Mathematical Theory and Modeling
ISSN 2224-5804 (Paper) ISSN 2225-0522 (Online)
Vol.5, No.2, 2015 www.iiste.org

Inference under Progressive Interval Censoring For Discrete Competing Risk Failure Model

Ms. Neha K. Gadhvi[1*] Dr. G.C.Bhimani[2]
1. H.L. Institute of commerce, Ahmedabad University, Nr. Commerce Six Road, Ahmedabad-380009, Gujarat.
2. Department of Statistics, Saurashtra University, University Road, Rajkot-360005 , Gujarat.
* E-mail of the corresponding author: nehakgadhvi@gmail.com

Abstract
In this paper geometric life time model is considered under competing risks. The causes of failures are assumed independent. Type-I progressive interval censoring scheme is used for inference. Point estimation and confidence intervals based on maximum likelihood and bootstrap methods are also proposed. Non parametric method for estimation and confidence interval for survival function is also considered. A real life example is provided to illustrate the theoretical results.
Keywords: Geometric distribution, competing risks model, maximum likelihood estimator, bootstrap method, survival function, non parametric estimation.

1. Introduction

A problem frequently faced by statistician is the analysis of time to event data such as failure time data and incubation time data. Such data arise in many fields including medicine, engineering, economics and several other fields like: the time to AIDS for HIV positive individuals, the time to death for cancer patients, life of electrical; or electronic component etc.
The failure of an item may be due to single cause or more than one cause. When failure is observed due to multiple causes the analysis of time to event data becomes more complicated. In medical studies or in the analysis of reliability data the failure of individuals or items may be attributable to more than one cause or factor. According to Hoel (1972), during a laboratory experiment, mice were given a dose of radiation of 6 weeks of age. The causes of death were recorded as Thymic Lymphoma, Reticulum Cell Sarcoma, or other. Another example due to Boag (1949) is from a breast cancer patient, where the cause of death was recorded as “cancer” or “other”. The data for these “competing risk models” consist of the failure time and an indicator random variable denoting the specific cause of failure of the individual or item. The causes of failure may be assumed independent or dependent. In most situations the analysis of competing risk data assumes independent causes of failure. Even though the assumption of dependence may be more realistic, there is some concern about the identifiability of the underlying model.
Suppose that a device exhibits k modes (risks) of failures. When the device begins operation, each failure mode simultaneously generate a random life that is independent of the other modes. Thus, in effect, k life times denoted by $X_1, X_2,...,X_k$, simultaneously begin, life time X_i corresponding to the i-th mode of failure.

Failure of the device occurs, as soon as any one of the life times say X_i is realized. Then if the life length of the device is denoted by a random variable X then X = min $\{X_1, X_2,...,X_k\} \equiv X_{(1)}$ and the cumulative

distribution function(cdf) of X, $F_X(x)$ is given by

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

Here each failure mode can have any failure distribution that not all the failure distribution need be alike; but all the k modes operate independently of each other such a model is called competing risk failure model.
In mixture model only one of the k possible modes generates a random life that causes part failure. The different identifiable causes of failures may be attributed to electrical, thermal, climatic sand mechanical stresses applied to an item.
Peck (1966)has considered epitomical transistors used in the telephone industry with two types of failure: electrical degradation of certain parts (cause I) and faulty banding of the leads (cause II). Boardman and Kendell (1970), Mendenhall and Hader(1958), Patel and Gajjar (1992) have considered continuous life time model for mixture or competing risk failure model.
Recently considerable attention and interest is being focuses on the analysis of discrete data. Discrete time data arise when we wish to investigate the ability of electron tubes of withstand successive voltage overloads, the performance of switches which are repeatedly turned on and off or the ability of a mattress to withstand repeated pounding in a torture test.

In each of these causes, failure can occurs at the x-th trial (X = 1,2,....) and it is often assumed that the probability of failure at the x-th trial is equal to some constant 1-q, provided the unit has not failed prior to that trial. Obviously the probability of failure at the x-th trial is given by the geometric distribution with probability function

$$f(x,q) = (1-q)q^{x-1}, \quad x = 1,2,....;\ 0 < q < 1. \tag{2}$$

Such discrete failure time model has been found to be useful in engineering, medical and biological studies. For instance, the number of genes in operon follows geometric distribution
(See: DeHoon et al (2004)). Yaqub and Khan (1981), Patel and Gajjar(1990), Patel and Patel (2006) have considered geometric distribution as a discrete life time model to study a problem of life testing.

2. Competing risk failure model

Suppose that the device exhibits 2 modes of failures and each failure mode simultaneously generate a random life time X_1 and X_2 respectively. i.e. X_i is the time of failure of an item due to cause i then the pmf of X_i is given by

$$f_i(x) = (1-q_i)q_i^{x-1}, \quad x = 1,2,....;\ 0 < q_i < 1,\ \ i = 1,2. \tag{3}$$

And corresponding cdf is

$$F_i(x) = P(X_i \le x) = 1 - q_i^{x}, \quad i = 1,2. \tag{4}$$

Let T be the time of failure of an item regardless of cause, then T = min(X_1, X_2). The pmf of T will be

$$h_T(t) = (1-q_1q_2)(q_1q_2)^{t-1}, \quad t = 1,2,..... \tag{5}$$

With its cdf, $$H_T(t) = P(T \le t) = 1-(q_1q_2)^t, \quad t = 1,2...... \tag{6}$$

Let $P_i(t)$ be the probability that an item failed by cause i at time t and it must not failed by the cause $i' \neq i$ up to time t with two independent causes only. Then

$$P_i(t) = f_{i(t)}[1 - F_{i'}(t)], \quad i' \neq i = 1,2. \tag{7}$$

Hence the probability of failure of an item is given by

$$\begin{aligned} P(t) &= P_1(t) + P_2(t) \\ &= f_1(t)[1-F_2(t)] + f_2(t)[1-F_1(t)] \\ &= (1-q_1)q_1^{t-1}(q_2^t) + (1-q_2)q_2^{t-1}(q_1^t) \\ &= q_2(1-q_1)(q_1q_2)^{t-1} + q_1(1-q_2)(q_1q_2)^{t-1} \end{aligned} \tag{8}$$

Hence the pmf of t regardless of cause of failure defined in (5) can be obtained from (8) after adjusting the normalizing factor C as

$$h_T(t) = CP(t) \tag{9}$$

Such that $$\sum_{t=1}^{\infty} h_T(t) = 1.$$

Which gives, $$C = \frac{1-q_1q_2}{q_2(1-q_1)+q_1(1-q_2)}.$$

Substituting C in (9), we have

$$h_T(t) = \frac{q_2(1-q_1)}{q_2(1-q_1)+q_1(1-q_2)}(1-q_1q_2)(q_1q_2)^{t-1} + \frac{q_1(1-q_2)}{q_2(1-q_1)+q_1(1-q_2)}(1-q_1q_2)(q_1q_2)^{t-1}$$
$$= g_1(t) + g_2(t) \tag{10}$$

The failure model in (10) is called geometric 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 verity of circumstances but are commonly encountered in medical studies, where patients are only monitored at regular interval (e.g. weekly or quarterly checkup). Several authors have discussed application of interval censoring in clinical, medical, biomedical and engineering studies like Odell et al (1992), Samuelson and Kongerud(1994), Scallan(1999), Rao (1998), Aggarwala (2001) etc.

Patel and Patel (2007) have considered progressive grouped censored samples from geometric competing risk failure model with different parameters at each stage of censoring.

Here we consider geometric competing risk failure model and apply progressive type-I interval censoring without changing the parameters at different stages of censoring.

Under progressive type-I interval censored sample the likelihood is proportional to expression

$$L(q_1,q_2) \propto \prod_{i=1}^{m}\left[G_1(N_i)-G_1(N_{i-1})\right]^{x_{1i}} \prod_{i=1}^{m}\left[G_2(N_i)-G_2(N_{i-1})\right]^{x_{2i}} \prod_{i=1}^{m}\left[1-H(N_i)\right]^{r_i} \tag{11}$$

With $N_0=0$, $r_m = n-\sum_{i=1}^{m} x_{1i} - \sum_{i=1}^{m} x_{2i} - \sum_{i=1}^{m-1} r_i$.

Based on $g_1(t)$, defined in (10), we define

$$\begin{aligned} G_1(N) &= P(X_1 \le N) \\ &= \sum_{t=1}^{N} g_1(t) \\ &= \frac{q_2(1-q_1)}{q_2(1-q_1)+q_1(1-q_2)}\left(1-(q_1q_2)^N\right) \end{aligned}$$

Similarly, $$G_2(N) = \frac{q_1(1-q_2)}{q_2(1-q_1)+q_1(1-q_2)}\left(1-(q_1q_2)^N\right)$$

Hence the likelihood function given in (11) becomes

$$L \propto C_1^{\sum_{i=1}^{m} x_{1i}} (q_1q_2)^{\sum_{i=1}^{m} x_{1i}N_{i-1}} C_2^{\sum_{i=1}^{m} x_{2i}} (q_1q_2)^{\sum_{i=1}^{m} x_{2i}N_{i-1}} (q_1q_2)^{\sum_{i=1}^{m} r_iN_i} \times \prod_{i=1}^{m}\left\{(1-(q_1q_2)^{S_i})^{x_{1i}}(1-(q_1q_2)^{S_i})^{x_{2i}}\right\} \tag{12}$$

where $C_1 = \dfrac{q_2(1-q_1)}{q_2(1-q_1)+q_1(1-q_2)}$ and $C_2 = \dfrac{q_1(1-q_2)}{q_2(1-q_1)+q_1(1-q_2)}$.

Hence we get $$\frac{\partial \log L}{\partial q_1} = \frac{-\sum_{i=1}^{m} x_{1i}}{1-q_1} - \frac{B(1-2q_2)}{q_1+q_2-2q_1q_2} + \frac{\sum_{i=1}^{m} x_{2i}}{q_1} + \frac{A}{q_1} - \frac{1}{q_1}\sum_{i=1}^{m}\frac{x_iS_i(q_1q_2)^{S_i}}{1-(q_1q_2)^{S_i}} \tag{13}$$

and, $$\frac{\partial \log L}{\partial q_2} = \frac{\sum_{i=1}^{m} x_{1i}}{q_2} - \frac{B(1-2q_1)}{q_1+q_2-2q_1q_2} - \frac{\sum_{i=1}^{m} x_{2i}}{1-q_2} + \frac{A}{q_2} - \frac{1}{q_2}\sum_{i=1}^{m}\frac{x_iS_i(q_1q_2)^{S_i}}{1-(q_1q_2)^{S_i}}. \tag{14}$$

where $A = \sum_{i=1}^{m} N_{i-1}(x_{1i}+x_{2i}) + \sum_{i=1}^{m} N_ir_i$, $B = \sum_{i=1}^{m}(x_{1i}+x_{2i})$ and $S_i = N_i - N_{i-1}$.

From (13) and (14), after some algebraic manipulation we get

$$A_1q_1^2 + A_2q_1 + A_3 = 0 \tag{15}$$

where

$$A_1 = -(1-q_2)\sum_{i=1}^{m} x_{1i} - q_2\sum_{i=1}^{m} x_{2i}$$

$$A_2 = (1-3q_2+2q_2^2)\sum_{i=1}^{m} x_{1i} + (3q_2-1)\sum_{i=1}^{m} x_{2i} + (1-q_2^2)B$$

and

$$A_3 = q_2(q_2-2)\sum_{i=1}^{m} x_{2i}$$

Solving equation (15) we get $$q_1 = \frac{-A_2 \pm \sqrt{A_2^2-4A_1A_3}}{2A_1} \tag{16}$$

Assuming equal length interval censoring with $N_i = iN$ i.e. $S_i = N$ = length of i-th interval, i=1,2,...,m, and again from (13), we get

$$q_2 = \left[\frac{N\sum_{i=1}^{m} x_i(q_1q_2)^N}{1-(q_1q_2)^N} - A + \frac{\sum_{i=1}^{m} x_{1i}}{1-q_1}\right]\left(\frac{q_1+q_2-2q_1q_2}{B}\right). \tag{17}$$

$$q_2 = \psi(q_2). \tag{18}$$

Solving equation (18) by any method of iteration like: Newton –Raphson etc we get MLE of q_2, say $\hat{q}_2$, and substituting it in (16) we get MLE of q_1, say $\hat{q}_1$,

The survival function at time t_0 is given by

$$S(t_0) = (q_1q_2)^{t_0},$$

whose MLE can be obtained by replacing the parameters by their MLEs

$$\text{i.e.} \quad \hat{S}(t_0) = (\hat{q}_1\hat{q}_2)^{t_0}. \tag{19}$$

4. Standard Errors of the Estimators:

The asymptotic variances and covariances of the MLE for the parameters q_1 and q_2 are given by elements of the inverse of the Fisher information matrix

$$I_{ij} = E\left[-\frac{\partial^2 \log L}{\partial q_i \partial q_j}\right]; \text{ i, j = 1, 2.}$$

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

$$\begin{bmatrix} \dfrac{-\partial^2 \log L}{\partial q_1^2} & \dfrac{-\partial^2 \log L}{\partial q_1 q_2} \\ \dfrac{-\partial^2 \log L}{\partial q_1 q_2} & \dfrac{-\partial^2 \log L}{\partial q_2^2} \end{bmatrix}^{-1}_{(\hat{q}_1,\hat{q}_2)} = \begin{bmatrix} V(\hat{q}_1) & Cov(\hat{q}_1,\hat{q}_2) \\ Cov(\hat{q}_1,\hat{q}_2) & V(\hat{q}_2) \end{bmatrix} \tag{20}$$

where

$$\frac{\partial^2 \log L}{\partial q_1^2} = \frac{-\sum_{i=1}^{m} x_{1i}}{(1-q_1)^2} + \frac{B(1-2q_2)^2}{(q_1+q_2-2q_1q_2)^2} - \frac{\sum_{i=1}^{m} x_{2i}}{q_1^2} - \frac{A}{q_1^2} - \frac{NB}{q_1^2}\left(\frac{(q_1q_2)^N}{1-(q_1q_2)^N}\right)\left(\frac{N}{1-(q_1q_2)^N} - 1\right)$$

$$\frac{\partial^2 \log L}{\partial q_2^2} = \frac{-\sum_{i=1}^{m} x_{1i}}{q_2^2} + \frac{B(1-2q_1)^2}{(q_1+q_2-2q_1q_2)^2} - \frac{\sum_{i=1}^{m} x_{2i}}{(1-q_2)^2} - \frac{A}{q_2^2} - \frac{NB}{q_2^2}\left(\frac{(q_1q_2)^N}{1-(q_1q_2)^N}\right)\left(\frac{N}{1-(q_1q_2)^N} - 1\right) \quad \text{and}$$

$$\frac{\partial^2 \log L}{\partial q_1 q_2} = \frac{B}{(q_1+q_2-2q_1q_2)^2} - \frac{BN^2(q_1q_2)^N}{q_1q_2(1-(q_1q_2)^N)^2}. \tag{21}$$

Hence the asymptotic variance of survival function is given by

$$V(\hat{S}(t_0)) = \left[\left(\frac{\partial S}{\partial q_1}\right)^2 V(\hat{q}_1) + \left(\frac{\partial S}{\partial q_2}\right)^2 V(\hat{q}_2) + 2\left(\frac{\partial^2 S}{\partial q_1 \partial q_2}\right) Cov(\hat{q}_1,\hat{q}_2)\right]_{(\hat{q}_1,\hat{q}_2)} \tag{22}$$

Standard errors of the estimators are nothing but a positive square root of their asymptotic variances.

5. Simulation Algorithm:

Simulation studies will aid in the competing risk modelling of the lifetime distribution and in investigating the behavior of the estimated parameters when samples are collected according to progressive interval censoring schemes. 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 if the progressive interval censoring.

$$X_{11} \approx B(n, G_1(N_1)) \text{ and } X_{21} \approx B(n, G_2(N_1))$$

And for i = 2,3, … , m

$$X_{j,i}\left|X_{j,i-1}, X_{j,i-2},\ldots, X_{j,1}, R_{i-1}, R_{i-2},\ldots, R_1\right.$$

$$\sim \mathrm{B}\left(n-\sum_{s=1}^{i-1}\left(X_{1s}+X_{2s}+R_s\right), \frac{G_j(N_i)-G_j(N_{i-1})}{1-\sum_{s=1}^{i-1}(G_j(N_i)-G_j(N_{i-1}))} \right)$$

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

Here B(n, p) denotes binomial distribution with parameters n and p, $0<p<1$.

On the basis of the algorithm given in Aggarwala (2001) we suggest the following algorithm for simulation for competing risk failure model.

1. Set I = 0, X_1sum = 0, X_2sum = 0, rsum = 0
2. Next i
3. If i = m+1, exit the algorithm
4. Generate X_{11} and X_{12} as binomial random variables with parameters $(n, G_1(N_1))$ and $(n, G_2(N_1))$ respectively.
5. Generate X_{1i} and X_{2i} as binomial variates with parameters (n- X_1sum - X_2sum – Xrsum, $\frac{G_1(N_i)-G_1(N_{i-1})}{1-G_1(N_{i-1})}$) and (n- X_1sum - X_2sum – Xrsum, $\frac{G_2(N_i)-G_2(N_{i-1})}{1-G_2(N_{i-1})}$) respectively.
6. Calculate R_i^{obs} = Floor[p_i(n- X_1sum - X_2sum – Xrsum – X_i)] or min(R_i, n- X_1sum - X_2sum –Xrsum – X_i)
7. Set X_1sum = X_1sum + X_{1i}, X_2sum = X_2sum + X_{2i}, Xrsum = Xrsum + R_i^{obs}
8. Go to step 2

This algorithm generates m binomial random variables. Here either the values p_1, p_2, …. , p_{m-1} or proposed values of R_1, R_2, ……., R_{m-1} are fixed in advance by the experimenter. Here $p_m = 1$ and

$R_m = n-\sum_{i=1}^{m} X_{1i} - \sum_{i=1}^{m} X_{2i} - \sum_{i=1}^{m-1} R_i$.

6. Confidence Interval Estimation:

6.1 Approximate confidence interval

The asymptotic normality of the MLE can be used to compute the approximate confidence intervals for parameters q_1 and q_2 and survival function.

Therefore, (1-α)100% confidence intervals for θ become

$$\theta \pm Z_{\alpha/2}\sqrt{\mathrm{var}(\theta)} \tag{23}$$

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

6.2 Bootstrap confidence intervals

A. Percentile Bootstrap Method

1. From the original data $\underline{x}$ compute the ML estimates of the parameters $\hat{q}_1$ and $\hat{q}_2$ by solving the equations (16) and (17).
2. Use $\hat{q}_1$ and $\hat{q}_2$ to generate a bootstrap sample $\underline{x}^*$ with the same values of N_i, r_i and m;(i = 1,2,….,m.) using the algorithm given in the Section 5.
3. As in step 1, based on $\underline{x}^*$ compute the bootstrap sample estimates of q_1, q_2 and $S(t_0)$, say $\hat{q}_1^*, \hat{q}_2^*$ and $\hat{S}^*(t_0)$.
4. Repeat steps 2-3 S times representing S bootstrap MLE's of (q_1, q_2, and $S(t_0)$) based on S different bootstrap samples.
5. Arrange all $\hat{q}_1^*, \hat{q}_2^*$ and $\hat{S}^*(t_0)$ in an ascending orders to obtain the bootstrap sample $\left(\varphi_l^{[1]}, \varphi_l^{[2]},\ldots,\varphi_l^{[S]}\right)$, l = 1, 2, 3 (where $\varphi_1 \equiv q_1^*$, $\varphi_2 \equiv q_2^*$ and $\hat{S}^*(t_0)$).
6. Let $G(z)=P(\varphi_l \le z)$ be the cumulative distribution function of φ_l. Define $\varphi_{lboot} = G^{-1}(z)$ for

given z. The approximate bootstrap 10(1-α)100% confidence interval of φ_l is given by

$$\left[\varphi_{lboot}(\frac{\alpha}{2}),\ \varphi_{lboot}(1-\frac{\alpha}{2}),\right]$$

B. Bootstrap-t method

1. From the original data x compute the ML estimates of the parameters $\hat{q}_1$ and $\hat{q}_2$ by solving the equations (16) and (17).
2. Use $\hat{q}_1$ and $\hat{q}_2$ to generate a bootstrap sample $\underline{x}^*$ with the same values of N_i, r_i and m;(i = 1,2,….,m.) using the algorithm given in the above method.
3. As in step 1, based on $\underline{x}^*$ compute the bootstrap sample estimates of q_1, q_2 and $S(t_0)$, say $\hat{q}_1^*, \hat{q}_2^*$ and $\hat{S}^*(t_0)$.
4. Compute the following statistics

$$T_1^* = \frac{\sqrt{n}(\hat{q}_1^* - \hat{q}_1)}{\sqrt{Var(\hat{q}_1^*)}},\quad T_2^* = \frac{\sqrt{n}(\hat{q}_2^* - \hat{q}_2)}{\sqrt{Var(\hat{q}_2^*)}} \quad \text{and} \quad T_3^* = \frac{\sqrt{n}(\hat{S}^*(t_0) - \hat{S}(t_0))}{\sqrt{Var(\hat{S}^*(t_0))}}$$

   where var(.) can be obtained using the Fisher information matrix in (20) and (22).
5. Repeat step 3 and 4 S boot times.
6. For T_i^*, i = 1, 2, 3 values obtained in step 4 determine the upper and lower bounds of the 100(1-α)% confidence interval of the parameters and survival function as follows:
   Let $H(x) = P(T_i^* \le x),\ i = 1,2,3$ be the cumulative distribution function of T_i^*. For a given x, define

$$\hat{q}_{1Boot-t}(x) = \hat{q}_1 + \sqrt{n}\sqrt{Var(\hat{q}_1)}H^{-1}(x)$$

The approximate (1-α) 100% confidence interval for q_1 can be constructed as

$$\left(\hat{q}_{1Boot-t}(\alpha/2),\ \hat{q}_{1Boot-t}(1-(\alpha/2))\right)$$

Similarly we can define for other parameters.

7. Non-parametric Estimation of Survival Function and its Confidence Interval Estimation:

Kaplan Meier estimate of survival function $S(t_0)$ can be obtained according to Miller et al (1981) (Page:47, 51) the estimate and estimate of its variance can be obtained as follow.

$$\hat{S}(t_0) = \prod_{Y_{(i)} \le t_0} \left(\frac{n-i}{n-i+1}\right)^{\delta(i)} \tag{24}$$

And,

$$Asy\hat{V}(\hat{S}(t_0)) = \hat{S}^2(t_0) \sum_{Y_{(i)} \le t_0} \frac{\delta(i)}{(n-i)(n-i+1)} \tag{25}$$

Using the results (24)-(25) in (23) (1-α)100% asymptotic confidence interval for $S(t_0)$ can be obtained.

8. APPLICATION:

The following example relates to the two types of failure namely electrical degradation of certain parts (Cause-I) and faulty bounding of the leads(Cause-II) leading to failure of transistors under accelerating testing considered by Peck(1966). The data is considered as censored at time 599 hrs. and 52 out of 369 items were censored at 599 hrs. Here we modify the data including withdrawals as follows.

Table:1. Failure data

Number i	**Time interval (N_{i-1}, N_i)**	**Failures due to Cause-I : x_{1i}**	**Failures due to Cause-II : x_{2i}**	**Withdrawals R_i**
1	0-49	25	14	2
2	50-99	27	15	3
3	100-149	27	20	3
4	150-199	35	10	3
5	200-249	15	10	3
6	250-299	20	07	2
7	300-349	10	06	2
8	350-399	09	05	2
9	400-449	10	03	2
10	450-499	06	04	1
11	500-549	05	01	1
12	550-599	08	02	52

As per our notations we have $N = S_i = N_i - N_{i-1} = 50$, $m = 12$, $N_m = 599$, $N_0 = 0$, $n = 369$, $r_m = 52$, Solving the equations (16) to (18) of section 3 we find MLEs of q_1 and q_2 as

$$\hat{q}_1 = 0.997986862 \quad \text{and} \quad \hat{q}_2 = 0.999017965. \qquad (26)$$

From (19) we get MLE for survival function at time $t_0 = 150$ as

$$\hat{S}(t_0 = 150) = 0.637849587$$

Using (20) the asymptotic variance –covariance matrix of the MLE s for parameter q_1 and q_2 is given by

$$\Sigma = \begin{bmatrix} 2.05478E\text{-}08 & 2.15179E\text{-}11 \\ 2.15179E\text{-}11 & 1.00332E08 \end{bmatrix} = \begin{bmatrix} V(\hat{q}_1) & Cov(\hat{q}_1, \hat{q}_2) \\ Cov(\hat{q}_1, \hat{q}_2) & V(\hat{q}_2) \end{bmatrix}$$

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

$SE(\hat{q}_1) = 0.000143345$ and $SE(\hat{q}_2) = 0.000100166$

Hence from (22) we get $SE(\hat{S}(t_0 = 150) = 0.016771357$

To apply bootstrap confidence interval estimation for the parameters we have made 1000 simulations based on the MLEs of q_1 and q_2 given in (26) and fixing the other values $N = S_i = N_i - N_{i-1} = 50$, $m = 12$, $N_m = 599$, $N_0 = 0$, $n = 369$ of the given real life data.

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

Based on the simulated results the confidence interval based on MLE and bootstrap confidence intervals for parameters and survival function are computed using the methods described in the Section 5.2, which are given in the following table.

Table 1. Confidence intervals

Method	Parmeter	Estimate	Confidence Interval	Length of the Interval
MLE	q_1	0.99798686	(0.99770590, 0.99826781)	(0.00056190)
	q_2	0.99901796	(0.99882163, 0.99921429)	(0.00039265)
	$S(t_0 = 150)$	0.63784956	(0.60497770, 0.67072144)	(0.06574374)
Percentile Bootstrap	q_1	0.99846346	(0.99818577, 0.99869169)	(0.00050592)
	q_2	0.99935030	(0.99918874, 0.99948722)	(0.00029848)
	$S(t_0 = 150)$	0.72046432	(0.68205851, 0.75095647)	(0.06889796)
Bootstrap - t	q_1	0.99846346	(0.99820016, 0.99895763)	(0.00075747)
	q_2	0.99935030	(0.99921542, 0.99972760)	(0.00051218)
	$S(t_0 = 150)$	0.72046432	(0.68392038, 0.77170121)	(0.08778083)

Table 2. Estimate of of survival (Reliability) function ($S(t_0)$). Its asymptotic variance and asymptotic confidence interval 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	
49	0.8633919	5.50077E-05	0.848855142	0.877928662	0.02907352
99	0.7432143	0.000166429	0.717928905	0.768499718	0.050570813
149	0.6397645	0.000279431	0.607000791	0.672528266	0.065527475
199	0.5507142	0.000369466	0.513040027	0.58838831	0.075348283
249	0.4740589	0.000428803	0.433472108	0.514645776	0.081173668
299	0.4080735	0.000458376	0.366110498	0.450036583	0.083926086
349	0.3512728	0.000463003	0.309098503	0.393447107	0.084348603

Table 3. Estimate of survival (Reliability) function ($S(t_0)$). Its asymptotic variance and asymptotic confidence interval using nonparametric estimation:

t_0	$\hat{S}(t_0)$	Asy V($\hat{S}(t_0)$)	95% Confidence interval for $\hat{S}(t_0)$		Difference of limits of CI
			Lower limit	Upper limit	
49	0.89431	2.55930E-04	0.862954288	0.925665712	0.062711424
99	0.78252	2.66980E-04	0.750494536	0.814545464	0.064050929
149	0.65302	2.97650E-04	0.619205029	0.686834971	0.067629941
199	0.52744	2.83200E-04	0.494456048	0.560423952	0.065967905
249	0.45902	1.75890E-04	0.433025789	0.485014211	0.051988423
299	0.38058	1.94160E-04	0.353269104	0.407890896	0.054621793
349	0.33338	1.22000E-04	0.311731092	0.355028908	0.043297815
399	0.29133	1.10360E-04	0.270739736	0.311920264	0.041180528
449	0.25146	1.05520E-04	0.231326306	0.271593694	0.040267388
499	0.22003	1.12910E-04	0.199203213	0.240856787	0.041653574
549	0.2009	5.57100E-05	0.18627073	0.21552927	0.02925854
599	0.1685	8.86400E-05	0.150046832	0.186953168	0.036906337

Here we see that from time 199 and onwards asymptotic variance of the estimate of survival function and length of confidence interval based on non parametric estimation are smaller than that of based on MLE.

REFERNCES

Aggarwala, R. (2001), “Progressive interval censoring: Some mathematical results with applications to inference, Communi. in Statistics”, Theo. & Meth., 30(8&9), 1921-1935.

Boag, J. W. (1949), “Maximum likelihood estimates of proportion of patients cured by Cancer Therapy”, Jr. of royal Statistical Society, B-11, 15-44.

Boardman, T. J. & Kendell, P. J. (1970), “Estimation in compound failure models”, Technometrics, 12, 891-908.

DeHoon, M.J.L., Imoto, S., Kobayashi, K., Ogasawara, N. and Miyano, S. (2004), “Predicting the operon structure of bacillus subtilis using operan length, intergene distance, and gene expression information”, Pacific Symposium on Biocomputing, 9, 276-287.

Hoel, D. G. (1972), “A representation of mortality data by competing risks”, Biometrics, 28, 475-488

Mendenhall, W. & Hader, R. J. (1958), “Estimation of the parameters of mixed exponentially distributed failure time distribution from censored life test data”, Biometrika, 45, 504-520.

Miller, R.G.; Gong, G. & Munoz, A. (1981), “Survival Analysis, John Wiley & Sons”, New York.

Odell, P. M., Anderson, K. M. & D’agostino, R. B. (1992), “Maximum likelihood estimation for interval censored data using a Weibull based accelerated failure model”, Biometrics, 48, 951-959.

Patel, M.N. & Gajjar, A.V. (1990), “Progressively censored samples from geometric distribution”,The Aligarh Journal of Statistics, Vol.10, 1-8.

Patel, M. N. & Gajjar, A. V. (1992), “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., 21(1), 2899-2908.

Patel, N.W. & Patel, M.N. (2006), “Some probabilistic properties of geometric life time model”, International Journal of Management and Systems, 22(2), 1-3.

Patel, N.W. & Patel, M.N. (2007), "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.,36, 2367-2375**.**

Peck, D. S. (1966), "Transistor failure studies at accelerated levels, IEEE, Bell Telephones Laboratories, Pennsylvania, Press, Pacific Grove", C.A.

Rao, M. B. (1998), "Interval-censored Type-II plan - a knell of the traditional Type-II plan?", In IISA International conference 1998, Abstracts, McMaster University, 55.

Samuelson, S.O. & Kongerud, J. (1994), "Interval-censoring longitudinal Data of respiratory of methods", Statistics in Medicine, 13, 1771-1780.

Scallan, A.J. (1999), "Regression modeling of interval –censored failure time data using the Weibull distribution", Journal of Applied Statistics, 26(5), 613-618.

Yaqub, M. & Khan, A. H. (1981), "Geometric failure low in life testing, Pure and Applied Mathematika Sciences", Vol. XIV, 1(2), 69-76.

Neha K. Gadhvi (MSc., MPhill.) is a doctoral student in the department of statistics at Saurastra University, Rajkot, Gujarat, India. She is currently working as an Lecturer at H.L. Institute of Commerce, Ahmedabad University, Gujarat.

Dr. G.C.Bhimani (MSc.,Ph.D.) is a professor of statistics at the department of statistics at Saurashtra University, Rakjot, Gujarat, India. Till date he has guided more than five students for the doctoral degree.