759(1.0448)	0.1555(0.7049)	0.3396(0.8938)	0.0388(0.3861)	0.0053(0.0040)	0.3086(0.0968)
	4	0.3454(0.9456)	0.1801(0.5425)	0.3454(0.9455)	0.1806(0.5426)	0.1756(0.8132)	0.0332(0.3189)	0.0049(0.0039)	0.3085(0.0966)
	5	0.3313(0.8899)	0.1106(0.4869)	0.3313(0.8899)	0.1101(0.4869)	0.1577(0.6952)	0.0311(0.5889)	0.0049(0.0037)	0.3083(0.0964)
50	0	0.2481(0.7113)	0.0950(0.3488)	–	–	0.0336(0.645)	0.0335(0.1833)	0.0040(0.0029)	0.3066(0.0856)
	1	0.2532(0.8041)	0.0967(0.3893)	0.2532(0.8041)	0.0776(0.3893)	0.0076(0.6998)	0.0157(0.2412)	0.0041(0.0030)	0.3068(0.0858)
	2	0.3326(0.9059)	0.1098(0.6149)	0.3145(0.9059)	0.1098(0.6149)	0.0954(0.7606)	0.0466(0.4659)	0.0048(0.0043)	0.3072(0.0863)
	3	0.3145(0.8993)	0.1195(0.4954)	0.3326(0.8993)	0.1195(0.4954)	0.0351(0.8321)	0.0126(0.4803)	0.0046(0.0037)	0.3070(0.0861)
	4	0.3009(0.8623)	0.1017(0.4742)	0.3009(0.8623)	0.1017(0.4742)	0.0879(0.7338)	0.0022(0.411)	0.0043(0.0034)	0.3068(0.0859)
	5	0.2499(0.8065)	0.0883(0.4323)	0.2499(0.8065)	0.0883(0.4323)	0.0242(0.7548)	0.0109(0.3039)	0.0041(0.0030)	0.3068(0.0858)

[a] Here, true value of reliability at time 1 is $S(1) = 0.3048$
[b] True value of hazard rate at time 1 is $H(1) = 1.7851$
[c] 0 means complete case, when no dropping and data points collected continuously

Table 4: Simulated bias (MSE) of estimates of parameters, reliability and hazard rate for various choice of parameters and fixed $n = 30$

α	θ	Dropping Scheme	α_{ML}	θ_{ML}	α_{EM}	θ_{EM}	α_B	θ_B	$S_{ML}(t=1)$	$H_{ML}(t=1)$
0.5	0.5	0	0.0334(0.0145)	0.032(0.0107)	–	–	0.0271(0.0131)	0.0213(0.0103)	0.0003(0.0007)	0.0491(0.0217)
	0.5	1	0.0367(0.0146)	0.0326(0.0123)	0.0365(0.0146)	0.0325(0.0123)	0.0310(0.0143)	0.0243(0.0116)	0.006(0.0008)	0.0504(0.0219)
	0.5	4	0.0448(0.0150)	0.0335(0.0131)	0.0449(0.0152)	0.0336(0.013)	0.0400(0.0147)	0.0306(0.0128)	0.0063(0.0009)	0.0543(0.0223)
	1.5	0	0.0346(0.0146)	0.0908(0.0959)	–	–	0.0320(0.0129)	0.0747(0.0908)	0.0029(0.0024)	0.1578(0.2476)
	1.5	1	0.0416(0.0151)	0.0912(0.0965)	0.0414(0.015)	0.0911(0.0965)	0.0407(0.0137)	0.0756(0.0951)	0.0082(0.0026)	0.1623(0.2478)
	1.5	4	0.0469(0.0152)	0.098(0.0965)	0.0471(0.0151)	0.0979(0.0965)	0.0451(0.0151)	0.0773(0.0958)	0.0095(0.0027)	0.1570(0.2479)
	2.5	0	0.0449(0.0152)	0.1452(0.2645)	–	–	0.0433(0.0150)	0.1305(0.2532)	0.0036(0.0045)	0.2796(0.7332)
	2.5	1	0.0455(0.0155)	0.1449(0.2655)	0.0455(0.0158)	0.1449(0.2654)	0.0431(0.0153)	0.1329(0.2557)	0.0042(0.0046)	0.2762(0.7333)
	2.5	4	0.0483(0.0164)	0.1456(0.2663)	0.0482(0.0164)	0.1455(0.2663)	0.0451(0.0158)	0.1333(0.2560)	0.0077(0.0048)	0.2726(0.7336)
1.5	0.5	0	0.1645(0.2268)	0.0363(0.0134)	–	–	0.1298(0.2241)	0.0308(0.0080)	0.0024(0.0021)	0.0080(0.0075)
	0.5	1	0.1711(0.2277)	0.0433(0.0136)	0.1708(0.2271)	0.0433(0.0135)	0.1339(0.2256)	0.0320(0.0095)	0.0038(0.0024)	0.0115(0.0075)
	0.5	4	0.1811(0.2311)	0.0441(0.0139)	0.1809(0.2313)	0.0440(0.0138)	0.1384(0.2282)	0.0408(0.0122)	0.0074(0.0026)	0.0045(0.0078)
	1.5	0	0.1877(0.2310)	0.1039(0.1192)	–	–	0.1418(0.2245)	0.0869(0.1119)	0.0020(0.0043)	0.0867(0.1136)
	1.5	1	0.1909(0.2321)	0.1067(0.1196)	0.1908(0.2319)	0.1066(0.1195)	0.1456(0.2243)	0.0869(0.1140)	0.0167(0.0044)	0.0850(0.1138)
	1.5	4	0.1953(0.2383)	0.1097(0.1203)	0.1954(0.2383)	0.1097(0.1203)	0.1517(0.2319)	0.0888(0.1189)	0.0182(0.0045)	0.0887(0.1140)
	2.5	0	0.1888(0.2350)	0.1877(0.3341)	–	–	0.1455(0.2339)	0.1540(0.3241)	0.0013(0.0048)	0.1845(0.3633)
	2.5	1	0.1926(0.2358)	0.1915(0.3346)	0.1926(0.2358)	0.1915(0.3346)	0.1469(0.2354)	0.1579(0.3238)	0.0026(0.005)	0.1831(0.3636)
	2.5	4	0.1991(0.2387)	0.1983(0.3352)	0.1993(0.2387)	0.1982(0.3352)	0.1481(0.2371)	0.1588(0.3251)	0.0039(0.0052)	0.1905(0.3638)
2.5	0.5	0	0.3226(0.8566)	0.0647(0.0282)	–	–	0.2821(0.8472)	0.0363(0.0267)	0.0049(0.0017)	0.0089(0.0038)
	0.5	1	0.3289(0.8574)	0.0717(0.0287)	0.3288(0.8574)	0.0717(0.0287)	0.2820(0.8475)	0.0379(0.0269)	0.0203(0.0018)	0.0098(0.0041)
	0.5	4	0.3355(0.8583)	0.0784(0.0293)	0.3356(0.8582)	0.0783(0.0293)	0.2867(0.8478)	0.0410(0.0278)	0.0224(0.0021)	0.0118(0.0044)
	1.5	0	0.3473(0.8631)	0.1681(0.2519)	–	–	0.2919(0.8414)	0.1268(0.2468)	0.0003(0.0038)	0.0689(0.0737)
	1.5	1	0.3498(0.8641)	0.1698(0.2528)	0.3497(0.864)	0.1698(0.2527)	0.3014(0.8453)	0.1319(0.2485)	0.0019(0.0039)	0.0504(0.0740)
	1.5	4	0.3523(0.8645)	0.1701(0.2528)	0.3524(0.8643)	0.1701(0.2527)	0.3080(0.8458)	0.1349(0.2499)	0.0026(0.0043)	0.0613(0.0740)
	2.5	0	0.3912(0.9119)	0.2835(0.7001)	–	–	0.3151(0.9069)	0.2349(0.6932)	0.0045(0.0051)	0.1418(0.3134)
	2.5	1	0.3916(0.9124)	0.2828(0.7004)	0.3916(0.9124)	0.2828(0.7003)	0.3223(0.9105)	0.2397(0.6971)	0.0062(0.0053)	0.1360(0.3141)
	2.5	4	0.3990(0.9203)	0.2859(0.7012)	0.3988(0.9102)	0.2859(0.7012)	0.3311(0.9148)	0.2420(0.6970)	0.0114(0.0054)	0.1258(0.3144)

Table 5: Average Bayes estimates(MSE in brackets) and 95% HPD intervals based on simulated data by dropping scheme 1 for different choice of prior parameters

$g_1(\alpha)$	$g_2(\theta)$	$\hat{\alpha}_B$	HPD Interval	$\hat{\theta}_B$	HPD Interval
G(4,2)[a]	G(4,2)	0.0848(0.8063)	(3.6263, 2.1042)	0.0112(0.2760)	(1.3451, 2.6745)
G(4,2)	G(1,0.5)	0.0970(0.8342)	(3.7263, 2.0055)	0.0276(0.3158)	(1.3071, 2.7326)
G(4,2)	G(0.4,0.2)	0.1139(0.9134)	(3.7963, 1.8692)	0.0412(0.3564)	(1.0128, 2.9745)
G(1,0.5)	G(4,2)	0.1134(0.9075)	(3.7701, 1.9506)	0.0198(0.3023)	(1.2016, 2.8618)
G(1,0.5)	G(1,0.5)	0.1329(0.9275)	(3.9263, 1.8069)	0.0389(0.3441)	(1.1129, 2.9045)
G(1,0.5)	G(0.4,0.2)	0.1458(0.9324)	(3.9292, 1.7864)	0.0500(0.3674)	(0.9976, 3.0198)
G(0.4,0.2)	G(0.4,0.2)	0.1461(0.9453)	(4.5138, 1.7001)	0.0567(0.3623)	(0.9900, 3.1199)

[a] $G(a,b)$ denotes the gamma prior with shape parameter a and scale parameter b

4 Comparison of the Estimators

In this section, we have compared the performances of the various estimators on the basis of their bias and mean square error (MSE). It may be mentioned here that the exact expressions for the bias and mean square errors can not be obtained because estimators are not in closed form. Therefore, biases and MSEs are estimated on the basis of Monte-Carlo simulation study of 2000 samples. For this purpose, we generated specified number of observations from the distribution given in equation (1) for arbitrarily fixed value of the parameters under the specified inspection and dropping scheme and calculated different estimates of α and θ following the procedure as described in the previous sections. This process was repeated 2000 times to obtain the simulated biases and MSEs. We have computed the MLEs by using Newton-Raphson algorithm as well as the EM algorithm. The estimates of (α, θ) obtained through Newton-Raphson algorithm and EM algorithm are denoted as(α_{ML}, θ_{ML}) and (α_{EM}, θ_{EM}) respectively. It is noted that Newton-Raphson algorithm has convergence rate of 85%-90%, whereas the EM algorithm converges most of the times. We have reported the results omitting these cases where the algorithms do not converge. To simulate progressive Type-I interval censored sample from the considered distribution, we have used the algorithm given by [3, pp.200] after modifying step ④ as : Determine the number of droppings at j^{th} stage by generating R_j from $Bin(n-x-r-d_j, p_j)$.

It may be noted here that the MSE and bias of these estimators will depend on sample size n, values of α, θ and hyper parameters λ_1, λ_2, ν_1 and ν_2. We considered a number of values for sample size n; namely n = 20, 30, 40 and 50. For the choice of the hyper-parameters of the prior distribution, we have considered one set of values as $\lambda_1 = \lambda_2 = \nu_1 = \nu_2 = 0$ which reduces the prior as non-informative prior. For informative prior, the hyper parameters are chosen on the basis of the information possessed by the experimenter. In most of the cases experimenter can have the notion that what is the expected value of the parameter and can always associate a degree of belief in this value. In other words, the experimenter can specify the prior mean and prior variance for the parameters. The prior mean reflects the experimenter's belief about the parameter in the form of its expected value and prior variance reflects his confidence in this expected value. Keeping this point in mind, we have chosen the hyper-parameters in such a way that the prior mean is equal to true value of the parameter and belief in the prior mean is either strong or weak i.e, the prior variance is small or large respectively; for details see [18]. The bias of the estimates of parameters, reliability and hazard rate with corresponding MSEs have been calculated and the results are summarized in tables 3, 4 and 5.

Table 3 provides absolute bias and MSE of estimates of the parameters along with reliability and hazard rate at time $t = 1$ for $\alpha = 2.5$, $\theta = 2$ and inspection times $0.2(0.2)1.6$. It can be seen from the table that in general the bias and MSEs decrease as n increases in all the considered cases. It can also be seen that MSE of MLE is more than that of corresponding Bayes in all the cases but, the difference between the MSEs of Bayes and ML estimates decreases for increase in the value of n. It is noted here that bias of the estimates and MSEs under censoring scheme 1 are approximately equal to that of complete sample case (denoted as scheme 0) and smaller than those under other schemes. In most of the cases it is observed that the bias and MSE under dropping scheme 1 is least ,followed by scheme 5,4,3 and 2 sequentially. The bias and the MSE for ML estimates and EM estimates are found to be same in almost all the considered cases. Bias and MSE of the reliability estimate shows a similar trend as observed for the parameter estimates.

Table 4 provides the absolute bias and MSE of the various estimators for different choices of model parameters. It is worthwhile to mention here that we have noted above that as sample size increases the Bias and MSE decrease, therefore we have reported the result in this table for n=30 only. Similarly, we have also noted above that the among the considered dropping schemes, under scheme 1 the performance of the estimates are as good as complete sample case and better than all other schemes. Therefore, we has reported the results for the complete sample case and scheme 1 and scheme 4 only. It may be seen from the table that bias and MSE of all the considered estimates of α, θ, reliability $S_{ML}(t = 1)$ and hazard rate $H_{ML}(t = 1)$ increases as α increases or/and as θ increases. It is interesting to note that the bias and MSEs of all the estimates are less when the proportion of droppings are less. All the estimates under scheme 1 have more or less similar bias and MSE as that of obtained for complete case; but biases and MSEs of the estimates under scheme 4 are little higher than those of others. Bias and MSEs of Bayes estimates for the variation of different prior choice are presented in table 5 and we see that, as prior confidence in the guessed value increases the MSE decreases.

5 Conclusions

In the present piece of work, we have considered both classical and Bayesian analysis for the progressive type-I interval censored data, when the lifetime of the items follows exponentiated exponential distribution. The ML estimates do not have explicit forms. Newton-Raphson and EM algorithm has been proposed to be used to compute the MLEs and it is found that although both work quite well, the EM algorithm provides better convergence rate. Therefore, we may conclude that EM algorithm be used for finding the MLE in censored sample cases. The Bayes estimates under the squared error loss function also do not exist in explicit form. But, Bayes estimates can be routinely obtained through the use of MCMC technique considering the shape and scale parameters having independent gamma priors. On the basis of this study, we may conclude that the proposed estimation procedures under progressive type-I interval censoring with specific choice of scheme, can be easily implemented. It is seen above that the censoring scheme and dropping schemes has an effect on the performance of the estimators. Thus if it is possible, it is better to choose a scheme resulting to less number of droppings. However, in most of the practical situations the dropping scheme are not controllable. In such situations, the inspection plan should be so designed as to result to less number of droppings. However, under any scheme proposed method can be used to obtain the estimates.

We have not considered any covariates in this paper. But in practice often the covariates may be present. It will be interesting to develop statistical procedures for the estimation of the unknown parameters in presence of covariates. Further, we have considered dropping probabilities at each stages to be fixed, but in real life phenomena these may be random and a suitable model to capture this randomness can be developed. The work in this direction is under process.

Appendix-A

$$
\begin{aligned}
L_{\alpha\alpha} = & \sum_{i=1}^{k} d_i \frac{\left(\phi_i{}^{\alpha}(\ln\phi_i)^2 - \phi_{i-1}^{\alpha}(\ln\phi_{i-1})^2\right)}{\left(\phi_i^{\alpha} - \phi_{i-1}^{\alpha}\right)} \\
& - d_i \frac{\left(\phi_i{}^{\alpha}\ln\phi_i - \phi_{i-1}^{\alpha}\ln\phi_{i-1}\right)^2}{\left(\phi_i^{\alpha} - \phi_{i-1}^{\alpha}\right)^2} \\
& - r_i \frac{\phi_i{}^{\alpha}(\ln\phi_i)^2}{1-\phi_i{}^{\alpha}} - r_i \frac{\phi_i{}^{2\alpha}(\ln\phi_i)^2}{(1-\phi_i{}^{\alpha})^2} \\
L_{\theta\theta} = & \sum_{i=1}^{k} -d_i \frac{\left(\phi_i^{\alpha-1}\alpha\,\xi_i - \phi_{i-1}^{\alpha-1}\alpha\,\xi_{i-1}\right)^2}{\left(\phi_i^{\alpha} - \phi_{i-1}^{\alpha}\right)^2} \\
& + d_i\alpha \frac{\phi_i^{\alpha-2}\xi_i(\alpha\xi_i - \phi_i T_i - \xi_i) - \phi_{i-1}^{\alpha-2}\xi_{i-1}(\alpha\xi_{i-1} - \phi_{i-1}T_{i-1} - \xi_{i-1})}{\left(\phi_i^{\alpha} - \phi_{i-1}^{\alpha}\right)} \\
& - r_i \frac{\phi_i^{\alpha-2}\alpha\xi_i(\alpha\xi_i - \phi_i T_i - \xi_i)}{1-\phi_i^{\alpha}} - r_i \frac{\phi_i^{2\alpha}\alpha^2\xi_i{}^2}{\phi_i^2(1-\phi_i^{\alpha})^2} \\
L_{\theta\alpha} = L_{\alpha\theta} = & \sum_{i=1}^{k} d_i \frac{\left(\psi_i\phi_i^{\alpha}\ln\phi_i - \psi_{i-1}\phi_{i-1}^{\alpha}\ln\phi_{i-1}\right)}{\left(\phi_i^{\alpha} - \phi_{i-1}^{\alpha}\right)} \\
& - \frac{(\psi_i - \psi_{i-1})(\phi_i^{\alpha}\ln\phi_i - \phi_{i-1}^{\alpha}\ln\phi_{i-1})}{\left(\phi_i^{\alpha} - \phi_{i-1}^{\alpha}\right)^2} \\
& - r_i \frac{\phi_i^{\alpha}\ln(1-\phi_i)(1-\phi_i)^{\alpha} + \psi_i^2}{(1-\phi_i^{\alpha})^2}
\end{aligned}
$$

where, $\phi_i = \phi_i(\theta, T_i) = 1 - e^{-\theta T_i}$ and $\psi_i = \frac{d}{d\theta}\phi_i = T_i e^{-\theta T_i}$

Appendix-B

The E-step requires computing the following conditional expectations using numerical integration methods,

$$E_{1i} = E\left[\tau|\tau \in [\,T_{i-1}, T_i\,)\right]$$
$$E_{2i} = E\left[\tau|\tau \in [\,T_i, \infty\,)\right]$$
$$E_{3i} = E\left[\ln\left(1 - e^{-\theta^{(k)}\tau}\right)|\tau \in [\,T_{i-1}, T_i\,)\right]$$
$$E_{4i} = E\left[\ln\left(1 - e^{-\theta^{(k)}\tau}\right)|\tau \in [T_i, \infty)\right]$$
$$E_{5i} = E\left[\frac{\tau}{e^{\theta^{(k)}\tau} - 1}\,|\,\tau \in [T_{i-1}, T_i)\right]$$
$$E_{6i} = E\left[\frac{\tau}{e^{\theta^{(k)}\tau} - 1}\middle|\,\tau \in [\,T_i, \infty\,)\right]$$

We can find these expectations of a doubly truncated from the left at a and from the right at b with $0 < a < b \leq \infty$ are as

$$E\left[\tau|\tau \in [\,a, b\,)\right] = \frac{\int_a^b \tau f(\tau; \alpha^{(k)}, \theta^{(k)})d\tau}{F(b; \alpha^{(k)}, \theta^{(k)}) - F(a; \alpha^{(k)}, \theta^{(k)})}$$
$$E\left[\ln\left(1 - e^{-\theta^{(k)}\tau}\right)|\tau \in [\,a, b\,)\right] = \frac{\int_a^b \ln\left(1 - e^{-\theta^{(k)}\tau}\right) f(\tau; \alpha^{(k)}, \theta^{(k)})d\tau}{F(b; \alpha^{(k)}, \theta^{(k)}) - F(a; \alpha^{(k)}, \theta^{(k)})}$$
$$E\left[\frac{\tau}{e^{\theta^{(k)}\tau} - 1}|\tau \in [a, b)\right] = \frac{\int_a^b \frac{\tau}{e^{\theta^{(k)}\tau} - 1} f(\tau; \alpha^{(k)}, \theta^{(k)})d\tau}{F(b; \alpha^{(k)}, \theta^{(k)}) - F(a; \alpha^{(k)}, \theta^{(k)})}$$

References

[1] R Aggarwala. Progressive interval censoring: Some mathematical results with applications to inference. *Communications in Statistics - Theory and Methods*, 30(8-9):1921–1935, jul 2001.
[2] S K Ashour and W M Afify. Statistical analysis of exponentiated weibull family under type i progressive interval censoring with random removals. *Journal of Applied Sciences Research*, 3(12):1851–1863, 2007.
[3] N Balakrishnan and E Cramer. *The Art of Progressive Censoring*. Springer New York, 2014.
[4] J O Berger and D Sun. Bayesian analysis for the poly-weibull distribution. *Journal of the American Statistical Association*, 88:1412–1418, 1993.
[5] James O Berger. *Statistical decision theory and Bayesian analysis*. Springer Science & Business Media, 2013.
[6] P P Carbone, L E Kellerhouse, and E A Gehan. Plasmacytic myeloma: A study of the relationship of survival to various clinical manifestations and anomalous protein type in 112 patients. *The American Journal of Medicine*, 42(6):937 – 948, 1967.
[7] D G Chen and Y L Lio. Parameter estimations for generalized exponential distribution under progressive type-i interval censoring. *Computational Statistics and Data Analysis*, 54(6):1581–1591, 2010.
[8] A P Dempster, N M Laird, and D B Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal statistical Society*, 39(1):1–38, 1977.
[9] A Gelmen, J Carlin, H Stern, and D Rubin. *Bayesian Data Analysis*. Text in Statistical Science, Chapman & Hall, 1995.
[10] B Gompertz. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society London*, 115:513–585, 1825.
[11] R D Gupta and D Kundu. Exponentiated Exponential Family : An Alternative to Gamma and Weibull Distributions. *Biometrical journal*, 43:117 – 130, 2001.
[12] R D Gupta and D Kundu. Generalized exponential distribution : Different method of estimations. *Journal of Statistical Computations and Simulations*, 69:315–337, 2001.
[13] J F Lawless. *Statistical Models and Methods for Lifetime Data*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, USA, 2002.
[14] I Natzoufras. *Bayesian Modeling Using WinBugs*. Series B, Wiley, 2009.
[15] M Z Raquab and M T Madi. Bayesian inference for the generalized exponential distribution. *Journal of Statistical computation and Simulation*, 75(10):841 – 852, 2005.

[16] C P Robert. The metropolishastings algorithm. *arXiv:1504.01896v1 [stat.CO]*, 2015.
[17] R Singh, S K Singh, U Singh, and G Prakash. Bayes Estimator of Generalized Exponential Parameters under LINEX Loss Function using Lindley's Approximation. *Data Science Journal*, 7:65–75, 2008.
[18] S K Singh, U Singh, and D Kumar. Estimation of parameters and reliability function of exponentiated exponential distribution: Bayesian approach under general entropy loss function. *Pakistan Journal of Statistics and Operational Research*, VII(2):199–216, 2011.
[19] S K Singh, U Singh, and Manoj Kumar. Estimation of parameters of exponentiated pareto model for progressive type-ii censored data with binomial removals using markov chain monte carlo method. *International Journal of Mathematics and Computation*, 21(4):88–102, 2013.
[20] S K Tse, C Yang, and H K Yuen. Statistical analysis of weibull distributed lifetime data under type ii progressive censoring with binomial removals. *Journal of Applied Statistics*, 27(8):1033–1043, November 2000.
[21] H K Yuen and S K Tse. Parameters estimation for weibull distributed lifetimes under progressive censoring with random removeals. *Journal of Statistical Computation and Simulation*, 55(1-2):57–71, September 1996.

Arun Kaushik is a CSIR senior research fellow in Department of Statistics, Institute of Science, Banaras Hindu University, India. He has published many peer-review research articles in the field of Statistical Inference, Bayesian Inference and Optimised censored plan.