Inference for Exponential Competing Risk Failure Model

Neha K Gadhvi[1*] and G. C. Bhimani[2]

[1] H. L. Institute of Commerce, Department of Statistics, Ahmedabad University, Near Commerce Six Road, Ahmedabad-380009, Gujarat, India,
[2] Department of Statistics, Saurashtra University, University Road, Rajkot-360005, Gujarat, India
Email: gcbhimani@yahoo.co.in

Abstract. In this paper the exponential life time model is used as a competing risk model. The causes of failures are assumed independent. For analysis of the data Type-I progressive interval censoring is used. Point estimation and interval estimation of the parameters and survival function are made using maximum likelihood and bootstrap methods. Non-parametric estimation is also considered. A real life example is provided to illustrate the results obtained in the paper.

Keywords: Progressive Type-I, interval censoring, maximum likelihood, bootstrap, survival function, non-parametric.

1   Introduction

Nowadays the analysis of time to event data is of great importance in the fields of medicine, engineering etc. For example, time to death of patients, and life of electrical components under test. Sometimes failure of an item or individual may be due to more than one cause. For example, Boag[1] studied the case of a breast cancer patient, where the cause of death was recorded as "cancer" or "other"; Peck[2] stated that failure of transistors may be due to electrical degradation of certain parts or by faulty bounding of the leads. When the items or individuals fail the failure time and an indicator number corresponding to the specific cause of failure were recorded. Here we assume the causes of failure are independent.

Suppose that an item exhibits k modes (causes) of failures. When the item begins operation, each failure mode simultaneously generates a random life that is independent of the other modes. Thus, in effect, k life times are denoted by $X_1, X_2, \ldots, X_i, \ldots, X_k$ which simultaneously begin, where life time X_i corresponds to the ith cause of failure, $i = 1, 2, \ldots, k$; and failure of the item occurs, as soon as any one of the k life times, say X_i is realized. Hence the life length of the unit, denoted by the random variable X is nothing but

$$X = \min\left\{X_1, X_2, \ldots, X_i, \ldots, X_k\right\} \cong X_{(1)}$$

and its cumulative distribution function (cdf) of X, say $F_X(x)$ is given by

$$F_X(x) = 1 - \prod_{i=1}^{k}\left[F_{X_i}(x)\right] \tag{1}$$

Such a model is called competing risk failure model and $F_X(x)$ becomes its cdf.

This model is different than the mixture model. In mixture model only one of the k possible causes of failure generates a random life that causes part failures.

In practice, the attributes like electrical thermal, climate and mechanical stresses applied to an item may be observed as causes of failures of an item.

Mendenhall and Hader[3], Patel and Gajjar[4], Boardman and Kendell[5], Patel[6], have considered progressive type I grouped (interval) censoring for various types of mixture and compound (competing risk) life time models.

2   Competing Risk Failure Model

Suppose that the device exhibits two modes of failures and each failure mode simultaneously generates a random life time X_1 and X_2 respectively i.e. X_i is the time of failure of the device due to cause I and the probability density function (pdf) of X_i is given by

$$f_i(x) = \frac{1}{\theta} e^{-x/\theta_i}; \ x > 0; \ \theta_i > 0; \ i = 1,2 \tag{2}$$

and the corresponding cdf is

$$F_i(x) = P(X_i \le x) = 1 - e^{-x/\theta_i}; \ i = 1,2 \tag{3}$$

Let T be the time of failure of the device regardless of cause, then $T = \min(X_1, X_2)$. So pdf of T will be

$$h_T(t) = \frac{1}{\theta} e^{-t/\theta} \tag{4}$$

and its cdf will be

$$H_T(t) = 1 - e^{-t/\theta}, \quad \text{where} \quad 1/\theta = 1/\theta_1 + 1/\theta_2. \tag{5}$$

Hence the survival function of the model will be

$$f(t) = 1 - H_T(t) = e^{-t/\theta}$$

According to Boardman & Kendell[5] the competing risk failure model can be obtained as follow:

Let $g_i(t)$ be the probability that an item fails by cause i and it does not fail by the cause $i \ne i$ up to time t with two independent causes only, then

$$g_i(t) = f_i(t)\left[1 - F_{i'}(t)\right], i' \ne i = 1,2$$

Hence the pdf of T can be obtained as

$$\begin{aligned}
h_T(t) &= c(g_1(t) + g_2(t)) \\
&= c\left\{ f_1(t)\left[1 - F_2(t)\right] + f_2(t)\left[1 - F_1(t)\right] \right\} \\
&= c\frac{1}{\theta_1} e^{-t/\theta_1}\left(e^{-t/\theta_2} \right) + c\frac{1}{\theta_2} e^{-t/\theta_2}\left(e^{-t/\theta_1} \right) \\
&= \frac{1}{\theta_1} e^{-t/\theta} + \frac{1}{\theta_2} e^{-t/\theta}
\end{aligned} \tag{6}$$

where c is a normalizing constant, here it will be 1. We call the failure model in (6) as the exponential competing risk failure model.

3   Interval Censoring

Interval censored data arise when observations are known to lie only in some interval between time points a and b. Here experimental units are not monitored continuously. Such data may arise in a variety of circumstances but are commonly encountered in medical studies, where patients are only monitored at regular interval, e.g. weekly or quarterly check-up. In engineering such kind of censoring is known as group censoring. The generalization of such censoring scheme are progressive Type-I interval censoring or progressive Type-II interval censoring schemes. Several authors have considered such generalized censoring schemes in life testing experiments. Some of them are Patel & Patel[7], Shah & Patel[8], Patel and Gajjar[9], Arora, Bhimani and Patel[10]. Several authors have used such interval censoring scheme in clinical, medical, biomedical engineering studies like Odell et al[11], Samuelson and Kongerud[12], Rao[13], Aggarwala[14], etc. Recently Gajjar &Patel[15], Gadhvi and Bhimani[16], Patel & Patel[17], have considered Progressive Type-I group censoring for discrete (geometric) life time model. Patel and Gajjar[4] have considered maximum likelihood estimation for compound exponential life time model based on progressively Type-I grouped censored sample. They have assumed different parameter at each stage of censoring.

Here we consider the estimation for exponential competing risk failure model under progressive Type-I interval censoring keeping the same parameter (unchanged) at each stage of censoring. MLE and Confidence interval estimation is considered. Bootstrap methods are also used to derive the confidence intervals. Non-parametric method for estimating survival function is also considered. A real life example is used to exemplify the theory.

The setup of m-stage progressive Type-I censoring scheme is described as follows:

Let n units are put on life test at time zero and observations on each of these units is continued until the unit fails or is censored. Units are observed at pre-set times $T_1, T_2, ..., T_m,$ where m is a fixed integer. Thus the time axis is partitioned into mutually exclusive intervals $I_i = (T_{i-1}, T_i], i = 1, 2, ..., m$ and $T_0 = 0,$ where T_m is the time when experimentation is scheduled to terminate.

Let x_{ji} denote the number of units falling in I_i due to cause j; $j = 1, 2$ and r_i be the set of live units removed at time $T_i; i = 1, 2, ..., m$ and r_m is equal to all the remaining units at time T_m. The values $r_1, r_2, ..., r_m$ may be pre-specified as positive integers or percentages $p_1, p_2, ..., p_m$, with $p_m = 100$ of the remaining live units.

Then based on the observed data $\{x_{11}, x_{12}, ..., x_{1m}, x_{21}, x_{22}, ..., x_{2m},\}$ the joint likelihood function will be proportional to the following expression:

$$L(\theta_1, \theta_2) \propto \prod_{i=1}^{m} \left[G_1(T_i) - G_1(T_{i-1}) \right]^{x_{1i}} \prod_{i=1}^{m} \left[G_2(T_i) - G_2(T_{i-1}) \right]^{x_{2i}} \prod_{i=1}^{m} \left[1 - H(T_i) \right]^{r_i} \tag{7}$$

with

$$r_m = n - \sum_{i=1}^{m} x_{1i} - \sum_{i=1}^{m} x_{2i} - \sum_{i=1}^{m-1} r_i \tag{8}$$

Based on $g_1(t)$, defined in (6), we define

$$G_1(T) = P(X_1 \leq T) = \int_0^T g_1(t)\, dt = \int_0^T \frac{1}{\theta_1} e^{-t/\theta}\, dt = \frac{\theta}{\theta_1}\left[1 - e^{-T/\theta} \right] \tag{9}$$

and

$$G_2(T) = P(X_2 \leq T) = \int_0^T g_2(t)\, dt = \int_0^T \frac{1}{\theta_2} e^{-t/\theta}\, dt = \frac{\theta}{\theta_2}\left[1 - e^{-T/\theta} \right] \tag{10}$$

Using (9) and (10) in (7), the likelihood function becomes

$$L = L(\theta_1, \theta_2) \propto \prod_{i=1}^{m} \left[\frac{\theta}{\theta_1}\left\{ e^{-T_{i-1}/\theta} - e^{-T_i/\theta} \right\} \right]^{x_{1i}} \prod_{i=1}^{m} \left[\frac{\theta}{\theta_2}\left\{ e^{-T_{i-1}/\theta} - e^{-T_i/\theta} \right\} \right]^{x_{2i}} \prod_{i=1}^{m} \left[e^{-T_i/\theta} \right]^{r_i}$$
$$= \theta^{\sum_1^m x_i} \theta_1^{-\sum_1^m x_{1i}} \theta_2^{-\sum_1^m x_{2i}} e^{-y/\theta} \prod_{i=1}^{m} \left[1 - e^{-S_i/\theta} \right]^{x_i} \tag{11}$$

where

$$x_i = x_{1i} + x_{2i}, S_i = T_i - T_{i-1}, y = \sum_1^m T_{i-1} x_i + \sum_1^m T_i r_i \tag{12}$$

To obtain maximum likelihood estimate of θ_1 and θ_2, we would like to maximize:

$$\log L = -\sum_1^m x_i \log\left(\frac{1}{\theta_1} + \frac{1}{\theta_2} \right) - \sum_1^m x_{1i} \log \theta_1 - \sum_1^m x_{2i} \log \theta_2 - y\left(\frac{1}{\theta_1} + \frac{1}{\theta_2} \right) + \sum_1^m x_i \log\left\{ 1 - e^{-S_i\left(\frac{1}{\theta_1} + \frac{1}{\theta_2} \right)} \right\} \tag{13}$$

Thus, the equations to be solved for the maximum likelihood estimate of θ_1 and θ_2 are

$$\frac{\partial \log L}{\partial \theta_1} = \frac{\sum_1^m x_i}{\theta_1^2 \left(\frac{1}{\theta_1} + \frac{1}{\theta_2} \right)} - \frac{\sum_1^m x_{1i}}{\theta_1} + \frac{y}{\theta_1^2} - \frac{1}{\theta_1^2} \sum_i^m \left\{ \frac{x_i S_i e^{-S_i\left(\frac{1}{\theta_1} + \frac{1}{\theta_2} \right)}}{1 - e^{-S_i\left(\frac{1}{\theta_1} + \frac{1}{\theta_2} \right)}} \right\} = 0 \tag{14}$$

and

$$\frac{\partial \log L}{\partial \theta_2} = \frac{\sum_{1}^{m} x_i}{\theta_2^2 \left(\frac{1}{\theta_1} + \frac{1}{\theta_2} \right)} - \frac{\sum_{1}^{m} x_{2i}}{\theta_2} + \frac{y}{\theta_2^2} - \frac{1}{\theta_2^2} \sum_{i}^{m} \left\{ \frac{x_i S_i e^{-S_i \left(\frac{1}{\theta_1} + \frac{1}{\theta_2} \right)}}{1 - e^{-S_i \left(\frac{1}{\theta_1} + \frac{1}{\theta_2} \right)}} \right\} = 0 \tag{15}$$

From (14) and (15), we get

$$\theta_2 = \frac{\theta_1 \sum_{1}^{m} x_{2i}}{\sum_{1}^{m} x_{1i}} \tag{16}$$

Again from (14) we can write,

$$\theta_1 = \frac{\sum_{1}^{m} x_i \left(\frac{\theta_1 \theta_2}{\theta_1 + \theta_2} \right) + y - \sum_{i}^{m} \left\{ \frac{x_i S_i e^{-S_i \left(\frac{1}{\theta_1} + \frac{1}{\theta_2} \right)}}{1 - e^{-S_i \left(\frac{1}{\theta_1} + \frac{1}{\theta_2} \right)}} \right\}}{\sum_{1}^{m} x_{1i}} \tag{17}$$

Substituting θ_2 from (16) in (17) we get the right hand side (17) as a function of parameter θ_1 only. Thus, using (16) & (17) we can write:

$$\theta_1 = \psi(\theta_1) = \text{ a function of } \theta_1 \tag{18}$$

Solving the equation (18), by any method of iteration we get MLE of θ_1 say $\hat{\theta}_1$, and substituting it in equation (16), we will get MLE of θ_2 say $\hat{\theta}_2$.

The MLE of survival function at time t_0 is given by

$$\hat{S}(t_0) = e^{-t \left(\frac{1}{\hat{\theta}_1} + \frac{1}{\hat{\theta}_2} \right)} \tag{19}$$

4   Standard Error of the Estimators

The asymptotic variances and covariance of the MLE of the parameters θ_1 and θ_2 can be obtained from the elements of the inverse of the Fisher information matrix (I) given by:

$$I^{-1} = \begin{pmatrix} -E\left(\frac{\partial^2 \log L}{\partial \theta_1^2} \right) & -E\left(\frac{\partial^2 \log L}{\partial \theta_1 \partial \theta_2} \right) \\ -E\left(\frac{\partial^2 \log L}{\partial \theta_1 \partial \theta_2} \right) & -E\left(\frac{\partial^2 \log L}{\partial \theta_2^2} \right) \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} V\left(\hat{\theta}_1 \right) & \text{cov}\left(\hat{\theta}_1, \hat{\theta}_2 \right) \\ \text{cov}\left(\hat{\theta}_1, \hat{\theta}_2 \right) & V\left(\hat{\theta}_2 \right) \end{pmatrix}$$

The exact mathematical expressions for the above expectations are difficult to obtain, therefore, we calculate observed asymptotic variance-covariance for the MLE by dropping the expectation 'E'.

$$\begin{pmatrix} V\left(\hat{\theta}_1 \right) & \text{cov}\left(\hat{\theta}_1, \hat{\theta}_2 \right) \\ \text{cov}\left(\hat{\theta}_1, \hat{\theta}_2 \right) & V\left(\hat{\theta}_2 \right) \end{pmatrix} \simeq \begin{pmatrix} -\frac{\partial^2 \log L}{\partial \theta_1^2} & -\frac{\partial^2 \log L}{\partial \theta_1 \partial \theta_2} \\ -\frac{\partial^2 \log L}{\partial \theta_1 \partial \theta_2} & -\frac{\partial^2 \log L}{\partial \theta_2^2} \end{pmatrix}_{(\hat{\theta}_1, \hat{\theta}_2)} \tag{20}$$

where

$$\frac{\partial^2 \log L}{\partial \theta_1^2} = \left(\frac{-\theta_2}{\theta_1^2}\right)\left(\frac{2\theta_1 + \theta_2}{\left(\theta_1 + \theta_2\right)^2}\right)\sum_1^m x_i + \frac{\sum_1^m x_{1i}}{\theta_1^2} - \frac{2y}{\theta_1^3} - \frac{2}{\theta_1^3}\sum_1^m \left\{\frac{x_i S_i}{\left(e^{S_i\left(\frac{1}{\theta_1}+\frac{1}{\theta_2}\right)} - 1\right)}\right\} - \frac{1}{\theta_1^4}\sum_1^m \left\{\frac{x_i S_i^2 e^{S_i\left(\frac{1}{\theta_1}+\frac{1}{\theta_2}\right)}}{\left(e^{S_i\left(\frac{1}{\theta_1}+\frac{1}{\theta_2}\right)} - 1\right)^2}\right\} \quad (21)$$

$$\frac{\partial^2 \log L}{\partial \theta_2^2} = \left(\frac{-\theta_1}{\theta_2^2}\right)\left(\frac{2\theta_1 + \theta_2}{\left(\theta_1 + \theta_2\right)^2}\right)\sum_1^m x_i + \frac{\sum_1^m x_{2i}}{\theta_2^2} - \frac{2y}{\theta_2^3} - \frac{2}{\theta_2^3}\sum_1^m \left\{\frac{x_i S_i}{\left(e^{S_i\left(\frac{1}{\theta_1}+\frac{1}{\theta_2}\right)} - 1\right)}\right\} - \frac{1}{\theta_2^4}\sum_1^m \left\{\frac{x_i S_i^2 e^{S_i\left(\frac{1}{\theta_1}+\frac{1}{\theta_2}\right)}}{\left(e^{S_i\left(\frac{1}{\theta_1}+\frac{1}{\theta_2}\right)} - 1\right)^2}\right\} \quad (22)$$

$$\frac{\partial^2 \log L}{\partial \theta_1 \partial \theta_2} = \frac{\sum_1^m x_i}{\left(\theta_1 + \theta_2\right)^2} - \frac{1}{\theta_1^2 \theta_2^2}\sum_1^m \left\{\frac{x_i S_i^2 e^{S_i\left(\frac{1}{\theta_1}+\frac{1}{\theta_2}\right)}}{\left(e^{S_i\left(\frac{1}{\theta_1}+\frac{1}{\theta_2}\right)} - 1\right)^2}\right\} \quad (23)$$

Hence the asymptotic variance of survival function can be obtained from the equation

$$V\left(\hat{S}(t_0)\right) = \left[\left(\frac{\partial S}{\partial \theta_1}\right)^2 V(\hat{\theta}_1) + \left(\frac{\partial S}{\partial \theta_2}\right)^2 V(\hat{\theta}_2) + 2\left(\frac{\partial S}{\partial \theta_1}\right)\left(\frac{\partial S}{\partial \theta_2}\right)\mathrm{cov}(\hat{\theta}_1, \hat{\theta}_2)\right]_{(\hat{\theta}_1, \hat{\theta}_2)} \quad (24)$$

5 Simulation Algorithm

A study of properties of maximum likelihood estimators based on progressively Type-I interval censored samples involves simulation. A short algorithm for simulating a random sample of size n put on a life test at time 0 is given below. Here we use the following properties of the progressive interval censoring

$$X_{11} \sim B\left(n, G_1\left(T_1\right)\right) \text{ and } X_{21} \sim B\left(n, G_2\left(T_1\right)\right)$$

where $G_1\left(T_1\right)$ and $G_2\left(T_1\right)$ are defined in (9) and (10) respectively and for $i = 2, 3, \ldots, m$

$$X_{ji} / X_{ji-1}, X_{ji-2}, X_{ji-3\ldots}, X_{j1}, R_{i-1}, R_{i-2}, \ldots, R_1 \sim B\left(n - \sum_{s=1}^{i-1}\left(X_{1s} + X_{2s} + R_s\right); \frac{G_j\left(T_i\right) - G_j\left(T_{i-1}\right)}{1 - \sum_{s=1}^{i-1}\left\{G_j\left(T_i\right) - G_j\left(T_{i-1}\right)\right\}}\right)$$

$$= B\left(n - \sum_{s=1}^{i-1}\left(X_{1s} + X_{2s} + R_s\right); \frac{G_j\left(T_i\right) - G_j\left(T_{i-1}\right)}{1 - G_j\left(T_{i-1}\right)}\right); \quad j = 1, 2$$

Here $B\left(n, p\right)$ denotes the binomial distribution with parameters n and p; $0 < p < 1$.

On the basis of the algorithm given in Aggarwala[14] for simulating a sample under progressive Type-I interval censoring scheme, Gadhvi and Bhimani[16] have modified the simulation scheme for competing risk failure model. The main steps of the algorithm are given below:

Algorithm 1.

1:     Set $\ i = 0, X_{1sum} = 0, X_{2sum} = 0, R_{sum} = 0$

2:     Next i

3:     If $\ i = m + 1,\ $ exit the algorithm.

4:     Generate $\ X_{11}\ $ and $\ X_{21}\ $ as binomial random variables with parameters
    $\left(n, G_1\left(T_1\right)\right)$ and $\left(n, G_2\left(T_1\right)\right)$ respectively.

5:     Generate $\ X_{1i}\ $ and $\ X_{2i}\ $ as binomial variables with parameter
$$\left(n - X_{1sum} - X_{2sum} - R_{sum}, \frac{G_1\left(T_i\right) - G_1\left(T_{i-1}\right)}{1 - G_1\left(T_{i-1}\right)}\right) \text{ and}$$
$$\left(n - X_{1sum} - X_{2sum} - R_{sum}, \frac{G_2\left(T_i\right) - G_2\left(T_{i-1}\right)}{1 - G_2\left(T_{i-1}\right)}\right) \text{ respectively}$$

6:     Calculate $\ R_i^{obs} = \text{Floor}\left[p_i \left(n - X_{1sum} - X_{2sum} - R_{sum}\right)\right]\ $ or
    $\min\left(R_i, n - X_{1sum} - X_{2sum} - R_{sum} - X_i\right)$

7:     Set $\ X_{1sum} = X_{1sum} + X_{1i}, X_{2sum} = X_{2sum} + X_{2i}, R_{sum} = R_{sum} + R_i^{obs}$

This algorithm generates m binomial random variables. Here either the values $p_1, p_2 \dots p_{m-1}$ or proposed values of $R_1, R_2 \dots R_m$ are fixed in advance by the experimenter. Here $p_m = 1$ and $R_m = n - \sum_1^m X_{1i} - \sum_1^m X_{2i} - \sum_1^{m-1} R_i$.

6   Confidence Interval Estimation

A.   Asymptotic Confidence Interval

Using the asymptotic normality property of maximum likelihood estimator, confidence interval for MLE can be obtained for parameters θ_1, θ_2 and for survival function $S\left(t_0\right)$ as follows:

$\left(1 - \propto\right)100\%$ asymptotic confidence interval for $\hat{\theta}_i$ becomes

$$\hat{\theta}_i \pm Z_{\alpha/2} \sqrt{V\left(\hat{\theta}_i\right)}, i = 1, 2$$

and for $\ S\left(t_0\right)$:

$$\hat{S}\left(t_0\right) \pm Z_{\alpha/2} \sqrt{V\left(\hat{S}\left(t_0\right)\right)}$$

where $V\left(\hat{\theta}_i\right)$ and $V\left(\hat{S}\left(t_0\right)\right)$ can be obtained from equations (20) & (24) and $Z_{\alpha/2}$ is $\left(\alpha/2\right)^{th}$ percentile of standard normal distribution.

B.   Bootstrap Confidence Interval

B-I: Percentile Bootstrap Method

Algorithm 2.

1: From the original data X compute the ML estimates of the parameters say $\hat{\theta}_1$ and $\hat{\theta}_2$ from (18) and (16).

2: Use $\hat{\theta}_1$ and $\hat{\theta}_2$ to generate a bootstrap sample $\underline{X}^*$ with the same values of T_i, r_i and m, $i = 1, 2, \ldots, m$ using the algorithm given in section 5.

3: As discussed in step 1, based on $\underline{X}^*$ compute the bootstrap sample estimator of θ_1, θ_2 and $S(t_0)$, say $\hat{\theta}_1^*$, $\hat{\theta}_2^*$ and $\hat{S}^*(t_0)$.

4: Repeat steps 2 and 3, S times representing S bootstrap MLE's of $(\theta_1, \theta_2, \mathbf{S}(t_0))$ based on S different bootstrap samples.

5: Arrange all $\hat{\theta}_1^*$, $\hat{\theta}_2^*$ and $\hat{S}^*(t_0)$ in an ascending order to obtain bootstrap sample $\left(\phi_l^{[1]}, \phi_l^{[2]}, \ldots, \phi_l^{[S]}\right), l = 1, 2, 3$ where $\phi_1 = \hat{\theta}_1^*, \phi_2 = \hat{\theta}_2^*$ and $\phi_3 = \hat{S}^*(t_0)$.

6: Let $G(Z) = P(\phi_l \leq Z)$ be the cumulative distribution function of ϕ_l. Define $\phi_{lboot} = G^{-1}(Z)$ for given Z. The approximate bootstrap $(1 - \alpha)100\%$ confidence interval of ϕ_l is given by $\left[\phi_{lboot}\left(\frac{\alpha}{2}\right), \phi_{lboot}\left(1 - \frac{\alpha}{2}\right)\right]$.

B-II.   Bootstrap-t Method

Algorithm 3.

1: From the original data $\underline{x}$ compute the MLE of the parameters say, $\hat{\theta}_1, \hat{\theta}_2$ and $\hat{S}(t_0)$.

2: Use $\hat{\theta}_1$ and $\hat{\theta}_2$ to generate bootstrap sample $\underline{X}^*$ with the same values of T_i, r_i and m, $i = 1, 2, \ldots, m$ using the algorithm given in B-I.

3: Based on $\underline{X}^*$ compute the bootstrap sample estimates of θ_1, θ_2 and $S(t_0)$, say $\hat{\theta}_1^*$, $\hat{\theta}_2^*$ and $\hat{S}^*(t_0)$.

4: Compute the following statistics:

$$T_1^* = \frac{\sqrt{n}\left(\hat{\theta}_1^* - \hat{\theta}_1\right)}{\sqrt{V\left(\hat{\theta}_1^*\right)}}, T_2^* = \frac{\sqrt{n}\left(\hat{\theta}_2^* - \hat{\theta}_2\right)}{\sqrt{V\left(\hat{\theta}_2^*\right)}}, T_3^* = \frac{\sqrt{n}\left(\hat{S}^*(t_0) - \hat{S}(t_0)\right)}{\sqrt{V\left(\hat{S}^*(t_0)\right)}}.$$

5: Repeat steps 3 and 4 S (boot)times.

6: From the values of $T_i^*, i = 1, 2, 3$ obtained in step 4, determine the upper and lower bounds of the $100(1 - \alpha)\%$ confidence interval of the parameters and survival function as follows:
Let $H(x) = P\left(T_i^* \leq x\right), i = 1, 2, 3$ be the cdf of T_i^*. For a given x define $\left(\hat{\theta}_{1Boot-t}\left(\frac{\alpha}{2}\right), \hat{\theta}_{1Boot-t}\left(1 - \frac{\alpha}{2}\right)\right)$. Similarly we can define other parameters.

---

7   Non-Parametric Estimation of Survival Function

Kaplan Meier estimate of survival function $S(t_0)$ can be obtained according to Miller et al[18]. Its estimate and the estimate of the variance can be obtained as follow:

$$\hat{S}(t_0) = \prod_{Y_{(i)} \le t_0} \left(\frac{n-i}{n-i+1} \right)^{\delta_{(i)}} \tag{25}$$

and

$$Asy\hat{V}\left(\hat{S}(t_0)\right) = \hat{S}^2(t_0) \sum_{y_{(i)} \le t_0} \frac{\delta_{(i)}}{(n-i)(n-i+1)} \tag{26}$$

Using the results (7.1) & (7.2), the $(1-\propto)100\%$ asymptotic confidence interval for $\hat{S}(t_0)$ can be obtained

$$\hat{S}(t_0) \pm Z_{\alpha/2} \sqrt{Asy\hat{V}\left(\hat{S}(t_0)\right)} \tag{27}$$

8   Application

The data about the failure time of radio transceivers considered by Mendenhall and Hader[3] is given below . Failures are classified as one of the two types: those confirmed on arrival at the maintenance centre (Type-I) and those unconfirmed (Type-II). The data consist of a failure time (in hours) and type for each failure until when observations cease after 600 hours. We have modified the data considering the withdrawals at the end of each failure time interval.

Number	Time interval	Type – I Failures:	Type – II Failures:	Withdrawals
i	(T_{i-1}, T_i)	x_{1i}	x_{2i}	R_i
1	0-50	26	15	3
2	50-100	27	14	2
3	100-150	28	20	0
4	150-200	35	13	0
5	200-250	17	11	0
6	250-300	20	8	1
7	300-350	10	7	1
8	350-400	11	5	1
9	400-450	11	3	0
10	450-500	7	4	0
11	500-550	6	1	0
12	550-600	9	2	51

As per our notations we have $S_i = T_i - T_{i-1} = 50$, $m = 12$, $T_m = 600$, $T_0 = 0$, $n = 369$, $r_m = 51$.
Solving the equations (16) to (17) of section 3 we find MLEs of θ_1 and θ_2 as

$$\widehat{\theta}_1 = 468.568470 \quad\text{and}\quad \widehat{\theta}_2 = 941.686148 \tag{28}$$

From (3.13) we get MLE for survival function at time $t_0 = 150$ as

$$\hat{S}(t_0 = 150) = 0.619147$$

Using (20) the asymptotic variance –covariance matrix of the MLEs for parameter θ_1 and θ_2 is given by

$$\begin{pmatrix} V\left(\widehat{\theta}_1\right) & Cov\left(\widehat{\theta}_1, \widehat{\theta}_2\right) \\ Cov\left(\widehat{\theta}_1, \widehat{\theta}_2\right) & V\left(\widehat{\theta}_2\right) \end{pmatrix} = \begin{pmatrix} 228.490456 & 64.098690 \\ 64.098690 & 1854.391245 \end{pmatrix}$$

Hence the asymptotic standard errors of the MLEs of the parameters will be

$$SE(\widehat{\theta_1}) = 15.115901 \quad \text{and} \quad SE(\widehat{\theta_2}) = 43.062643$$

Hence from (24) we get $SE\left(\hat{S}\left(t_0 = 150\right)\right) = 0.008179$.

To apply bootstrap confidence interval estimation for the parameters we have made 1000 simulations based on the MLEs of θ_1 and θ_2 given in (28) with the other values fixed as follows:

$S_i = T_i - T_{i-1} = 50$, $m = 12$, $T_m = 600$, $T_0 = 0$, $n = 369$ of the given real life data.

The summary statistics for our simulation are given in the following table:

Table 1. Summary statistics for simulation

	min	0.025	0.25	0.50	0.75	0.975	max
$\widehat{\theta_1}$	453.881921	534.49941	600.519587	635.009587	666.823853	738.014578	815.657104
$\widehat{\theta_2}$	1103.24111	1196.47891	1380.32981	1490.57503	1605.86791	1888.97570	2276.43244
$\hat{S}\left(t_0\right)$	0.640287	0.674061	0.701207	0.714147	0.723910	0.744047	0.763701
$T\widehat{\theta_1}$	-19.556925	70.818664	123.872344	146.354047	163.912440	195.743508	216.159550
$T\widehat{\theta_2}$	58.444155	83.146751	118.733825	134.023486	146.362393	166.596258	177.334380
$T\hat{S}\left(t_0\right)$	50.028425	136.05393	213.647116	252.157978	283.773046	349.591892	427.982516

Based on the simulation results the confidence interval based on MLE and the bootstrap confidence intervals for parameters and the survival function are computed using the methods described in Section 6, which are given in the following table.

Table2. Estimates and confidence intervals for the parameters based on MLE and Boot strap

Method	Parameter	Estimate	Confidence Interval	Length of the Interval
MLE	θ_1	634.857929	(591.910652, 677.805205)	85.894553
	θ_2	1502.724540	(1345.933796, 1659.515285)	313.581489
	S($t_0 = 150$)	0.712668	(0.698474, 0.726863)	0.028389
Percentile Bootstrap	θ_1	634.857929	(534.499419, 738.014578)	203.515159
	θ_2	1502.724540	(1196.478918, 1888.975706)	692.496788
	S($t_0 = 150$)	0.712668	(0.674061, 0.744047)	0.069986
Bootstrap - t	θ_1	634.857929	(524.295848, 622.599502)	98.303654
	θ_2	1502.724540	(1128.080524, 1315.153596)	187.073072
	S($t_0 = 150$)	0.712668	(0.677079, 0.768004)	0.090925

The following tables give estimate of the survival (Reliability) function ($S(t_0)$) of an item at time t_0. Its asymptotic variance and asymptotic 95% confidence interval of $S(t_0)$ based on MLE and nonparametric estimation are calculated as discussed in sections 6 and 7.

Table 3. Estimate of survival (Reliability) function ($S(t_0)$), with its asymptotic variance and asymptotic confidence interval calculated using MLE

t_0	$\hat{S}(t_0)$	AsyV($\hat{S}(t_0)$)	95% Confidence interval for $\hat{S}(t_0)$		Length of CI
			Lower limit	Upper limit	
50	0.852311	1.408663E-05	0.844954	0.859667	0.014713
100	0.726433	4.093199E-05	0.713894	0.738973	0.025079
150	0.619147	6.690232E-05	0.603115	0.635178	0.032063
200	0.527705	8.640014E-05	0.509486	0.545924	0.036438
250	0.449769	9.806867E-05	0.430359	0.469179	0.03882
300	0.383343	0.000103	0.363491	0.403195	0.039704
350	0.326727	0.000101	0.306987	0.346467	0.03948
400	0.278473	9.624045E-05	0.259245	0.297701	0.038456
450	0.237345	8.848273E-05	0.218909	0.255782	0.036873
500	0.202292	7.935409E05	0.184832	0.219752	0.034920
550	0.172416	6.975100E-05	0.156046	0.188785	0.032739
600	0.146952	6.030084E-05	0.131732	0.162171	0.030439

Table 4. Estimate of survival (Reliability) function ($S(t_0)$), with its asymptotic variance and asymptotic confidence interval calculated using nonparametric estimation

t_0	$\hat{S}(t_0)$	AsyV($\hat{S}(t_0)$)	95% Confidence interval for $\hat{S}(t_0)$		Length of CI
			Lower limit	Upper limit	
50	0.888889	0.000268	0.856823	0.920955	0.064132
100	0.776752	0.000268	0.744665	0.808839	0.064174
150	0.644539	0.000302	0.610467	0.678611	0.068143
200	0.512326	0.000289	0.478979	0.545673	0.066694
250	0.435202	0.000180	0.408872	0.461531	0.052659
300	0.358077	0.000175	0.332165	0.38399	0.051825
350	0.310889	0.000114	0.289987	0.33179	0.041803
400	0.266076	0.000107	0.245762	0.28639	0.040628
450	0.226448	9.54653E-05	0.207297	0.245598	0.038301
500	0.195311	7.60164E-05	0.178222	0.2124	0.034177
550	0.175497	5.03960E-05	0.161583	0.189411	0.027828
600	0.14436	7.24981E-05	0.127672	0.161049	0.033377

Here we see that up to 450 hours asymptotic variance of the estimate of survival function and length of confidence interval based on MLE are smaller than that based on non-parametric estimation.

References

1. Boag, J. W. Maximum Likelihood Estimates of Proportion of Patients Cured by Cancer Therapy, *Jr. of royal Statistical Society*, B-11, PP. 15-44, 1949.

2. Peck, D. S. Transistor Failure Studies at Accelerated Levels, *IEEE*, Bell Telephones Laboratories, Pennsylvania, Press, Pacific Grove", C.A., 1966.

3. Mendenhall, W. & Hader, R. J. Estimation of the Parameters of Mixedexponentially Distributed Failure Time Distribution from Censored Life Test Data, Biometrika, 45, PP. 504-520, 1958.

4. Patel, M.N. and Gajjar, A. V. Maximum Likelihood Estimation in Compound Exponential Failure Model with Changing Failure Rates from Type-I Progressively Censored and Group Censored Samples, *Communi. in Statistics, Theo. & Meth.*, Vol.21(1), PP. 2899-2908,1992.

5. Boardman, T. J. & Kendell, P. J. Estimation in Compound Failure Models, *Technometrics*, Vol. 12, PP. 891-908, 1970.

6. Patel, M.N. Progressive Censored Samples from Mixture of Two Exponential distribution with Changing Parameter, *IAPQR Trans.* Vol. 23(2), PP. 83-93,1998.

7. Patel, N.W. and Patel, M.N. Estimation of Parameters of Mixed Geometric Failure Model from Type-I Progressive Group Censored Sample, *IAPQR Trans.*, Vol.28(1), PP. 33-42, 2003.

8. Shah, J.B. and Patel. M.N. Bayes estimation of Parameters of mixed geometric failuremodel from Progressive Type-I grouped Censored Sample, Jr. of Applied Statistics, Vol.36(6), PP. 495-506,(2009).

9. Patel, M.N. and Gajjar K.A. Inference Under Progressive Interval Censoring for geometric Distribution, *Int. Journal of Statistics and Economics*, Vol. 5 (A-10), PP. 21-36, 2010.

10. Arora, S.H.; Bhimani, G.C. and Patel M.N. Estimation Under Progressive Interval Type-I Censoring for Reciprocal Exponential Distribution, *Wulfenia*, Vol. 20(4), PP. 388-405,2013.

11. Odell, P. M., Anderson, K. M. And D'agostino, R. B. Maximum Likelihood Estimation for Interval Censored Data Using a Weibull Based Accelerated Failure Model, *Biometrics*, PP. 48, 951-959,1992.

12. Samuelson, S.O. and Kongerud, J. Interval-Censoring Longitudinal Data of Respiratory of Methods, *Statistics in Medicine*, 13, 1771-1780,1994.

13. Rao, M. B. Interval-censored Type-II Plan - A Knell of the Traditional Type-II plan? In *IISA International conference* 1998, Abstracts, McMaster University, 55,1998.

14. Aggarwala, R. Progressive Interval Censoring: Some Mathematical Results with Applications to Inference, *Communi. in Statistics*, Theo. & Meth., 30(8&9), 1921-1935, 2001.

15. Gajjar, K. A and Patel, M. N. Some Results On Mle of Parameters of Geometric Distribution Under Grouped and Ungrouped Progressive Type-I Censoring, *Jr. of Probability and Statistical Science*, Vol. 10(1), PP. 77-92, 2012.

16. Gadhvi N.K. and Bhimani, G.C. Inference Under Progressive Interval Censoring for Discrete Competing Risk Failure Model, Mathematical Theory & Modelling, Vol. 5(2), PP.153-161,2015.

17. Patel, N. W. and Patel, M. N. Maximum Likelihood Estimation in Compound Geometric Failure Model with Changing Parameters from Type-I Two-Stage Progressively Censored and Group Censored Samples, *Communi. in Statistics, Theo. & Meth.*, Vol. 36(13), PP.2367-2375, 2007.

18. Miller, R. G., Gong, G and Munoz, *A. Survival Analysis*, John Wiley and Sons, NY, 1981, pp.47-51.