

# On estimating flexible weibull parameters with Type I progressive interval censoring with random removal using data of cancerous tumors in blood

## Abstract

In this paper, the maximum likelihood and the Bayes estimators of the two unknown parameters of the flexible Weibull distribution have been obtained for progressive Interval type-I censoring scheme with binomial random removal. Point estimation and confidence intervals based on maximum likelihood and bootstrap method are also proposed. A Bayesian approach using Markov chain Monte Carlo (MCMC) method to generate from the posterior distributions and in turn computing the Bayes estimators are developed. To illustrate the proposed methods will discuss an example with the real data. Finally, comparing the two techniques through comparisons between the maximum likelihood using bootstrap method and different Bayes estimators using MCMC study.

**Keywords:** flexible weibull distribution, progressive interval type-I censoring, random removal, percentile bootstrap, bayesian and non-bayesian approach, markov chain monte carlo (MCMC)

Volume 4 Issue 5 - 2016

## Afify WM

Department of Head of Statistics, Mathematics & Insurance, Kafr El-sheikh University, Egypt

**Correspondence:** Afify WM, Department of Head of Statistics, Mathematics & Insurance, Kafr El-sheikh University, Faculty of Commerce, Egypt, Email waleedafify@yahoo.com

**Received:** September 09, 2016 | **Published:** October 14, 2016

## Introduction

Censoring is very common in life tests in the past several decades; the experimenter may be unable to obtain complete information on failure times of all experimental items. For this reason, Aggarwalla<sup>1</sup> suggested a useful type of censoring, namely, a progressively Type I interval censored data, which is a union of Type I interval and progressive censoring. This method of lifetime data collection can be useful to a biological experimenter, particularly when the experimental units are humans, as continuous monitoring is often not possible to implement, and withdrawal rates from such studies may high.

In progressive censored the number of units being removed from the test at each failure time may occur at random. For example; the number of patients who drop out of clinical test at each stage is random and cannot be predetermined. That is why to display a more general censoring scheme called progressive progressively Type I interval censored with random removal. It can be described as follows: suppose  $n$  units are put on life test at time  $T_0=0$  and under inspection at  $m$  pre-specified times  $T_1 < T_2 < \dots < T_m$  where  $T_m$  is scheduled time to terminate the experiment. The number,  $k_i$ , of failures within  $(T_{i-1}, T_i]$  is recorded and  $r_i$  surviving items are randomly removed from the life testing at the  $i^{\text{th}}$  inspection time,  $T_i$ , for  $i = 1, 2, \dots, m$ . Since the number,  $r_i$ , of surviving items is a random variable and exact number of items with drawn should not be greater than  $r_i$  at time schedule  $T_i$ ,  $r_i$ 's are random. Such a censoring mechanism is termed as progressive interval type-I censoring with random removal scheme. If we assume that probability of removal of a unit at every stage is  $\pi$  for each unit then  $r_i$  can be considered to follow a binomial distribution i.e.,  $r_i \approx B(n-m-\sum_{j=0}^{i-1} r_j, \pi)$  for  $i = 1, 2, \dots, m$ . The main difference between progressive interval type I censoring with fixed removal and progressive interval type I censoring with random removals is that the removals are predetermined in the former case while they are random in the latter case. Note that  $m$  is pre-determined in both cases. However, many practical applications suggest that it is more flexible to have removals random to accommodate the unexpected drop out of experimental subjects.

Although progressive censoring occurs frequently in many applications, there are relatively few works on it. Some early works can be found in Cohen,<sup>2</sup> Readers can refer to the book Balakrishnan & Aggarwala<sup>3</sup> for more details on the methods and applications of this topic. However, all these works assumed that the number of units being removed from the test is fixed in advance. In practice, it is impossible to pre-determine the removal pattern. Thus, Yuen & Tse<sup>4</sup> and Yang et al.,<sup>5</sup> considered the estimation problem when lifetimes collected under a Type II progressive censoring with random removals and Kendell & Anderson<sup>6</sup> point out that the expected duration under grouped data. Progressive type-I interval censored sampling is an important practical problem that has received considerable attention in the past several years. Based on the progressive type-I interval censored sampling, Ashour & Afify<sup>7</sup> derived the maximum likelihood estimators of parameters of the exponentiated Weibull family and their asymptotic variances under random removal. Lin et al.,<sup>8</sup> determined optimally spaced inspection times for the log-normal distribution, while Ng & Wang<sup>9</sup> and Chen & Lio<sup>10</sup> compared three classical estimation methods, the maximum likelihood estimators the moment method and the probability plot method in terms of the Weibull distribution and generalized exponential respectively.

In Bayesian approach, It is too difficult to find integrate over the posterior distribution and the problem is that the integrals are usually impossible to evaluate analytically. But in MCMC technique, the MCMC methodology provided a convenient and efficient way to sample from complex, high-dimensional statistical distributions. Recently, application of the MCMC method to the estimation of parameters or some other vital properties about statistical models is very common. Green et al.,<sup>11</sup> using the MCMC method for estimating the three parameters Weibull distribution, and they showed that the MCMC method is better than the ML method, when given a proper prior distribution of the parameters. As a generalization of the two parameter Weibull model, Gupta et al.,<sup>12</sup> gave a complete Bayesian analysis of the Weibull extension model using MCMC simulation and complete sample. Lin & Lio<sup>13</sup> discussed Bayesian inference under

progressive type I interval censoring by using MCMC.

A random variable is said to have a Flexible Weibull Distribution with parameters  $\lambda, \beta > 0$  if its probability density function, cumulative function, survival function and hazard function are given by

$$f(x; \lambda, \beta) = \left( \lambda + \frac{\beta}{x^2} \right) e^{\lambda x - \frac{\beta}{x}} e^{-\lambda x - \frac{\beta}{x}} \quad (1)$$

$$F(x; \lambda, \beta) = 1 - e^{-\lambda x - \frac{\beta}{x}} \quad (2)$$

$$\bar{F}(x; \lambda, \beta) = e^{-\lambda x - \frac{\beta}{x}} \quad (3)$$

In this paper we consider the Bayesian inference of the scale parameters for progressive interval type-I censored data when both parameters are unknown. We assumed that the both scale parameters  $\lambda$  and  $\beta$  have gamma prior and they are independently distributed. As expected in this case also, the Bayes estimates cannot be obtained in closed form. We propose to use the Gibbs sampling procedure to generate MCMC samples, and then using the Metropolis–Hastings algorithms, we obtain the Bayes estimates of the unknown parameters. We perform some simulation experiments to see the behavior of the proposed Bayes estimators and compare their performances with the maximum likelihood estimators.

The rest of the paper is organized as follows. In the next section, the ML estimators of the unknown parameters and approximate confidence intervals are presented. The corresponding parametric bootstrap confidence intervals for the parameters are given in Section 3. In Section 4, we cover Bayes estimates and construction of credible intervals using the MCMC techniques. In Section 5, for illustrative purposes, we performed a real data analysis. Comparisons among estimators are investigated through Monte Carlo simulations in Section 6. Finally, conclusions appear in Section 7.

## Classical estimation and percentile bootstrap algorithm (Boot-p)

Classical estimation (maximum likelihood estimators) of the unknown parameters and approximate confidence intervals are presented. Also, the corresponding parametric bootstrap confidence intervals using percentile bootstrap Algorithm (Boot-p) for the parameters are given in this section.

### Classical estimation

Suppose a progressively Type-I interval censored sample is collected as described above, beginning with a random sample of units with a continuous lifetime distribution  $F(x)$  and let  $k_1, k_2, \dots, k_m$  denote the number of units known to have failed in the intervals  $(0, T_1], (T_1, T_2], \dots, (T_{m-1}, T_m]$ , respectively. Then, based on this observed data, the joint likelihood function will be Aggarwala.<sup>1</sup>

$$L_1(X; \lambda, \beta | R) = C \prod_{i=1}^m [F(T_i; \lambda, \beta) - F(T_{i-1}; \lambda, \beta)]^{k_i} [1 - F(T_i; \lambda, \beta)]^{r_i} \quad (4)$$

Where  $C$  is constant. Clearly, if  $r_i = 0$  for  $i=1, 2, \dots, m-1$  and  $r_m = n-k$  equation (4) reduces to the likelihood function for interval type I censoring data is defined as follows:

$$L(X; \lambda, \beta) = C (1 - F(T_m; \lambda, \beta))^{n-k} \prod_{i=1}^m (F(T_i; \lambda, \beta) - F(T_{i-1}; \lambda, \beta))^{k_i}$$

Where  $k = \sum_{i=1}^m k_i$  and  $k_1, k_2, \dots, k_m$  are the number of units known to have failed in the intervals  $(0, T_1], (T_1, T_2], \dots, (T_{m-1}, T_m]$ , respectively.

For type I progressive Interval censoring, supposed that  $r_i$  is independent of  $X_i$  for all  $i$ ; Wu & Chang<sup>14</sup> suggested the following likelihood function of a progressive interval censoring with binomial removals

$$L(X, R; \lambda, \beta) = L_1(X; \lambda, \beta | R=r) \times P(R) \quad (5)$$

Where  $L_1(X; \theta | R=r)$  is the likelihood function for a progressive type I interval censored with fixed removal (4) and  $P(R)$  will be

$$P(R) = P(R_{m-1} = r_{m-1} \setminus R_{m-2} = r_{m-2}, \dots, R_1 = r_1) P(R_{m-2} = r_{m-2} \setminus R_{m-3} = r_{m-3}, \dots, R_1 = r_1) \dots P(R_2 = r_2 \setminus R_1 = r_1) P(R_1 = r_1) \dots P(R_1 = r_1)$$

Such as

$$P(R) = \frac{(n-m)!}{\prod_{i=1}^m r_i! (n-m - \sum_{j=1}^{m-1} r_j)!} \pi^{\sum_{j=1}^{m-1} r_j} (1-\pi)^{(m-1)(n-m) - \sum_{j=1}^{m-1} (m-j)r_j} \quad (6)$$

and  $f(\cdot), F(\cdot)$  are the same as defined before in (1) and (2) respectively. The log likelihood function with random removal can be written as

$$\log L(X, R; \lambda, \beta) = \log C + \sum_{i=1}^m k_i \log [F(T_i) - F(T_{i-1})] + \sum_{i=1}^m r_i \log [1 - F(T_i)] + \sum_{j=1}^{m-1} r_j \log \pi + (m-1)(n-m) - \sum_{j=1}^{m-1} (m-j)r_j \log (1-\pi) \quad (7)$$

The maximum likelihood estimations of  $\lambda$  and  $\beta$  are the simultaneous solutions of following normal equations

$$\sum_{i=1}^m k_i \frac{\frac{\partial F(T_i)}{\partial \lambda} \frac{\partial F(T_{i-1})}{\partial \lambda}}{F(T_i) - F(T_{i-1})} + \sum_{i=1}^m r_i \frac{\frac{\partial}{\partial \lambda} [1 - F(T_i)]}{[1 - F(T_i)]} = 0 \quad (8)$$

$$\sum_{i=1}^m k_i \frac{\frac{\partial F(T_i)}{\partial \beta} \frac{\partial F(T_{i-1})}{\partial \beta}}{F(T_i) - F(T_{i-1})} + \sum_{i=1}^m r_i \frac{\frac{\partial}{\partial \beta} [1 - F(T_i)]}{[1 - F(T_i)]} = 0 \quad (9)$$

Note that  $P(R)$  does not involve the parameters. Therefore, the MLE  $\hat{\pi}$  of  $\pi$  can be found by maximizing  $P(R)$  directly, that is,

$$\hat{\pi} = \frac{1}{\sum_{j=1}^{m-1} r_j} \left( (m-1)(n-m) - \sum_{j=1}^{m-1} (m-j)r_j \right) = 0$$

Therefore, the maximum likelihood estimation of parameter  $\hat{\pi}$  is given by

$$\hat{\pi} = \frac{\sum_{j=1}^{m-1} r_j}{(m-1)(n-m) - \sum_{j=1}^{m-1} (m-j)r_j} \quad (10)$$

It may be noted that (9) and (10) cannot be solved simultaneously

to provide a nicely closed form for the estimators. Therefore, we propose to use fixed point iteration method for solving these equations. Using Fisher information matrix  $I(\hat{\lambda}, \hat{\beta}, \hat{\pi})$  in the Appendix and the asymptotic normality of the maximum likelihood estimators can be used to compute the approximate confidence intervals (ACI) for parameters  $\lambda$ ,  $\beta$  and  $\pi$ . Therefore,  $(1-\gamma)100\%$  confidence intervals for parameters  $\lambda$ ,  $\beta$  and  $\pi$  will be become

$$\hat{\lambda} \pm Z_{\gamma/2} \sqrt{\text{Var}(\hat{\lambda})}, \hat{\beta} \pm Z_{\gamma/2} \sqrt{\text{Var}(\hat{\beta})} \text{ and } \hat{\pi} \pm Z_{\gamma/2} \sqrt{\text{Var}(\hat{\pi})}$$

Where  $Z_{\gamma/2}$  is percentile of the standard normal distribution with right-tail probability  $\gamma/2$ .

### Data algorithm

The data generation is based on the algorithm proposed by Aggarwala<sup>1</sup> to simulate the numbers,  $k_i$  of failed items in each subinterval  $(T_{i-1}, T_i]$ ,  $i = 1, \dots, m$ , from an initial sample of size  $n$  putting on life testing at time 0. This algorithm, which is an extension from the procedure developed by Kemp & Kemp<sup>15</sup> for the multinomial distribution, involves generating  $m$  binomial random variables. A procedure to generate a progressively type I interval censored data with random removal,  $(k_i, r_i, T_i)$ ,  $i = 1, \dots, m$ , from the flexible Weibull distribution can be described as follows briefly:  $k_0 = 0$  let and  $r_0 = 0$  and for  $i = 1, \dots, m$ ,

**Step 1:** set  $i = 0$  and let  $k \text{ sum} = r \text{ sum} = 0$ .

**Step 2:**  $i = i + 1$

- Using initial  $\pi$  to generate a sample  $R = r_i$ ,  $i = 1, \dots, m$  using binomial distribution, where  $r_1$  following the binomial  $(n-m, \pi)$  distribution and the variables  $r_i/r_1, r_2, \dots, r_{i-1}$  follow the binomial  $(n-m - \sum_{j=1}^{i-1} r_j, \pi)$  distribution for  $i = 2, 3, \dots, m-1$ .

$$\text{Set } r_m = \begin{cases} n-m - \sum_{j=1}^{m-1} r_j & \text{if } n-m - \sum_{j=1}^{m-1} r_j > 0 \\ 0 & \text{otherwise} \end{cases}$$

- Generate  $k_i$  as a binomial random variable with parameters  $n \text{ k sum} \text{ r sum}$  and

$$p = \left( e^{-e^{\lambda T_{i-1} - \frac{\beta}{T_{i-1}}}} - e^{-e^{\lambda T_i - \frac{\beta}{T_i}}} \right) / \left( 1 - e^{-e^{\lambda T_{i-1} - \frac{\beta}{T_{i-1}}}} \right)$$

**Step 3:** Set  $k \text{ sum} = k \text{ sum} + k_i$  and  $r \text{ sum} = r \text{ sum} + r_i$ .

**Step 4:** If  $i < m$ , go to step 2; otherwise, stop.

### Percentile bootstrap algorithm (Boot-p)

We can increase information about the population value more than does a point estimate by using a parametric bootstrap interval. We propose to use confidence intervals based on the parametric bootstrap methods using percentile bootstrap Algorithm (Boot-p) based on the idea of Efron.<sup>16</sup>

The algorithm for estimating the confidence intervals is illustrated as follows:

Before progressing further, we first describe how we generate

progressively interval Type I censored data with binomial random removals. The following algorithm is followed to obtain these samples.

- Specify the values of  $n, m, T$ .
- Specify the values of  $\lambda, \beta$  and  $\pi$ .
- Form data algorithm; compute the maximum likelihood estimates of the parameters  $\hat{\lambda}$ ,  $\hat{\beta}$  and  $\hat{\pi}$ , by solving the likelihood equations simultaneously in (8), (9) and (10).
- Use  $\hat{\lambda}$ ,  $\hat{\beta}$  and  $\hat{\pi}$ , to generate a bootstrap sample  $k^*$  with the same values of  $r_i, m$ ;  $i = 1, 2, \dots, m$  using algorithm presented in Balakrishnan & Sandhu.<sup>17</sup>
- As in step 3, based on  $k^*$  compute the bootstrap sample estimates of  $\hat{\lambda}$ ,  $\hat{\beta}$  and  $\hat{\pi}$ , say  $\hat{\lambda}^*$ ,  $\hat{\beta}^*$  and  $\hat{\pi}^*$ .
- Repeat steps 4-5  $B$  times representing  $B$  bootstrap maximum likelihood estimators of  $\lambda$ ,  $\beta$  and  $\pi$  based on  $B$  different bootstrap samples.
- Arrange all  $\hat{\lambda}^*$ ,  $\hat{\beta}^*$  and  $\hat{\pi}^*$ , in an ascending order to obtain the bootstrap sample  $(\varphi_1^{[1]}, \varphi_1^{[2]}, \dots, \varphi_1^{[B]})$ ,  $i = 1, 2, 3$  (where  $\varphi_1 \equiv \hat{\lambda}^*$ ,  $\varphi_2 \equiv \hat{\beta}^*$  and  $\varphi_3 \equiv \hat{\pi}^*$ ).

Let  $G(z) = P(\varphi_1 \leq z)$  be the cumulative distribution function of  $\varphi_1$ .

Define  $\varphi_{boot} = G^{-1}(z)$  for given  $z$ . The approximate bootstrap  $100(1-2\gamma)\%$  confidence interval (ABCI) of  $\pi$  is given by

### Bayesian estimation and MCMC technique

In this section, we will focus to Bayesian approach using Markov chain Monte Carlo (MCMC) method to generate from the posterior distributions and in turn computing the Bayes estimators are developed.

#### Bayesian estimation

In Bayesian scenario, we need to assume the prior distribution of the unknown model parameters to take into account uncertainty of the parameters. The informative prior densities for  $\lambda$  and  $\beta$  are given as

$$g_1(\lambda) \propto \lambda^{b-1} e^{-\lambda a}, a, b, \lambda > 0,$$

$$g_2(\beta) \propto \beta^{d-1} e^{-\beta c}, c, d, \beta > 0,$$

and  $\pi$  has a

$$g_3(\pi) \propto \pi^{A-1} (1-\pi)^{B-1}, 0 < \pi < 1; A, B > 0$$

Note that the parameters  $\lambda$ ,  $\beta$  and  $\pi$  behave as independent random variables. The joint informative prior probability density function of  $\lambda$ ,  $\beta$  and  $\pi$  is

$$g(\lambda, \beta, \pi) \propto g_1(\lambda) \times g_2(\beta) \times g_3(\pi) \\ (\lambda, \beta, \pi) \propto \lambda^{b-1} e^{-\lambda a} \beta^{d-1} e^{-\beta c} \pi^{A-1} (1-\pi)^{B-1} \quad (11)$$

where  $a, b, c, d, A$  and  $B$  are assumed to be known and are chosen to reflect prior knowledge about  $\lambda$ ,  $\beta$  and  $\pi$ . Note that when

$a = b = c = d = A = B = 0$ , (we call it prior 0) they are the non-informative  $\lambda$ ,  $\beta$  and  $\pi$  respectively.

It follows from (4), (6) and (11) that the joint posterior density function of  $\lambda$ ,  $\beta$  and  $\pi$  given  $x$  is thus

$$\pi^*(\lambda, \beta, \pi) \propto \prod_{i=1}^m \left[ 1 - \frac{e^{-u(T_i)}}{e^{-u(T_{i-1})}} \right]^{k_i} e^{-\{k_i u(T_{i-1}) + r_i u(T_i)\}} \pi^{\sum_{j=1}^{m-1} r_j} \cdot (1-\pi)^{(m-1)(n-m) - \sum_{j=1}^{m-1} (m-j)r_j} \times \lambda^{b-1} e^{-\lambda a} \beta^{d-1} e^{-\beta c} \pi^{A-1} (1-\pi)^{B-1} \quad (12)$$

where  $u(T_i) = e^{\frac{\lambda T_i - \beta}{T_i}}$  and  $u(T_{i-1}) = e^{\frac{\lambda T_{i-1} - \beta}{T_{i-1}}}$ .

It is not possible to compute (12) analytically. The problem is that the integrals in (12) are usually impossible to evaluate analytically, and the numerical methods may fail. The MCMC method provides an alternative method for parameter estimation. In the following subsections, we propose using the MCMC technique to obtain Bayes estimates of the unknown parameters and construct the corresponding credible intervals.

## MCMC technique

Computer simulation of Markov chains in the space of parameter will depend on Markov chain Monte Carlo (MCMC) Gilks et al.<sup>18</sup> The Markov chains are defined in such a way that the posterior distribution in the given statistical inference problem is the asymptotic distribution. However, the posterior likelihood usually does not have a closed form for a given progressively type-I interval-censored data. Moreover, a numerical integration cannot be easily applied in this situation. A lot of standard approaches to display like Markov chains exist, including Gibbs sampling, Metropolis-Hastings (M-H) and reversible jump. The M-H algorithm is a very general MCMC method first expansion by Metropolis et al.<sup>19</sup> and later extended by Hastings.<sup>20</sup> it is possible to use these algorithms by implement posterior simulation in essentially any problem which allow point wise evaluation of the prior distribution and likelihood function. It can be used to obtain random samples from any arbitrarily complicated target distribution of any dimension that is known up to a normalizing constant. In fact, Gibbs sampler is just a special case of the M-H algorithm.

In order to use the method of MCMC for estimating the parameters of the flexible Weibull distribution and random removal, namely,  $\lambda$ ,  $\beta$  and  $\pi$ . Let us consider independent priors as in (10), the full conditional distribution for any parameter can be obtained, to within a constant, by factoring out from the likelihood function  $L(X, R; \lambda, \beta)$  any terms containing the relevant parameter and multiplying by its prior. From (11), the full posterior conditional distribution for is proportional to

$$\pi^*(\lambda/x, \beta, \pi) \propto \prod_{i=1}^m \left[ 1 - \frac{u(T_i)}{u(T_{i-1})} \right]^{k_i} e^{-\{k_i u(T_{i-1}) + r_i u(T_i)\}} \times \lambda^{b-1} e^{-\lambda a} \quad (12)$$

Also, the full posterior conditional distribution for  $\beta$  is proportional to

$$\pi^*(\beta/x, \lambda, \pi) \propto \prod_{i=1}^m \left[ 1 - \frac{u(T_i)}{u(T_{i-1})} \right]^{k_i} e^{-\{k_i u(T_{i-1}) + r_i u(T_i)\}} \times \beta^{d-1} e^{-\beta c} \quad (13)$$

$$\pi^*(\lambda, \beta, \pi) \propto \frac{L_1(X; \lambda, \beta / R = r) \times P(R) \times g(\lambda, \beta, \pi)}{\int_0^\infty \int_0^\infty \int_0^1 L_1(X; \lambda, \beta / R = r) \times P(R) g(\lambda, \beta, \pi) d\lambda d\beta d\pi}$$

Similarly, the marginal posterior density of  $\pi$  is proportional to

$$\pi^*(\pi/x, \lambda, \beta) \propto \pi^{A + \sum_{j=1}^{m-1} r_j - 1} (1-\pi)^{B + (m-1)(n-m) - \sum_{j=1}^{m-1} (m-j)r_j - 1} \quad (14)$$

It is noted that the posterior distribution of  $\pi$  is beta with parameters  $A^*$  and  $B^*$  where  $A^* = A + \sum_{j=1}^{m-1} r_j$  and  $B^* = B + (m-1)(n-m) - \sum_{j=1}^{m-1} (m-j)r_j$  and, therefore, samples of  $\pi$  can be easily generated using any beta generating routine. But the conditional posterior distribution of  $\lambda$  and  $\beta$  equations (12) and (13) respectively, cannot be reduced analytically to well-known distributions and therefore it is not possible to sample directly by standard methods, but the plot of it show that it is similar to normal distribution. So to generate random numbers from this distribution, we use the M-H method with normal proposal distribution.

## MCMC process

Now, we propose the following scheme to generate  $\lambda$ ,  $\beta$  and  $\pi$  from density functions and in turn obtain the Bayes estimates and the corresponding credible intervals.

1. Start with an  $\lambda^{(0)} = \hat{\lambda}$ ,  $\beta^{(0)} = \hat{\beta}$  and  $M = \text{burn-in}$ .
2. Set  $t = 1$ .
3. Generate  $\pi^{(t)}$  from beta distribution  $\pi^*(\pi/x, \lambda, \beta)$ .
4. Using M-H algorithm Metropolis et al.<sup>19</sup>  $\lambda^{(t)}$  from  $\pi^*(\lambda/x, \beta, \pi)$  with the  $N(\lambda^{(t-1)}, \sigma_\lambda^2)$  proposal distribution where  $\sigma_\lambda^2$  is the variance of  $\lambda$  obtained using variance-covariance matrix; similarly,  $\beta^{(t)}$  from  $\pi^*(\beta/x, \lambda, \pi)$  with the  $N(\beta^{(t-1)}, \sigma_\beta^2)$  proposal distribution where  $\sigma_\beta^2$  is the variance of  $\beta$  obtained using variance-covariance matrix.
5. Compute  $\lambda^{(t)}$ ,  $\lambda^{(0)} = \hat{\lambda}$ ,  $\beta^{(0)} = \hat{\beta}$  and  $\pi^{(t)}$ .
6. Set  $t = t + 1$ .
7. Repeats Steps 3-6  $N$  times.
8. Obtain the Bayes estimates of  $\lambda$ ,  $\beta$  and  $\pi$  with respect to the squared error loss function as  $\hat{E}(\lambda/x) = \frac{1}{N-M} \sum_{i=M+1}^N \lambda_i$ ,  $\hat{E}(\beta/x) = \frac{1}{N-M} \sum_{i=M+1}^N \beta_i$  and  $\hat{E}(\pi/x) = \frac{1}{N-M} \sum_{i=M+1}^N \pi_i$

9. To compute the credible intervals  $\lambda$  of  $\beta$ , and  $\pi$ , order  $\lambda_1, \dots, \lambda_{N-M}, \beta_1, \dots, \beta_{N-M}$  and  $\pi_1, \dots, \pi_{N-M}$  as  $\lambda_1 < \dots < \lambda_{N-M}$ ,  $\beta_1 < \dots < \beta_{N-M}$ , and  $\pi_1 < \dots < \pi_{N-M}$ . Then the  $100(1-\gamma)\%$  symmetric credible intervals (SCI) of  $\lambda$ ,  $\beta$  and  $\pi$  become:

$$[\lambda_{(N-M)\gamma/2}, \lambda_{(N-M)(1-\gamma/2)}], [\beta_{(N-M)\gamma/2}, \beta_{(N-M)(1-\gamma/2)}]$$

and

$$[\pi_{(N-M)\gamma/2}, \pi_{(N-M)(1-\gamma/2)}]$$

## Real data analysis

To conduct a study within the Institute of Oncology in Tanta - Egypt. This study is concerned with the treatment of cancerous tumors in blood and studies their impact on the overall health of the patient. Underwent the study 228 patients and they had varying degrees of

$$T_0 = 0, T_1 = 16, T_2 = 31, T_3 = 46, T_4 = 61, T_5 = 76, T_6 = 91, T_7 = 106, T_8 = 121, T_9 = 136, T_{10} = 151, T_{11} = 166 \text{ and } T_{12} = 181.$$

The last inspection time,  $T_{12} = 181$ , is the scheduled time to terminate the experiment. The lifetime distribution is flexible Weibull with parameters  $\lambda$  and  $\beta$  where the simulation input parameters are selected close to the maximum likelihood estimators of flexible Weibull parameters for modeling the real data in (Table 1). The performance of parameter estimation under progressively type I interval censored with random removal  $\pi$  is compared via the maximum likelihood, bootstrap method and MCMC procedure developed in this paper. The summary for 1000 simulation runs is shown in (Tables 2-5).

Bayes estimates of  $\lambda$ ,  $\beta$  and  $\pi$  using MCMC method, we assume that informative priors  $a = 2, b = 3, c = 4, d = 2, A = 2$  and  $B = 3$  on  $\lambda$ ,  $\beta$  and  $\pi$  in (Table 4). Also, by non-informative prior using MCMC procedure with Bayes estimation will be obtained on estimates of parameters in (Table 5).

disease. Patients were examined every 15 days for 6 consecutive months. Of course there were cases of withdrawal (death - interruption of treatment for different reasons)

As we know on the basis of a single sample, one cannot make a general statement regarding the behavior of proposed estimators, therefore we present a simulation study for the study of the behavior of the estimators in the next section.

## Simulation

The simulation is conducted using the R version 3.2.2 (for more information about R programming, the reader may refer to this manual of R, version 3.3.0 under development (2015-10-30) Copyright 2000 –2015 R Core Team). The simulation setup is parallel to the real data given in (Table 1). To be specific, each replication of the simulation generates a progressively type-I interval-censored data within twelve subintervals which have pre-specified inspection times (in terms of half month),

Both of density functions of  $\pi^*(\lambda/x, \beta, \pi)$  and  $\pi^*(\beta/x, \lambda, \pi)$  can be approximated by normal distribution functions but density function of  $\pi^*(\pi/x, \lambda, \beta)$  will be beta as mentioned in subsection (3.3) which are plotted in (Figure 1& 2) Chain of MCMC outputs of  $\lambda$ ,  $\beta$  and  $\pi$ , using 100 000 MCMC samples. This was done with 1000 bootstrap sample and 100 000 MCMC sample and discard the first 50000 values as ‘burn-in’. The Bayes estimators can be seen to have the smaller risks than classical estimators for all the considered cases. It may also be noted that the Bayes estimators obtained under informative prior are more efficient than those obtained under non-informative priors. This indicates that the Bayesian procedure with accurate prior information provides more precise estimates. Also, The Length of the SCI (using informative prior) is smaller than the Length of the ACI and ABCI.

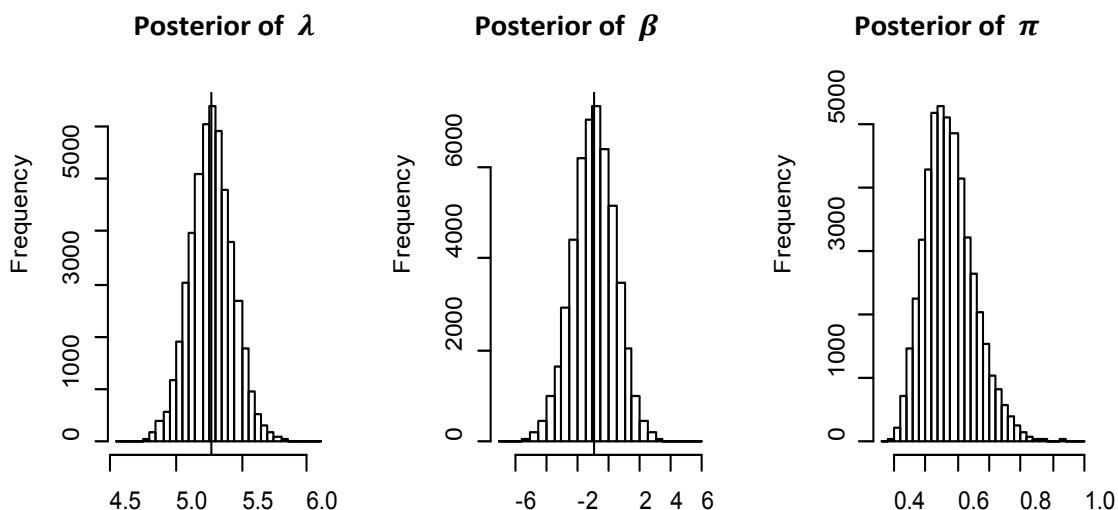
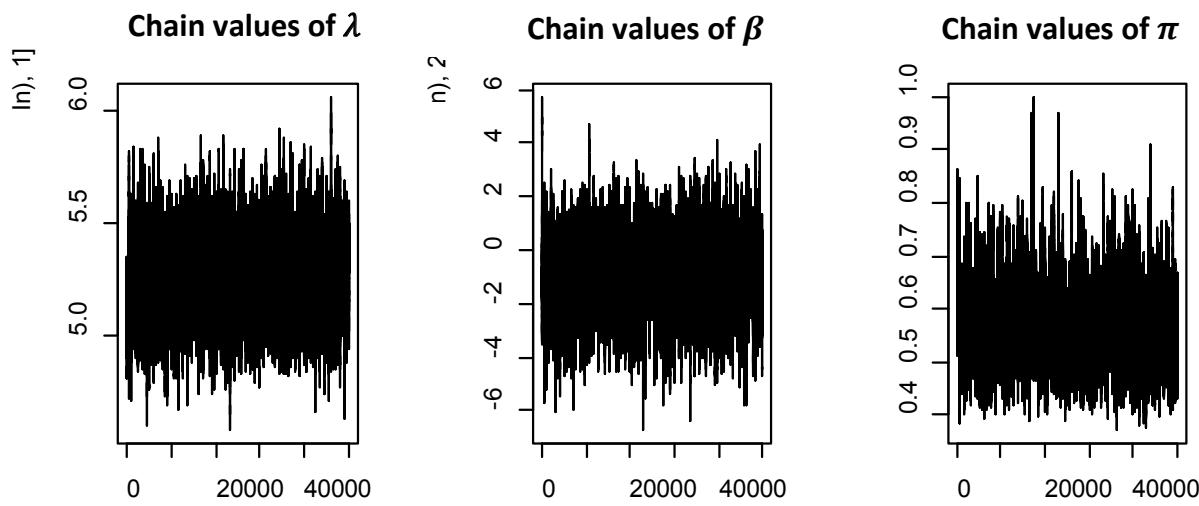


Figure 1 Posterior density function of  $\lambda$ ,  $\beta$  and  $\pi$ .



**Figure 2** Chain of MCMC outputs of  $\lambda$ ,  $\beta$  and  $\pi$ .

**Table 1** Examine patients every 15 days

m	Cases of withdrawal			Number of random removals $r_i$
	Interval in hours $T_i$	Number at risk	Number of failure $k_i$	
1	[0,16)	228	25	2
2	[16,31)	201	39	2
3	[31,46)	160	25	1
4	[46,61)	134	20	3
5	[61,76)	111	11	1
6	[76,91)	99	14	2
7	[91,106)	83	11	3
8	[106,121)	69	17	0
9	[121,136)	52	6	2
10	[136,151)	44	31	1
11	[151,166)	12	6	1
12	[166,181)	5	5	0

**Table 2** Progressively type I interval censored with random removal via the, maximum likelihood

	Different parameters		
	$\lambda$	$\beta$	$\pi$
Average	6.441	0.0841	0.6312
MSE	0.0134	0.0743	0.0484
Bias	0.0231	0.1073	0.0094
Variance	0.0129	0.0627	0.0483
ACI	[5.0132,7.9801]	[-0.1736,0.0901]	[0.4421,0.7782]
Length ACI	2.9669	0.2637	0.3361

**Table 3** Progressively type I interval censored with random removal via the, bootstrap method

Different Parameters			
	$\lambda$	$\beta$	$\pi$
Average	5.966	0.099	0.7058
MSE	0.1174	0.0984	0.0487
Bias	0.0346	0.1764	0.0109
Variance	0.1162	0.0673	0.0486
ABCI	[5.0117,8.0412]	[-0.1811,0.0884]	[0.4434,0.7992]
Length ABCI	3.0295	0.2695	0.3558

**Table 4** Progressively type I interval censored with random removal via the, MCMC procedure developed (*Informative Priors*)

Different Parameters			
	$\lambda$	$\beta$	$\pi$
Average	4.902	0.0083	0.5118
MSE	0.0035	0.0277	0.04804
Bias	0.0049	0.0833	0.0017
Variance	0.0035	0.0207	0.04803
SCI	[5.1023,7.6421]	[-0.1075,0.0826]	[0.4927,0.6524]
Length SCI	2.5398	0.1901	0.1597

**Table 5** Progressively type I interval censored with random removal via the, MCMC procedure developed (*Non-Informative Priors*)

Different Parameters			
	$\lambda$	$\beta$	$\pi$
Average	4.671	0.0398	0.6501
MSE	0.0673	0.0559	0.0656
Bias	0.0174	0.0304	0.0093
Variance	0.067	0.0549	0.0655
SCI	[4.8821,8.0307]	[-0.1010,0.0721]	[0.3881,0.7061]
Length SCI	3.1486	0.1731	0.318

## Conclusion

The methodology developed in this paper will be very useful to the researchers, engineers, statisticians and in the field of medical where such type of life test is needed and especially where the Weibull distribution is used. we have considered the problem of estimation for flexible Weibull distribution in the presence of Progressive Type-I Interval censored sample with Binomial removals. The scope of this censoring scheme in clinical trials has been discussed. We have found that Bayesian procedure provides estimates of the unknown parameters of flexible Weibull model with smaller MSE. The length of SCI is smaller than that of the ACI and ABCI. Applying the MCMC process through the application of the MH algorithm to deal with the Bayesian estimation for another lifetime distributions under type I progressive interval censoring with random removal could be a fruitful future research.

## Appendix

The asymptotic variance-covariance matrix of the maximum likelihood estimators for parameters  $\lambda$ ,  $\beta$  and  $\pi$  are given by elements of the inverse of the Fisher information matrix with random removal will be

$$I(\hat{\lambda}, \hat{\beta}, \hat{\pi}) = \begin{bmatrix} I_1(\hat{\lambda}, \hat{\beta}) & 0 \\ 0 & I_2(\hat{\pi}) \end{bmatrix},$$

Unfortunately, the exact mathematical expressions for the above expectations are very difficult to obtain. Therefore, we give the approximate (observed) asymptotic variance-covariance matrix for the maximum likelihood estimators, which is obtained by dropping the expectation operator E, where

$$I_1^{-1}(\hat{\lambda}, \hat{\beta}) = \begin{bmatrix} \left( \frac{\partial^2 \ln L(\pi)}{\partial \lambda^2} \right) & \left( \frac{\partial^2 \ln L(\pi)}{\partial \lambda \partial \beta} \right) \\ \left( \frac{\partial^2 \ln L(\pi)}{\partial \lambda \partial \beta} \right) & \left( \frac{\partial^2 \ln L(\pi)}{\partial \beta^2} \right) \end{bmatrix}_{\lambda=\hat{\lambda}_2, \beta=\hat{\beta}}^{-1} \begin{bmatrix} V(\hat{\lambda}) & Cov(\hat{\lambda}, \hat{\beta}) \\ Cov(\hat{\lambda}, \hat{\beta}) & V(\hat{\beta}) \end{bmatrix}$$

We explained how to find Fisher information matrix  $I_1(\hat{\lambda}, \hat{\beta})$  in the Appendix B.

$$I_2(\hat{\pi}) = E\left(-\frac{\partial^2 \ln L(\pi)}{\partial \pi^2}\right), \text{ and } \frac{\partial^2 \ln P(R)}{\partial \pi^2} = \frac{-1}{\pi^2} \sum_{j=1}^{m-1} r_j - \frac{1}{(1-\pi)^2} \left[ (m-1)(n-m) - \sum_{j=1}^{m-1} (m-j)r_j \right].$$

Numerical technique is needed to obtain the Fisher information matrix and the variance-covariance matrix. Note that under fixed and random removal the estimates based on intervals with equal length when the intervals are of equal length, so that monitoring and

censoring occur periodically say  $T_i = i \cdot t$ .

We determine the second partials by differentiating the first partials, equations (8) and (9), obtaining

$$\begin{aligned} \frac{\partial^2 \log L}{\partial \lambda \partial \beta} &= \sum_{i=1}^m k_i \left\{ \frac{\frac{\partial^2 F(T_i)}{\partial \lambda \partial \beta} \frac{\partial^2 F(T_{i-1})}{\partial \lambda \partial \beta}}{F(T_i) - F(T_{i-1})} \left[ \frac{\partial F(T_i)}{\partial \lambda} \frac{\partial F(T_{i-1})}{\partial \lambda} \right] \left[ \frac{\partial F(T_i)}{\partial \beta} \frac{\partial F(T_{i-1})}{\partial \beta} \right] \right\} + \sum_{i=1}^m r_i \left\{ \frac{\frac{\partial^2}{\partial \lambda \partial \beta} [1-F(T_i)]}{[1-F(T_i)]} \left[ \frac{\partial}{\partial \lambda} [1-F(T_i)] \right] \left[ \frac{\partial}{\partial \beta} [1-F(T_i)] \right] \right\} \\ \frac{\partial^2 \log L}{\partial \lambda^2} &= \sum_{i=1}^m k_i \left\{ \frac{\frac{\partial^2 F(T_i)}{\partial \lambda^2} \frac{\partial^2 F(T_{i-1})}{\partial \lambda^2}}{F(T_i) - F(T_{i-1})} \left[ \frac{\partial F(T_i)}{\partial \lambda} \frac{\partial F(T_{i-1})}{\partial \lambda} \right]^2 \right\} + \sum_{i=1}^m r_i \left\{ \frac{\frac{\partial^2}{\partial \lambda^2} [1-F(T_i)]}{[1-F(T_i)]} \left[ \frac{\partial}{\partial \lambda} [1-F(T_i)] \right]^2 \right\} \\ \frac{\partial^2 \log L}{\partial \beta^2} &= \sum_{i=1}^m k_i \left\{ \frac{\frac{\partial^2 F(T_i)}{\partial \beta^2} \frac{\partial^2 F(T_{i-1})}{\partial \beta^2}}{F(T_i) - F(T_{i-1})} \left[ \frac{\partial F(T_i)}{\partial \beta} \frac{\partial F(T_{i-1})}{\partial \beta} \right]^2 \right\} + \sum_{i=1}^m r_i \left\{ \frac{\frac{\partial^2}{\partial \beta^2} [1-F(T_i)]}{[1-F(T_i)]} \left[ \frac{\partial}{\partial \beta} [1-F(T_i)] \right]^2 \right\} \end{aligned}$$

$$F(T_i) = 1 - e^{-e^{\lambda T_i} - \frac{\beta}{T_i}}$$

$$\frac{\partial^2 F(T_i)}{\partial \beta^2} = \frac{1}{T_i} e^{\lambda T_i - \frac{\beta}{T_i}} \left\{ \frac{1}{T_i} [1-F(T_i)] + \frac{\partial F(T_i)}{\partial \beta} \right\}$$

$$\frac{\partial F(T_i)}{\partial \lambda} = \frac{-1}{T_i} e^{\lambda T_i - \frac{\beta}{T_i}} e^{-e^{\lambda T_i - \frac{\beta}{T_i}}} = \frac{-1}{T_i} e^{\lambda T_i - \frac{\beta}{T_i}} \{1-F(T_i)\}$$

$$\frac{\partial^2 F(T_i)}{\partial \lambda \partial \beta} = -e^{\lambda T_i - \frac{\beta}{T_i}} \left\{ [1-F(T_i)] + T_i \frac{\partial F(T_i)}{\partial \beta} \right\}$$

$$\frac{\partial F(T_i)}{\partial \beta} = \frac{-1}{T_i} e^{\lambda T_i - \frac{\beta}{T_i}} e^{-e^{\lambda T_i - \frac{\beta}{T_i}}} = \frac{-1}{T_i} e^{\lambda T_i - \frac{\beta}{T_i}} \{1-F(T_i)\}$$

$$\frac{\partial F(T_i)}{\partial \lambda} = [1-F(T_i)] = -\frac{\partial F(T_i)}{\partial \lambda}$$

$$\frac{\partial F(T_i)}{\partial \beta} = [1-F(T_i)] = -\frac{\partial F(T_i)}{\partial \beta}$$

$$\frac{\partial^2 F(T_i)}{\partial \lambda^2} = \frac{1}{T_i} e^{\lambda T_i - \frac{\beta}{T_i}} \left\{ T_i [1-F(T_i)] + \frac{\partial F(T_i)}{\partial \lambda} \right\}$$

Note that:

$$T_i \frac{\partial F(T_i)}{\partial \beta} = -\frac{1}{T_i} \frac{\partial F(T_i)}{\partial \lambda}$$

## Acknowledgement

None.

## Conflict of interest

None.

## References

1. Aggarwala R. Progressive interval censoring: Some mathematical results with applications to inference. *Communications in Statistics - Theory and Methods*. 2001;30(8-9):1921–1935.
2. Cohen AC. Progressively censored samples in life testing. *Technometrics*. 1963;(3):327–339.
3. Balakrishnan N, Aggarwala R. *Progressive Censoring: Theory, Methods, and Applications*, Boston, MA: Birkhäuser Boston. 2000.
4. Yuen HK, Tse SK. Parameters estimation for Weibull distributed lifetimes under progressive censoring with random removals. *Journal of Statistical Computation and Simulation*. 1996;55(1-2):57–71.
5. Yang C, Tse SK, Yuen HK. Statistical analysis of Weibull distributed life time data under type II progressive censoring with binomial removals. *Journal of Applied Statistics*. 2000;27(8):1033–1043.
6. Kendall PJ, Anderson RL. An Estimation Problem in Life Testing. *Technometrics*. 1971;13(2):289–301.
7. Ashour SK, Afify WM. Statistical analysis of exponentiated Weibull family under type-i progressive interval censoring with random removals. *Journal of Applied Sciences Research*. 2007;3(12):1851–1863.
8. Lin CT, Wu SJS, Balakrishnan N. Planning life tests with progressively type-I interval censored data from the lognormal distribution. *Journal of Statistical Planning and Inference*. 2009;139(1):54–61.
9. Ng HKT, Wang Z. Statistical estimation for the parameters of Weibull distribution based on progressively type-I interval censored sample. *Journal of Statistical Computation and Simulation*. 2009;79(2):145–159.
10. Chen DG, Lio YL. Parameter estimations for generalized exponential distribution under progressive type-I interval censoring. *Computational Statistics and Data Analysis*. 2010;54(6):1581–1591.
11. Green EJ, Roesch FA, Smith AFM, et al. Bayesian estimation for the three-parameter Weibull distribution with tree diameter data. *Biometrika*. 1994;81(1):254–269.
12. Gupta A, Mukherjee B, Upadhyay SK. Weibull extension model: A Bayes study using Markov Chain Monte Carlo simulation. *Reliability Engineering & System Safety*. 2008;93(10):1434–1443.
13. Lin YJ, Lio YL. Bayesian inference under progressive type-I interval censoring. *Journal of Applied Statistics*. 2012;39(8):1811–1824.
14. Wu SJ, Chang CT. Parameter Estimations Based on Exponential Progressive Type II Censored with Binomial Removals. *International Journal of Information and Management Sciences*. 2002;13(3):37–46.
15. Kemp CD, Kemp W. Rapid generation of frequency tables. *Journal of the Royal Statistical Society*. 1987;36(3):277–282.
16. Efron B. The bootstrap and other resampling plans. *CBMS-NSF Regional Conference Series in Applied Mathematics*. SIAM, Philadelphia, USA.1982:38.
17. Balakrishnan N, Sandhu RA. A simple simulational algorithm for generating progressive type-II censored samples. *The American Statistician*. 1995;49(2):229–230.
18. Gilks WR, Richardson S, Spiegelhalter DJ. *Markov chain Monte Carlo in Practices*, Chapman and Hall, London. 1996.
19. Metropolis N, Rosenbluth AW, Rosenbluth MN, et al. Equations of state calculations by fast computing machine. *Journal of Chemical Physics*. 1953;21:1087–1091.
20. Hastings WK. Monte carlo sampling methods using markov chains and their applications. *Biometrika*.1970;57(1):97–109.