On classical inference of a flexible semi-parametric class of distributions under a joint balanced progressive censoring scheme

Dhrubasish Bhattacharyya [a], Debasis Kundu[b]

[a]*Department of Mathematical Sciences, Rajiv Gandhi Institute of Petroleum Technology, Jais, Amethi - 229304, Uttar Pradesh, India.*
[b]*Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India.*

Abstract

The paper deals with the estimation procedures for proportional hazard class of distributions under a two-sample balanced joint progressive censoring scheme. The baseline hazard function is assumed to be piecewise constant, instead of any specific form. This adds flexibility to the proposed model and the shape of the underlying hazard function is completely data-driven. Since the complicated form of the likelihood function does not yield closed form estimators, we propose a variant of Expectation-Maximization algorithm, known as Expectation Conditional Maximization (ECM) algorithm for obtaining maximum likelihood estimates of the model parameters. This leads to explicit expressions for the iterative constrained maximization steps of the algorithm. Extension to the case when the cut points are unknown, has also been considered for dealing with problems involving real data. Simulation results and illustrations using real data have also been presented.

Keywords: Expectation-Maximization algorithm, Expectation Conditional Maximization algorithm, Maximum likelihood estimators, Bootstrap confidence interval.
AMS 2010 classification code : 62N01, 62N02, 62F10.

1. Introduction

In order to make life testing experiments cost-effective and less time-consuming, it is a standard practice to employ different censoring schemes. The most commonly used censoring schemes are Type-I and Type-II. In Type-I censoring, the experiment is terminated at a pre-specified time. This has the possibility that no failure is observed during the time. To ensure a certain number of failures in a life testing experiment, Type-II censoring had been introduced in the literature where the experiment is terminated after a specific number of failures are observed. During the last two decades, the progressive censoring scheme has received a widespread attention in statistical literature. The progressive censoring schemes add flexibility to other censoring schemes by allowing withdrawals of items during experiment. The Type-II progressive censoring is generalization of Type-II censoring. For a comprehensive review on progressive censoring schemes and related issues, one may refer to Balakrishnan

Email addresses: dhrubasish018@gmail.com (Dhrubasish Bhattacharyya), kundu@iitk.ac.in (Debasis Kundu)

and Cramer (2014). Although extensive work has been done on different aspects of the progressive censoring schemes for one-sample, not much work has been done related to two-sample problems. Rasouli and Balakrishnan (2010) introduced the joint progressive Type-II censoring for two samples. Recently, Mondal and Kundu (2019b) have introduced a new balanced joint progressive censoring (BJPC) which is quite useful to compare the lifetime distribution of products from different units which are being manufactured by two different lines in the same facility. They also provided exact inference procedures for two exponential populations under balanced joint progressive censoring schemes. It is pertinent to mention that the BJPC scheme is a generalization of Self Reallocated Design by Srivastava (1987). Moreover, It produces more efficient estimators in less time, see Mondal and Kundu (2019a) for details. We now briefly describe the BJPC scheme. Suppose two independent samples of respective sample sizes m and n from two populations are put on a life testing experiment simultaneously. Let $k < \min(m, n)$ be the total number of failures to be observed and let $R_1, \ldots, R_{k-1}$ be the nonnegative integers such that $\sum_{i=1}^{k-1}(R_i + 1) < \min(m, n)$. We also define the random variables Z_i, $i = 1, \ldots, k$ such that $Z_j = 1$, if the j-th failure occurs from Population 1 and $Z_j = 0$, otherwise. Then BJPC proposed by Mondal and Kundu (2019b), can be described as follows. For $j = 1, \ldots, k-1$, at the j-th failure, $R_j + 1 - Z_j$ items are randomly chosen from remaining items of Population 1 and are removed from the study, and simultaneously $R_j + Z_j$ items are randomly chosen from remaining items of Population 2 and are removed from the study. Finally, at the time of k-th failure, all the remaining items are withdrawn from the experiment. Hence for the BJPC scheme, the observations will be of the form $(\boldsymbol{W}, \boldsymbol{Z})$, where $\boldsymbol{W} = (W_1, \ldots, W_k)$, $W_1 \leq \ldots \leq W_k$ are the observed failure times and $\boldsymbol{Z} = (Z_1, \ldots, Z_k)$.

In this article, we have considered the estimation procedures for a general semi-parametric family of distributions, namely for Proportional Hazard Class, under the BJPC schemes. Let us recall that a class of distribution functions $\{F(x;\theta), \theta > 0\}$, is said to be a proportional hazard class of distributions if its survival function $S(x;\theta)$ can be written as $S(x;\theta) = (S_0(x))^{\theta}$. Here $S_0(x)$ is known as the baseline hazard function. Throughout this article we will denote this proportional hazard class by PHC(S_0, θ). Note that if $f_0(x)(f(x;\theta))$ and $h_0(x)(h(x;\theta))$ are the probability density function (pdf) and hazard rate function respectively, associated with $S_0(x)(S(x;\theta))$, then it follows that $h(x;\theta) = \theta h_0(x)$ and $f(x;\theta) = \theta \{S_0(x)\}^{\theta} h_0(x)$. The assumption of a proportional hazard rate is quite reasonable for the lifetime distribution of products manufactured by two different lines within the same facility. This assumption holds especially true when the causes of failure for different items are of the same kind. In this context, we propose estimation procedures without making any specific assumptions about the baseline hazard function. Instead, we assume that the baseline hazard function belongs to a class of distributions characterized by piecewise constant hazard rates. This assumption introduces significant flexibility into our model, making it adaptable to a wide range of real-world scenarios. A key strength of our proposed model is its data-driven approach to the shape of the failure rate function. We do not impose any predefined form, such as an increasing or decreasing failure rate. Our model also generalizes the work of Mondal and Kundu (2019b), extending their constant baseline hazard function to a more versatile piecewise constant form. We consider the Maximum Likelihood Estimation (MLE) of the model parameters assuming that the number of cut points as well as the positions of the cut points are known. Although exact inferential procedures are car-

ried out for exponential populations, our model's complicated likelihood function precludes such exact inferential procedures. The intricate form of the likelihood function yields no closed form expression of the MLEs, necessitating the solution of a multi-dimensional, complex optimization problem. Standard numerical optimization methods, which are sensitive to initial guesses, become impractical as the problem's dimensionality increases. Moreover, the Expectation Maximization (EM) algorithm is not suitable because it fails to produce an explicit M-step for the proposed model. So, in order to circumvent this issue, we propose an Expectation Conditional Maximization algorithm (ECM), which belongs to generalized Expectation Maximization algorithms, in order to obtain the MLEs. This approach provides explicit expressions for the iterative constrained maximization steps needed to determine the MLEs of the model parameters effectively. Recognizing that the assumption of known cut points is rare in practical applications, we also introduce a methodology that accommodates unknown cut points. This advancement makes our model highly practical and easy to implement for real-world data problems, significantly enhancing its utility and applicability.

The Expectation Maximization (EM) algorithm introduced by Dempster (1977) is a two-step iterative procedure used to obtain MLEs of unknown parameters of a likelihood function. The first step, i.e., the E-step, involves computation of the expected complete log likelihood function conditioned based on knowledge of the data and current estimates of the parameters. In the second step, i.e., in the M-step, this log-likelihood function is maximized with respect to the unknown parameters. This two-step procedure is repeated until it converges. The main important feature of the EM algorithm is that it enables one to estimate parameters under incomplete observations, e.g., when the observations are censored. In spite of its widespread applicability, the EM algorithm often leads to complicated and unattractive maximization step. In order to circumvent this issue, Meng and Rubin (1993) introduced a version of EM algorithm, known as the Expectation Conditional Maximization (ECM) algorithm where the single M-step is replaced with a series of maximization steps (known as CM-steps) subject to constraints. In the proposed model, the M-step also leads us to complicated M-step. So, we choose to employ the conditional M-steps which yields two iterative constrained maximization steps.

The organization of the rest of the paper is as follows. In Section 2, we obtain expression of the likelihood function under the balanced joint progressive censoring. The complicated form of the likelihood function does not yield closed form expressions for the MLEs. So in Section 3, we propose an efficient ECM algorithm for obtaining the MLEs with the assumption that the number of cut points as well as the positions of the cut points are known. The convergence properties are also shown. In Section 4, we carry out a simulation study in order to assess the performance of the proposed estimators. Section 5 deals with a methodology capable of incorporating the assumption of unknown cut point positions, making it effective for addressing problems involving real data. Illustrative examples involving real data sets are presented in Section 6. Finally we conclude the paper in Section 7.

2. Maximum Likelihood Estimation

Here we assume that $X_1, \ldots, X_m$ denote the lifetimes of the items of Population 1 and $Y_1, \ldots, Y_n$ denote the lifetimes of the items of Population 2. We also assume that $X_1, \ldots, X_m \overset{\text{iid}}{\sim} \text{PHC}(S_0, \theta_1)$ and $Y_1, \ldots, Y_n \overset{\text{iid}}{\sim} \text{PHC}(S_0, \theta_2)$. Now we will discuss the maxi-

mum likelihood estimation of the model parameters based on the piecewise constant hazard function of S_0. The piecewise constant baseline hazard function can be written as

$$h_0(t) = \sum_{j=1}^{M} c_j I_{[\tau_{j-1},\tau_j)}(t), \text{ where } I_{[\tau_{j-1},\tau_j)}(t) = \begin{cases} 1, & \text{if } t \in [\tau_{j-1}, \tau_j) \\ 0, & \text{otherwise} \end{cases}$$

and $0 = \tau_0 < \tau_1 < \cdots < \tau_{M-1} < \tau_M = \infty$ are known cut points, and the c_j's are unknown quantities. In order to avoid identifiability issue, we assume $c_M = 1$ without loss of generality. Observe that $M = 1$ corresponds to the exponential model. The cumulative hazard function of the baseline distribution can be obtained as

$$H_0(t) = \sum_{j=1}^{M} c_j \left\{\min(t, \tau_j) - \tau_{j-1}\right\} I_{[\tau_{j-1},\tau_M)}(t).$$

The survival function $S_0(t)$ and the pdf $f_0(t)$ can then be obtained from the following formulas

$$S_0(t) = e^{-H_0(t)} \text{ and } f_0(t) = h_0(t)e^{-H_0(t)}.$$

For a given sampling scheme m, n, k and $R_1, \ldots, R_{k-1}$, the likelihood function based on the observations $(\boldsymbol{w}, \boldsymbol{z})$ can be written as

$$\begin{aligned} L(\boldsymbol{\theta}, \boldsymbol{c}|\boldsymbol{w}, \boldsymbol{z}) =& C \prod_{i=1}^{k-1} \left\{ f_X(w_i)^{z_i} f_Y(w_i)^{1-z_i} \left\{S_X(w_i)\right\}^{R_i+1-z_i} \left\{S_Y(w_i)\right\}^{R_i+z_i} \right\} \times \\ & f_X(w_k)^{z_k} f_Y(w_k)^{1-z_k} \left\{S_X(w_k)\right\}^{m-\sum_{j=1}^{k-1}(R_j+1)-z_k} \left\{S_Y(w_k)\right\}^{n-\sum_{j=1}^{k-1}(R_j+1)-1+z_k}, \end{aligned}$$

where $C = \prod_{i=1}^{k} \left[\left(m - \sum_{j=1}^{i-1}(R_j+1) \right) z_i + \left(n - \sum_{j=1}^{i-1}(R_j+1) \right) (1 - z_i) \right]$ is the normalizing constant. After simplification, the log-likelihood function can be written as

$$\begin{aligned} \log L(\boldsymbol{\theta}, \boldsymbol{c}|\boldsymbol{w}, \boldsymbol{z}) \propto\ & m_k \log\theta_1 + n_k \log\theta_2 - (\theta_1 + \theta_2) \sum_{i=1}^{k-1} (R_i + 1) H_0(w_i) + \sum_{i=1}^{k} \log h_o(w_i) \\ & - \theta_1 \left(m - \sum_{j=1}^{k-1}(R_j+1) \right) H_0(w_k) - \theta_2 \left(n - \sum_{j=1}^{k-1}(R_j+1) \right) H_0(w_k), \end{aligned}$$

where $m_k = \sum_{i=1}^{k} z_i$ and $n_k = \sum_{i=1}^{k}(1 - z_i) = k - m_k$.

Note that one needs to solve a $(M + 1)$ dimensional nonlinear optimization problem in order to obtain the maximum likelihood estimates of the parameters. Standard optimization methods such as Newton-Raphson method may be used to solve the problem. However, these methods have some disadvantages, specifically, these methods are quite sensitive to initial guesses. Consequently, it is hard to choose the initial guesses particularly if M is large. In order to address this issue, we propose an efficient Expectation Conditional Maximization algorithm for obtaining the MLEs of the parameters.

3. The Expectation Conditional Maximization Algorithm

We now treat the BJPC model as an incomplete data problem and propose an Expectation Conditional Maximization (ECM) algorithm in order to obtain the maximum likelihood estimators of the unknown parameters. The observed data is denoted by the vector $\boldsymbol{W} = (W_1, \ldots, W_k)$. We denote the missing data by $\boldsymbol{M} = \left(\boldsymbol{M}^X, \boldsymbol{M}^Y\right)$, where $\boldsymbol{M}^X = \left(\boldsymbol{M}_1^X, \ldots, \boldsymbol{M}_k^X\right)$ is the vector of missing data from Population 1 and $\boldsymbol{M}^Y = \left(\boldsymbol{M}_1^Y, \ldots, \boldsymbol{M}_k^Y\right)$ correspond to the missing data from Population 2. Note that $\boldsymbol{M}_i^X$ is a $1 \times (R_i + 1 - Z_i)$ vector with $\boldsymbol{M}_i^x = \left(M_{i,1}^X, \ldots, M_{i,R_i+1-Z_i}\right)$ for $i = 1, \ldots, k-1$ and $\boldsymbol{M}_k^X$ is a $1 \times \left(m - \sum_{j=1}^{k-1}(R_j + 1) - z_k\right)$ vector with $\boldsymbol{M}_k^x = \left(M_{k,1}^X, \ldots, M_{k,m-\sum_{j=1}^{k-1}(R_j+1)-Z_k}^X\right)$. Similarly, $\boldsymbol{M}_i^Y$ is a $1 \times (R_i + Z_i)$ vector with $\boldsymbol{M}_i^Y = \left(M_{i,1}^Y, \ldots, M_{i,R_i+Z_i}\right)$ for $i = 1, \ldots, k-1$ and $\boldsymbol{M}_k^Y$ is a $1 \times \left(n - \sum_{j=1}^{k-1}(R_j + 1) - 1 + Z_k\right)$ vector with $\boldsymbol{M}_k^Y = \left(M_{k,1}^Y, \ldots, M_{k,m-\sum_{j=1}^{k-1}(R_j+1)-Z_k}^Y\right)$. The complete data can be obtained by combining $\boldsymbol{W}$ and $\boldsymbol{M}$ and is given by $\boldsymbol{Y} = (\boldsymbol{W}, \boldsymbol{M}) = \left(\boldsymbol{W}, \boldsymbol{M}^X, \boldsymbol{M}^Y\right)$. The following theorem which as an extension of the result of Ng et al. (2002), provides the conditional distribution of missing data $\boldsymbol{M}^X$ and $\boldsymbol{M}^Y$ given the the observations $(\boldsymbol{w}, \boldsymbol{z})$.

Theorem 3.1. *For any fixed $i = 1, \ldots, k$, the random variables $M_{i,j}^X$ $(M_{i,j}^Y)$ and $M_{i,k}^X$ $(M_{i,k}^Y)$ are conditionally independent and identically distributed for all $j \neq k$, given the observations $(\boldsymbol{w}, \boldsymbol{z})$ on $(\boldsymbol{W}, \boldsymbol{Z})$. Also, the conditional distributions of $M_{i,j}^X$ and $M_{i,j}^Y$ are given by*

$$f_{M^X|\boldsymbol{W},\boldsymbol{Z}}\left(M^X = m | W_1 = w_1, \ldots, W_k = w_k\right) = \frac{f_X(m)}{S_X(w_i)}, \ m > w_i$$

and

$$f_{M^Y|\boldsymbol{W},\boldsymbol{Z}}\left(M^Y = m | W_1 = w_1, \ldots, W_k = w_k\right) = \frac{f_Y(m)}{S_Y(w_i)}, \ m > w_i$$

respectively.

Proof: The proof closely follows the arguments used in Ng et al. (2002) and hence omitted for conciseness. □

Now we will briefly review the ECM algorithm which will be useful for further developments of the article. Suppose $\boldsymbol{X} = (X_1, \ldots, X_m)$ is a set random observations from a distribution with pdf $f(x; \theta)$ where $\theta \in \mathbb{R}^d$ is a set of unknown parameters. Further suppose that $\boldsymbol{Z} = (Z_1, \ldots, Z_n)$ is another set of incomplete observations on an unobservable random variable Z and $\boldsymbol{Y} = (\boldsymbol{X}, \boldsymbol{Z})$ is the set of complete observations. Let $f(\boldsymbol{y}; \theta) = f(\boldsymbol{x}, \boldsymbol{z}; \theta)$ be the pdf governing $\boldsymbol{Y}$ and $f(\boldsymbol{z}|\boldsymbol{x}; \theta)$ be the conditional pdf of $\boldsymbol{Z}|\boldsymbol{X}$. Then the EM algorithm consists of iterations of the following two steps:

(i) The E-step computes the expected value of the log-likelihood function of the complete data with respect to the conditional pdf $f(\boldsymbol{z}|\boldsymbol{x}; \theta^{(h)})$, i.e.,

$$Q\left(\theta|\boldsymbol{x}; \theta^{(h)}\right) = \int \left\{\log f(\boldsymbol{x}, \boldsymbol{z}; \theta)\right\} f(\boldsymbol{z}|\boldsymbol{x}; \theta^{(h)}) d\boldsymbol{z}.$$

(ii) The M-step finds $\theta^{(h+1)} = \arg\max_{\theta} Q\left(\theta|\boldsymbol{x}; \theta^{(h)}\right)$.

Here $\theta^{(1)}, \ldots, \theta^{(h)}$ are the solutions obtained from h iterations and $\theta^{(0)}$ is the initial guess of θ. The ECM algorithm replaces the the single M-step (ii) with a number of conditional M-steps, see Meng and Rubin (1993) for details. The ECM algorithm is a generalized EM algorithm, see Neal and Hinton (1998) for details. We now proceed to obtain the ECM algorithm in order to obtain the MLEs of the proposed model.

3.1. The Expectation (E)-Step

Suppose $\lambda(\boldsymbol{\theta}, \boldsymbol{c}|\boldsymbol{Y})$ denote the log-likelihood function corresponding to the complete data. Then the E-step of the ECM algorithm requires the computation of the expectation of $\lambda(\boldsymbol{\theta}, \boldsymbol{c}|\boldsymbol{Y})$, with respect to the conditional distribution of $\boldsymbol{M}$, given $\boldsymbol{W}$ and the current estimate of the parameters $\boldsymbol{\theta^{(h)}}, \boldsymbol{c^{(h)}}$. Under the proposed model, the expectation of $\lambda(\boldsymbol{\theta}, \boldsymbol{c}|\boldsymbol{Y})$, given $\boldsymbol{W}$ and $\boldsymbol{\theta^{(h)}}, \boldsymbol{c^{(h)}}$ can be written as

$$\begin{aligned} Q\left(\boldsymbol{\theta}, \boldsymbol{c}|\boldsymbol{\theta}^{(h)}, \boldsymbol{c}^{(h)}\right) &= m \log \theta_1 + n \log \theta_2 - \theta_1 \left\{ z_i H_0(w_i) + \sum_{i=1}^{k} m_R^i E\left[H_0(M_i^X)\right] \right\} \\ &\quad - \theta_2 \left\{ (1 - z_i) H_0(w_i) + \sum_{i=1}^{k} n_R^i E\left[H_0(M_i^Y)\right] \right\} + \sum_{i=1}^{k} \log h_0(w_i) \\ &\quad + \sum_{i=1}^{k} m_R^i E\left[\log h_0(M_i^X)\right] + \sum_{i=1}^{k} n_R^i E\left[\log h_0(M_i^Y)\right], \end{aligned} \tag{1}$$

where $m_R^i = R_i + 1 - z_i$, $n_R^i = R_i + z_i$ for $i = 1, \ldots, k-1$ and $m_R^k = m - \sum_{J=1}^{k-1}(R_j + 1) - z_k$, $n_R^k = n - \sum_{J=1}^{k-1}(R_j + 1) - 1 + z_k$. Also note that the conditional distributions of M_i^X and M_i^Y, given $\boldsymbol{w}$ and the current estimates of the parameters $\boldsymbol{\theta}^{(h)}, \boldsymbol{c}^{(h)}$ are given by

$$f_{M_i^X|\boldsymbol{w}, \boldsymbol{\theta}^{(h)}, \boldsymbol{c}^{(h)}}(x) = \frac{\theta_1^{(h)}}{S_X^{(h)}(w_i)} h_0^{(h)}(x) e^{-\theta_1^{(h)} H_0^{(h)}(x)}, \;\; x > w_i$$

and

$$f_{M_i^Y|\boldsymbol{w}, \boldsymbol{\theta}^{(h)}, \boldsymbol{c}^{(h)}}(x) = \frac{\theta_2^{(h)}}{S_Y^{(h)}(w_i)} h_0^{(h)}(x) e^{-\theta_2^{(h)} H_0^{(h)}(x)}, \;\; x > w_i$$

respectively. We now proceed to obtain the expression for $Q\left(\boldsymbol{\theta}, \boldsymbol{c}|\boldsymbol{\theta}^{(h)}, \boldsymbol{c}^{(h)}\right)$.

If $w_i \in [\tau_{j-1}, \tau_j)$, $j = 1, \ldots, M$, then $E\left[H_0(M_i^X)\right]$ can be written as

$$\begin{aligned} E\left[H_0(M_i^X)\right] &= \frac{\theta_1^{(h)}}{S_X^{(h)}(w_i)} \left[c_j^{(h)} \int_{w_i}^{\tau_j} H_0(x) e^{-\theta_1^{(h)} H_0^{(h)}(x)} dx + \sum_{k=j}^{M-1} c_{k+1}^{(h)} \int_{\tau_k}^{\tau_{k+1}} H_0(x) e^{-\theta_1^{(h)} H_0^{(h)}(x)} dx \right] \\ &= \frac{\theta_1^{(h)}}{S_X^{(h)}(w_i)} \left[I_1 + I_2 + \cdots + I_{M-j+1} \right], \text{ say.} \end{aligned}$$

Here $H_0^{(h)}(x)$ and $S_X^{(h)}(x)$ are cumulative hazard function and survival function respectively, corresponding to the current estimates of the parameters $\boldsymbol{\theta} = \boldsymbol{\theta}^{(h)}$, $\boldsymbol{c} = \boldsymbol{c}^{(h)}$. Now integration

by parts followed by simplification yields

$$I_1 = -\frac{1}{\theta_1^{(h)}}\left[\left\{H_0(\tau_j)S_X^{(h)}(\tau_j) - H_0(w_i)S_X^{(h)}(w_i)\right\} + \frac{c_j}{\theta_1^{(h)}c_j^{(h)}}\left\{S_X^{(h)}(\tau_j) - S_X^{(h)}(w_i)\right\}\right],$$

$$\begin{aligned} I_k = -\frac{1}{\theta_1^{(h)}}\Bigg[&\left\{H_0(\tau_{j+k-1})S_X^{(h)}(\tau_{j+k-1}) - H_0(\tau_{j+k-2})S_X^{(h)}(\tau_{j+k-2})\right\} \\ &+ \frac{c_{j+k-1}}{\theta_1^{(h)}c_{j+k-1}^{(h)}}\left\{S_X^{(h)}(\tau_{j+k-1}) - S_X^{(h)}(\tau_{j+k-2})\right\}\Bigg] \end{aligned}$$

for $k = 2, \ldots, M-j$; and

$$I_{M-j+1} = \frac{1}{\theta_1^{(h)}}\left[H_0(\tau_{M-1})S_X^{(h)}(\tau_{M-1}) + \frac{c_M}{\theta_1^{(h)}c_M^{(h)}}S_X^{(h)}(\tau_{M-1})\right].$$

Further simplification shows that

$$E\left[H_0(M_i^X)\right] = \frac{1}{S_X^{(h)}(w_i)}\left[H_0(w_i)S_X^{(h)}(w_i) - \sum_{r=j}^{M}\frac{c_r}{\theta_1^{(h)}c_r^{(h)}}\left\{S_X^{(h)}(\tau_r) - S_X^{(h)}\left(\max(w_i, \tau_{r-1})\right)\right\}\right].$$

Hence it follows that

$$E\left[H_0(M_i^X)\right] = \frac{1}{S_X^{(h)}(w_i)}\sum_{j=1}^{M}\left[H_0(w_i)S_X^{(h)}(w_i) - \sum_{r=j}^{M}\frac{c_r}{\theta_1^{(h)}c_r^{(h)}}\left\{S_X^{(h)}(\tau_r) - S_X^{(h)}\left(\max(w_i, \tau_{r-1})\right)\right\}\right] I_{[\tau_{j-1},\tau_j)}(w_i).$$

Similar calculations show that

$$E\left[H_0(M_i^Y)\right] = \frac{1}{S_Y^{(h)}(w_i)}\sum_{j=1}^{M}\left[H_0(w_i)S_Y^{(h)}(w_i) - \sum_{r=j}^{M}\frac{c_r}{\theta_2^{(h)}c_r^{(h)}}\left\{S_Y^{(h)}(\tau_r) - S_Y^{(h)}\left(\max(w_i, \tau_{r-1})\right)\right\}\right] I_{[\tau_{j-1},\tau_j)}(w_i).$$

If $w_i \in [\tau_{j-1}, \tau_j)$, $j = 1, \ldots, M$, then it follows that

$$\begin{aligned} E\left[\log h_0(M_i^X)\right] &= \frac{\theta_1^{(h)}}{S_X^{(h)}(w_i)}\left[c_j^{(h)}\log c_j\int_{w_i}^{\tau_j} e^{-\theta_1^{(h)}H_0^{(h)}(x)}dx + \sum_{k=j}^{M-1}c_{k+1}^{(h)}\log c_{k+1}\int_{\tau_k}^{\tau_{k+1}} e^{-\theta_1^{(h)}H_0^{(h)}(x)}dx\right] \\ &= \frac{\theta_1^{(h)}}{S_X^{(h)}(w_i)}\left[I_1 + I_2 + \cdots + I_{M-j+1}\right], \text{ say.} \end{aligned}$$

Again integration followed by simplification yields $I_1 = -\log c_j\left\{S_X^{(h)}(\tau_j) - S_X^{(h)}(w_i)\right\}/\theta_1^{(h)}$, $I_k = -\log c_{j+k-1}\left\{S_X^{(h)}(\tau_{j+k-1}) - S_X^{(h)}(\tau_{j+k-2})\right\}/\theta_1^{(h)}$ for $k = 2, \ldots, M-j$; and $I_{M-j+1} =$

$S_X^{(h)}(\tau_{M-1}) \log c_M / \theta_1^{(h)}$. Further simplification yields

$$E\left[\log h_0(M_i^X)\right] = -\frac{1}{S_X^{(h)}(w_i)} \sum_{r=j}^{M} \log c_r \left\{ S_X^{(h)}(\tau_r) - S_X^{(h)}\left(\max(w_i, \tau_{r-1})\right) \right\} \text{ if } w_i \in [\tau_{j-1}, \tau_j)\,.$$

Hence it follows that

$$E\left[\log h_0(M_i^X)\right] = -\frac{1}{S_X^{(h)}(w_i)} \sum_{j=1}^{M} \sum_{r=j}^{M} \left[\log c_r \left\{ S_X^{(h)}(\tau_r) - S_X^{(h)}\left(\max(w_i, \tau_{r-1})\right) \right\}\right] I_{[\tau_{j-1}, \tau_j)}(w_i).$$

Similarly, it follows that

$$E\left[\log h_0(M_i^Y)\right] = -\frac{1}{S_Y^{(h)}(w_i)} \sum_{j=1}^{M} \sum_{r=j}^{M} \left[\log c_r \left\{ S_Y^{(h)}(\tau_r) - S_Y^{(h)}\left(\max(w_i, \tau_{r-1})\right) \right\}\right] I_{[\tau_{j-1}, \tau_j)}(w_i)$$

Substituting these quantities in (1), one can obtain $Q\left(\boldsymbol{\theta}, \boldsymbol{c} | \boldsymbol{\theta}^{(h)}, \boldsymbol{c}^{(h)}\right)$ in the E-step of the proposed algorithm.

3.2. The Conditional Maximization (CM)-Steps

We now formalize the CM-steps for the proposed model. Let $\boldsymbol{\phi} = (\boldsymbol{\theta}, \boldsymbol{c})$ and suppose that we have the solutions $\boldsymbol{\phi^{(0)}}, \ldots \boldsymbol{\phi^{(h)}}$ from h iterations, where $\boldsymbol{\phi}^{(h)} = (\boldsymbol{\theta}^{(h)}, \boldsymbol{c}^{(h)})$. We consider the constraint sets Θ_s as given by

$$\Theta_s\left(\boldsymbol{\phi}^{(h+(s-1)/2)}\right) := \left\{\boldsymbol{\phi} : g_s(\boldsymbol{\phi}) = g_s\left(\boldsymbol{\phi}^{(h+(s-1)/2)}\right)\right\},\ s = 1, 2;$$

where $g_s(\boldsymbol{\phi})$, $s = 1, 2$ are preselected vector valued functions of ϕ. In our case, $g_1(\boldsymbol{\phi}) = \boldsymbol{\theta}$ and $g_2(\boldsymbol{\phi}) = \boldsymbol{c}$. Here $\boldsymbol{\phi}^{(h+(s-1)/2)}$, $s = 1, 2$ is a sequence of solutions to the maximization problem described as follows. The $(h+1)$-th iteration of the algorithm contains two sub-iterations as given below:

1) For $s = 1$, we solve

$$\arg\max_{\boldsymbol{\phi}} Q\left(\boldsymbol{\theta}, \boldsymbol{c} | \boldsymbol{\theta}^{(h)}, \boldsymbol{c}^{(h)}\right) \text{ subject to } \boldsymbol{\phi} \in \Theta_1\left(\boldsymbol{\phi}^{(h)}\right)$$

and denote this solution by $\boldsymbol{\phi}^{(h+1/2)}$.

2) For $s = 2$, we solve

$$\arg\max_{\boldsymbol{\phi}} Q\left(\boldsymbol{\theta}, \boldsymbol{c} | \boldsymbol{\theta}^{(h)}, \boldsymbol{c}^{(h)}\right) \text{ subject to } \boldsymbol{\phi} \in \Theta_2(\boldsymbol{\phi}^{(h+1/2)})$$

and denote this solution by $\boldsymbol{\phi}^{(h+1)}$.

This completes the $(h+1)$-th iteration of the algorithm. In order to facilitate the ECM algorithm, we now proceed to obtain the explicit expressions for the CM-steps. Note that

the partial derivatives of $E\left[H_0(M_i^X)\right]$ and $E\left[\log h_0(M_i^X)\right]$ with respect to c_q are given by

$$\frac{\partial E\left[H_0(M_i^X)\right]}{\partial c_q} = \frac{1}{S_X^{(h)}(w_i)} \sum_{j=1}^{M} A_{\theta_1}^{(h)}(w_i, j, q)$$

and

$$\frac{\partial E\left[\log h_0(M_i^X)\right]}{\partial c_q} = -\frac{1}{c_q S_X^{(h)}(w_i)} \sum_{j=1}^{M} B_{\theta_1}^{(h)}(w_i, j, q),\ q = 1, \ldots, M-1;$$

where

$$A_{\theta_1}^{(h)}(w_i, j, q) = \Bigg[\left\{\min\left(w_i, \tau_q\right) - \tau_{q-1}\right\} S_X^{(h)}(w_i) I_{[\tau_{q-1}, \tau_M)}(w_i) - \frac{1}{\theta_1^{(h)} c_q^{(h)}} \left\{ S_X^{(h)}(\tau_q) - S_X^{(h)}\left(\max(w_i, \tau_{q-1})\right) \right\} I\{q \geq j\} \Bigg] I_{[\tau_{j-1}, \tau_j)}(w_i)$$

and

$$B_{\theta_1}^{(h)}(w_i, j, q) = \left[\left\{ S_X^{(h)}(\tau_q) - S_X^{(h)}\left(\max(w_i, \tau_{q-1})\right) \right\} I\{q \geq j\} \right] I_{[\tau_{j-1}, \tau_j)}(w_i).$$

Similarly, the partial derivatives of $E\left[H_0(M_i^Y)\right]$ and $E\left[\log h_0(M_i^Y)\right]$ with respect to c_q can be obtained by replacing θ_1 and $S_X^{(h)}$ with θ_2 and $S_Y^{(h)}$, respectively and We have

$$\frac{\partial E\left[H_0(M_i^Y)\right]}{\partial c_q} = \frac{1}{S_Y^{(h)}(w_i)} \sum_{j=1}^{M} A_{\theta_2}^{(h)}(w_i, j, q)$$

and

$$\frac{\partial E\left[\log h_0(M_i^Y)\right]}{\partial c_q} = -\frac{1}{c_q S_Y^{(h)}(w_i)} \sum_{j=1}^{M} B_{\theta_2}^{(h)}(w_i, j, q),\ q = 1, \ldots, M-1.$$

Also note that

$$\frac{\partial H_0(w_i)}{\partial c_q} = p(i, q) = \left\{\min\left(w_i, \tau_q\right) - \tau_{q-1}\right\} I_{[\tau_{q-1}, \tau_M)}(w_i) \text{ and } \frac{\partial \log h_0(w_i)}{\partial c_q} = \frac{I_{[\tau_{q-1}, \tau_q)}(w_i)}{c_q}.$$

Now using (1), the iterative CM-steps at $(h+1)$-th iteration of the algorithm can be computed as follows:

1) For $s = 1$, the solution of the constrained maximization problem can be obtained as $\boldsymbol{\phi}^{(h+1/2)} = \left(\boldsymbol{\theta}^{(h+1/2)}, \boldsymbol{c}^{(h+1/2)}\right)$, where $\boldsymbol{\theta}^{(h+1/2)} = \left(\theta_1^{(h)}, \theta_2^{(h)}\right)$ and

$$c_q^{(h+1/2)} = \frac{N^{(h)}\left(\boldsymbol{w}, \boldsymbol{z}\right)}{D^{(h)}\left(\boldsymbol{w}, \boldsymbol{z}\right)},\ q = 1, \ldots, M-1;$$

where

$$N^{(h)}(\boldsymbol{w},\boldsymbol{z}) = \sum_{i=1}^{k} I_{[\tau_{q-1},\tau_q)}(w_i) - \sum_{i=1}^{k}\sum_{j=1}^{M} \frac{m_R^i}{S_X^{(h)}(w_i)} B_{\theta_1}^{(h)}(w_i,j,q) - \sum_{i=1}^{k}\sum_{j=1}^{M} \frac{n_R^i}{S_Y^{(h)}(w_i)} B_{\theta_2}^{(h)}(w_i,j,q)$$

and

$$\begin{aligned} D^{(h)}(\boldsymbol{w},\boldsymbol{z}) = \theta_1^{(h)} &\left\{ \sum_{i=1}^{k} z_i p(i,q) + \sum_{i=1}^{k}\sum_{j=1}^{M} \frac{m_R^i}{S_X^{(h)}(w_i)} A_{\theta_1}^{(h)}(w_i,j,q) \right\} \\ &+ \theta_2^{(h)} \left\{ \sum_{i=1}^{k} (1-z_i) p(i,q) + \sum_{i=1}^{k}\sum_{j=1}^{M} \frac{n_R^i}{S_Y^{(h)}(w_i)} A_{\theta_2}^{(h)}(w_i,j,q) \right\}. \end{aligned}$$

2) For $s = 2$, the solution to the constrained maximization problem can be obtained as $\boldsymbol{\phi}^{(h+1)} = \left(\boldsymbol{\theta}^{(h+1)}, \boldsymbol{c}^{(h+1)}\right)$, where $\boldsymbol{c}^{(h+1)} = \boldsymbol{c}^{(h+1/2)}$. The solution $\boldsymbol{\theta}^{(h+1)}$ can be obtained as

$$\theta_1^{(h+1)} = m \Big/ \left(\sum_{i=1}^{k} z_i \, H_0^{(h+1)}(w_i) \sum_{i=1}^{k} m_R^i \, E\left[H_0^{(h+1)}(M_i^X) \right] \right)$$

and

$$\theta_2^{(h+1)} = n \Big/ \left(\sum_{i=1}^{k} (1-z_i) \, H_0^{(h+1)}(w_i) \sum_{i=1}^{k} n_R^i \, E\left[H_0^{(h+1)}(M_i^Y) \right] \right).$$

This completes the $(h+1)$-th iteration of the algorithm. We propose the obvious initialization of the parameters as $c_1 = \cdots = c_{M-1} = \theta_1 = \theta_2 = 1$.

Remark 3.1. *It is interesting to observe that the existence and uniqueness of the the solutions in the conditional maximization steps of the proposed algorithm are contingent upon the fulfillment of the condition* $\sum_{i=1}^{k} I_{[\tau_{q-1},\tau_q)}(w_i) \geq 1$ *for all* $q = 1,\ldots,M$*, i.e., each of the sub-interval* $[\tau_{q-1},\tau_q)$*,* $q = 1,\ldots,M$ *contains at least one of the progressive censored samples* $W_1 \leq \cdots \leq W_k$*.*

Remark 3.2. *We now discuss the convergence property of the proposed algorithm using the results of Meng and Rubin (1993). The convergence of an ECM algorithm depends on the existence of unique solution at each CM stage along with an additional condition on the* S *constraint sets, known as the 'space filling' condition. The second condition ensures that at any point inside the parameter space* Θ*, the algorithm is able to search for maximum in all possible directions, and consequently, the solutions in different stages of the ECM algorithm are not trapped in any of the constraint sets. This condition can be easily verified when the vector valued functions* g_s*,* $s = 1,\ldots,S$ *are differentiable and the matrix* $\nabla g_s(\phi)$ *whose* i*-th column is the gradient of the* i*-th element of* $g_s(\boldsymbol{\phi})$*, is of full column rank for* $\phi \in \Theta_0$*, the interior of* Θ *and for all* $s = 1,\ldots,S$*. Under these assumptions, the set of constraints* $G = \{g_s, s = 1,\ldots,S\}$ *is 'space filling' at* ϕ *if* $J(\phi) \equiv \cap_{s=1}^{S} J_s(\phi) = \{\mathbf{0}\}$*, where* $J_s(\phi)$ *is the column space of* $\nabla g_s(\phi)$*, i.e.,* $J_s(\phi) = \{\nabla g_s(\phi)\lambda : \lambda \in \mathbb{R}^{d_s}\}$ *and* d_s *is the dimension of the vector valued function* $g_s(\phi)$*. Note that for the proposed algorithm,* $\nabla g_s(\phi)$*,* $s = 1, 2$ *are*

given by

$$\nabla g_1(\phi) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{pmatrix} \text{ and } \nabla g_2(\phi) = \begin{pmatrix} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

It is clear that both $\nabla g_1(\phi)$ *and* $\nabla g_2(\phi)$ *are of full column ranks. Moreover,* $J_1(\phi)$ *and* $J_2(\phi)$ *are orthogonal complement to each other and hence* $\cap_{s=1}^{S} J_s(\phi) = \{\mathbf{0}\}$*. Further the existence of unique solution at each CM step is assured by the condition* $\sum_{i=1}^{k} I_{[\tau_{q-1},\tau_q)}(w_i) \geq 1$ *for all* $q = 1, \ldots, M$*. Hence from Theorem 3 of Meng and Rubin (1993), it follows that the proposed ECM algorithm converges to a stationary point of the observed likelihood function* $L(\boldsymbol{\theta}, \boldsymbol{c}|\boldsymbol{w}, \boldsymbol{z})$*.*

4. Simulation Study

In this section, we carry out a simulation study in order to assess the performance of the proposed estimation procedure. Here we consider the cases of $M = 2$ and $M = 3$. For certain choices of the parameters $\boldsymbol{c}$, $\boldsymbol{\theta}$ and the cut points, we generate balanced joint progressive censored samples corresponding to different values of m, n and different censoring schemes. In each case, we compute MLEs of the parameters along with associated mean squared errors (MSEs) based on 1000 replications. These results, corresponding to $M = 2$, are presented in Table 1. The results corresponding to $M = 3$ can be found in the Supplementary Material. For each generated sample, we also evaluate 90% bootstrap confidence interval (see Efron (1982)). The parametric percentile bootstrap confidence intervals for $M = 2$ are obtained using the following steps:

Step 1: For a generated sample, compute the maximum likelihood estimates $\hat{c}_1$, $\hat{\theta}_1$ and $\hat{\theta}_2$.
Step 2: Generate bootstrap samples $(W_1^*, Z_1^*), \ldots, (W_k^*, Z_k^*)$ corresponding to the parameters $\hat{c}_1$, $\hat{\theta}_1$ and $\hat{\theta}_2$ and the censoring scheme k, $R_1, \ldots, R_{k-1}$.
Step 3: Compute the MLEs $\hat{c}_1^*$, $\hat{\theta}_1^*$ and $\hat{\theta}_2^*$ based on the bootstrap sample.
Step 4: Repeat Steps 2 and 3 B times to obtain $\{\hat{c}_{11}^*, \ldots, \hat{c}_{1B}^*\}$, $\{\hat{\theta}_{11}^*, \ldots, \hat{\theta}_{1B}^*\}$ and $\{\hat{\theta}_{21}^*, \ldots, \hat{\theta}_{2B}^*\}$.
Step 5: Construct a $100(1-\alpha)\%$ confidence interval for c_1, θ_1 and θ_2 as $\left(\hat{c}_{1k}^*, \hat{c}_{1k'}^*\right)$, $\left(\hat{\theta}_{1k}^*, \hat{\theta}_{1k'}^*\right)$ and $\left(\hat{\theta}_{2k}^*, \hat{\theta}_{2k'}^*\right)$ respectively, where $k = \left[\frac{\alpha}{2}B\right]$, $k = \left[1 - \frac{\alpha}{2}B\right]$ and $[x]$ denotes the largest integer less than or equal to x.

Similar procedure is carried out for obtaining confidence intervals corresponding to $M = 3$. We report average length (AL), median length (ML) and associated coverage probability (CP) corresponding to $M = 2$ and $M = 3$ in the Supplementary Material. From now on, we adopt a notation to denote a particular censoring scheme. For instance, $\left(0_{(18)}, 5\right)$ implies $R_1 = \ldots R_{18} = 0$, $R_{19} = 5$.

It is evident from the simulation study that the estimates are quite satisfactory in case of all the censoring schemes even when k is small. It is also clear that the MSEs and the

Table 1: AE and MSE of the estimators ($M = 2$)

m, n	Censoring Scheme	Estimates					
		c_1		θ_1		θ_2	
		AE	MSE	AE	MSE	AE	MSE
$\tau_1 = 0.6, c_1 = 2, \theta_1 = 0.1, \theta_2 = 0.2$							
40, 40	$k = 20, R = \left(0_{(18)}, 5\right)$	2.0742	1.1327	0.1180	0.0108	0.2574	0.0405
	$k = 20, R = \left(5, 0_{(18)}\right)$	2.0845	1.1968	0.1073	0.0038	0.2306	0.0119
	$k = 20, R = \left(0_{(9)}, 5, 0_{(9)}\right)$	2.0870	1.3677	0.1142	0.0074	0.2440	0.0266
60, 60	$k = 30, R = \left(0_{(28)}, 5\right)$	2.0442	0.8556	0.1041	0.0065	0.2227	0.0190
	$k = 30, R = \left(5, 0_{(28)}\right)$	1.9997	0.5222	0.1073	0.0019	0.2190	0.0056
	$k = 30, R = \left(0_{(14)}, 5, 0_{(14)}\right)$	2.0075	0.5927	0.1092	0.0027	0.2286	0.0102
$\tau_1 = 1.5, c_1 = 0.04, \theta_1 = 2, \theta_2 = 4$							
30, 40	$k = 20, R = \left(0_{(18)}, 5\right)$	0.0418	0.0006	2.2188	1.8437	4.6533	5.4900
	$k = 20, R = \left(5, 0_{(18)}\right)$	0.0415	0.0004	2.2049	1.6354	4.3986	2.6930
	$k = 20, R = \left(0_{(9)}, 5, 0_{(9)}\right)$	0.0411	0.0005	2.2862	2.8920	4.5796	4.4492
60, 50	$k = 30, R = \left(0_{(28)}, 5\right)$	0.0398	0.0002	2.1046	0.6691	4.4626	2.4204
	$k = 30, R = \left(5, 0_{(28)}\right)$	0.0395	0.0002	2.0685	0.5884	4.4779	2.3360
	$k = 30, R = \left(0_{(14)}, 5, 0_{(14)}\right)$	0.0403	0.0003	2.1065	0.5962	4.4478	2.4457

lengths of the confidence intervals are decreasing with increasing values of m, n and k. The coverage probabilities are also quite close to the nominal values.

5. Determination of Cut Points

Throughout the article, we have assumed that both the number and the positions of the cut points in the baseline hazard function are known. However, in most practical situations, it is more realistic to assume that the positions of the cut points are unknown. One common approach to handle unknown cut points is to choose equidistant cut points, though this method may not always be practical or accurate. In this context, Goodman et al. (2011) proposed estimators for both the number and positions of cut points using the maximum likelihood principle. However, their methodology is complex and not easily implemented in practical settings.

In this section, we introduce a novel methodology that is both easy to implement in practice and flexible enough to generalize the case of equidistant cut points. Our approach addresses the practical challenges of determining cut point positions, offering a straightforward and effective solution for real-world applications. We are mainly concerned about the shape of the hazard rate function in $[0, \tau]$, where $\tau = w_k$. Here we model the positions of the cut points as the arrival times of events in $[0, \tau]$, from a non-homogeneous Poisson process $\{N(t) : t \geq 0\}$, with intensity function $u(t) = \alpha t^{\alpha-1}$, $\alpha > 0$. Since the arrival times of the events from $N(t)$ are not observed, we treat them as missing values. Hence for a given $M-1$ number of cut points, the likelihood of the observations becomes a function of the parameters $\boldsymbol{c}$, $\boldsymbol{\theta}$ and α, where the arrival times of the events from $N(t)$, being missing, are replaced with their expected values.

Suppose $\{N(t) : t \geq 0\}$ is a non-homogeneous Poisson process with intensity function $u(t) = \alpha t^{\alpha-1}$, $\alpha > 0$. Let $T_1, T_2, \ldots$ be the ordered arrival times of events from $N(t)$. We now proceed to provide the expected values of the arrival times of the events from $N(t)$. To this end, we first note that that the joint pdf of $T_1, T_2, \ldots, T_n$, given that $N(\tau) = n$, can be written as

$$f_{T_1,\ldots T_n|N(\tau)=n}(t_1, \ldots, t_n) = \begin{cases} \frac{n!\alpha^n}{\tau^{n\alpha}} t_1^{\alpha-1} \ldots t_n^{\alpha-1}, & 0 < t_1 < \cdots < t_n < \tau \\ 0 & \text{otherwise} \end{cases}.$$

Hence, for $j = 1, \ldots, n$, we have

$$\begin{aligned} E\left(T_j | N(\tau) = n\right) &= \frac{n!\alpha^n}{\tau^{n\alpha}} \int_0^{\tau} \int_0^{t_n} \cdots \int_0^{t_3} \int_0^{t_2} t_1^{\alpha-1} \ldots t_{j-1}^{\alpha-1} t_j^{\alpha} t_{j+1}^{\alpha-1} \ldots t_n^{\alpha-1} \, dt_1 dt_2 \ldots dt_{n-1} dt_n \\ &= \frac{n!\alpha^n}{\tau^{n\alpha}} \times \frac{\tau^{n\alpha+1}}{\alpha(2\alpha)\ldots((j-1)\alpha)\left(j\alpha+1\right)\left((j+1)\alpha+1\right)\ldots(n\alpha+1)} \\ &= \frac{n!}{(j-1)!} \frac{\alpha^{n-j+1}\tau}{\left(j\alpha+1\right)\left((j+1)\alpha+1\right)\ldots(n\alpha+1)} \\ &= \frac{\tau}{\left(1+\frac{1}{j\alpha}\right)\ldots\left(1+\frac{1}{n\alpha}\right)}. \end{aligned}$$

For homogeneous Poisson process, i.e., for $\alpha = 1$, $E\left(T_j | N(\tau) = n\right) = j\tau/(n+1)$. So in this case, the cut points are equidistant to each other.

Note that for each value of α, the estimates of the parameters can be obtained using the method discussed earlier. Now one can numerically vary the values of α and obtain different estimates of the parameters $\boldsymbol{c}$, $\boldsymbol{\theta}$ along with associated values of the likelihood corresponding the observations. Hence the value of α, and consequently, the positions of the cut points, should be chosen which maximizes the likelihood as a function of α. This numerical procedure can be easily carried out in connection with situations involving real data. One can also choose M using some information theoretic criteria, such as Akike Information Criterion (AIC) and Bayes Information Criterion (BIC).

6. Applications to Real Data

We now apply the proposed estimation methodology to real data sets for the purpose of illustration. Here we consider the cases $M = 1, 2$ and 3 since these choices of M works quite well in practice, see for instance, Ibrahim et al. (2001) and Balakrishnan et al. (2016).

Data Set 1: We choose the data set from a skin-painting study carried out by the Tobacco Research Council, UK (see Lee (1979)). The data consists of time (in weeks) to visible skin tumor and time to death without skin tumor of three groups of Carworth Swiss female mice painted with condensate from standard cigarettes at dose levels 180, 136 and 103 mg per week. For illustrative purpose, here we choose the times to visible skin tumor of the two groups of mice, painted with condensate from standard cigarettes at dose levels 180 mg and 136 mg per week, respectively. The times to visible skin tumor of the two groups are presented in Tables 2 and 3 respectively. We now generate joint progressive censored samples $(\boldsymbol{W}, \boldsymbol{Z})$ from the data set corresponding to different censoring schemes and the generated samples are presented in Table 4. Here we also assume that the times to visible skin tumor of the two groups belong to a PHR class of life distributions with the baseline distribution having a piecewise constant hazard function with one $(M = 2)$ or two $(M = 3)$ cut points. We report the maximum likelihood estimates of the parameters along with associated estimates of the cut points corresponding to that value of α which maximizes the observed likelihood function. We also report the MLEs corresponding to exponential model $(M = 1)$. Further we obtain 90% bootstrap confidence intervals based on 500 bootstrap samples corresponding to each of the joint progressive censored samples. These results corresponding to $M = 1$, $M = 2$ and $M = 3$ are reported in Tables 5, 6 and 7, respectively. Here also the range of α is chosen in such a way that each of the sub-intervals contain at least one joint progressive censored samples. It is observed the MLEs of the parameters, as well as the associated confidence intervals, exhibit small variation across different censoring schemes. Moreover, it is interesting to observe that the estimates of the parameters exhibit increasing failure rate of the underlying distribution functions in all the cases of censoring scheme corresponding to $M = 2$ and $M = 3$, which is quite expected. Further it is observed that in all the cases, the failure rate function corresponding to Group 1 uniformly dominates that corresponding to Group 2. This is in accordance with the fact that the dose of condensate for Group 1 is higher than the dose for Group 2.

Data Set 2: We now choose the data sets from Bader and Priest (1982), which was recently analyzed by Sultana et al. (2024) and Kundu and Gupta (2006). These data sets represent the strength measured in GPa (giga-Pascals) for single carbon fibers, and impregnated 1000 carbon fiber tows. Single fibers were tested under tension at gauge lengths of 20mm (Group-I) and 10 mm (Group-II) and are presented in Tables 8 and 9. Again we generate joint progressive censored samples $(\boldsymbol{W}, \boldsymbol{Z})$ from the data set corresponding to different censoring schemes and the generated samples are presented in Table 10. As in the case of Data Set 1, here we also assume that the times to visible skin tumor of the two groups belong to a PHR class of life distributions with the baseline distribution having a piecewise constant hazard function with one $(M = 2)$ or two $(M = 3)$ cut points. We report the maximum likelihood estimates of the parameters along with associated estimates of the cut points

corresponding to that value of α which maximizes the observed likelihood function. We also report the MLEs corresponding to exponential model ($M = 1$). Further we obtain 90% bootstrap confidence intervals based on 500 bootstrap samples corresponding to each of the joint progressive censored samples. These results corresponding to $M = 1$, $M = 2$ and $M = 3$ are reported in Tables 11, 12 and 13, respectively. Here also the range of α is chosen in such a way that each of the sub-intervals contain at least one joint progressive censored samples. Much like in the case of Data Set 1, the MLEs of the parameters, as well as the associated confidence intervals, exhibit small variation across different censoring schemes. Again in this case, It is observed that the estimates of the parameters exhibit increasing failure rate of the underlying distribution functions in all the cases of censoring scheme corresponding to $M = 2$ and $M = 3$, which is quite expected. Moreover, it is quite interesting to observe that in all the cases, the failure rate function corresponding to Group I uniformly dominates that corresponding to Group II.

We calculate AIC and BIC corresponding to different choices of M for the data sets considered. These results are reported in Tables 14 and 15 corresponding to Data Set 1 and Data Set 2, respectively. It is observed that, regardless of the censoring scheme, in the case of both the data sets, the choice of $M = 3$ is recommended based on the information criterion.

We also compare the proposed methodology with the Newton-Raphson method for $M = 2$ and $M = 3$. Both methods yield the same solution, so we have compared the number of iterations required for convergence, using identical initial values and accuracy levels. The results are presented in Tables 16 and 17, corresponding to Data Sets 1 and 2, respectively. These tables show that, in most cases, the proposed ECM algorithm requires more iterations than the Newton-Raphson method, a phenomenon similarly observed by Ng et al. (2002). However, it is important to highlight that the proposed algorithm is computationally much simpler than the Newton-Raphson method, especially for larger M. This simplicity arises because the proposed algorithm provides explicit constrained maximization steps for any general M, whereas the Newton-Raphson method involves computing the inverse of the Hessian matrix, which is exceedingly difficult to derive explicitly for general M.

We now compare the proposed model with a parametric model, assuming that the lifetimes of the two populations follow Weibull distributions with a common shape parameter but different scale parameters. Weibull-distributed lifetimes are widely used in reliability studies, and such a parametric model has been analyzed by Mondal and Kundu (2020) in the context of BJPC. The pdf of the Weibull distribution with shape parameter α and scale parameter λ is given by

$$f(t;\alpha,\lambda) = \alpha\lambda t^{\alpha-1}e^{-\lambda t^{\alpha}},\ t > 0 \text{ and } \alpha,\lambda > 0.$$

For the Weibull model, we assume that $X_1,\ldots,X_m \overset{\text{iid}}{\sim} \text{Weibull}(\alpha,\lambda_1)$ and $Y_1,\ldots,Y_n \overset{\text{iid}}{\sim} \text{Weibull}(\alpha,\lambda_2)$. We apply this model to the datasets considered here and report the maximum observed log-likelihood values (based on BJPC samples under three censoring schemes). These maximum values are then compared to those obtained using the proposed semi-parametric model (denoted by PHC), across different choices of M. The results are presented in Tables 18 and 19, corresponding to Datasets 1 and 2, respectively.

These results demonstrate that, for both datasets, the Weibull model consistently achieves

higher observed likelihood values than the exponential model ($M = 1$). However, as expected, for $M \geq 2$, the proposed semi-parametric model outperforms the Weibull model across all cases. Notably, even though the Weibull and the proposed model for $M = 2$ have the same number of parameters, the proposed model exhibits significant gain in the observed likelihood. The highest gains in log-likelihood are observed for $M = 3$ in all the cases, this is also in accordance with the findings based on the the information criterion discussed earlier. Moreover, Table 19 shows substantial gain in log-likelihood when $M = 3$. These results highlight the flexibility and advantage of the semi-parametric model, which benefits from the very flexible baseline hazard function afforded by the piecewise constant hazard function.

Table 2: Time (in weeks) to skin cancer of the mice in Group 1

24	24	29	31	32	34	34	38	41	41	41	41	43	48	48
49	51	57	58	59	61	61	62	62	62	63	64	65	66	66
68	72	72	75	76	79	79	81	84	85	88	88	100		

Table 3: Time (in weeks) to skin cancer of the mice in Group 2

32	38	46	46	50	61	61	63	64	66	67	69	71	71	72
72	75	75	75	79	79	79	80	81	81	81	82	84	84	85
88	88	89	90	93	98									

Table 4: Generated balanced joint progressive censored samples from Data Set 1

Censoring Scheme		Joint Progressive Censored Samples														
$k = 30, R = (0_{(28)}, 5)$	$\boldsymbol{W}$:	24	29	31	32	32	34	38	41	43	46	48	49	51	57	58
		59	61	61	62	63	64	65	66	68	71	72	72	75	76	79
	$\boldsymbol{Z}$:	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1
		1	1	0	1	1	1	1	1	1	0	1	0	0	1	0
$k = 30, R = (5, 0_{(28)})$	$\boldsymbol{W}$:	24	29	32	32	34	38	41	43	46	48	50	51	57	58	59
		61	62	63	64	65	66	66	68	72	72	76	79	84	85	85
	$\boldsymbol{Z}$:	1	1	1	0	1	1	1	1	0	1	0	1	1	1	1
		1	1	0	1	1	1	0	1	1	0	1	0	0	1	0
$k = 30, R = (0_{(14)}, 5, 0_{(14)})$	$\boldsymbol{W}$:	24	29	31	32	32	34	38	38	41	43	46	48	49	50	51
		57	58	59	61	62	63	63	64	65	66	72	76	79	79	84
	$\boldsymbol{Z}$:	1	1	1	1	0	1	1	0	1	1	0	1	1	0	1
		1	1	1	1	1	1	0	1	1	1	1	1	1	0	0

Table 5: MLE and 90% CI of the parameters for $M = 1$ (Data Set 1)

Parameters			
θ_1		θ_2	
MLE	90% CI	MLE	90% CI
Censoring Scheme: $k = 30, R = \left(0_{(28)}, 5\right)$			
0.0136	(0.0078, 0.02016)	0.0094	(0.0057, 0.0175)
Censoring Scheme: $k = 30, R = \left(5, 0_{(28)}\right)$			
0.0099	(0.0056, 0.0149)	0.0094	(0.0056, 0.0175)
Censoring Scheme: $k = 30, R = \left(0_{(14)}, 5, 0_{(14)}\right)$			
0.0138	(0.0079, 0.0224)	0.0100	(0.0060, 0.0189)

Table 6: MLE and 90% CI of the parameters for $M = 2$ (Data Set 1)

Parameters					
c_1		θ_1		θ_2	
MLE	90% CI	MLE	90% CI	MLE	90% CI
Censoring Scheme: $k = 30, R = \left(0_{(28)}, 5\right)$; $\hat{\alpha} = 3.38$, $\hat{\tau}_1 = 60.96$					
0.1472	(0.0720, 0.2371)	0.0348	(0.0227, 0.0579)	0.0149	(0.0087, 0.0359)
Censoring Scheme: $k = 30, R = \left(5, 0_{(28)}\right)$; $\hat{\alpha} = 2.86$, $\hat{\tau}_1 = 62.98$					
0.1547	(0.0809, 0.2703)	0.0326	(0.0205, 0.0535)	0.0214	(0.0118, 0.0427)
Censoring Scheme: $k = 30, R = \left(0_{(14)}, 5, 0_{(14)}\right)$; $\hat{\alpha} = 2.64$, $\hat{\tau}_1 = 60.92$					
0.1679	(0.0831, 0.2689)	0.0347	(0.0231, 0.0582)	0.0164	(0.0097, 0.0359)

Table 7: MLE and 90% CI of the parameters for $M = 3$ (Data Set 1)

Parameters							
c_1		c_2		θ_1		θ_2	
MLE	90% CI	MLE	90% CI	MLE	90% CI	MLE	90% CI
Censoring Scheme: $k = 30, R = \left(0_{(28)}, 5\right)$; $\hat{\alpha} = 4.42$, $\hat{\tau}_1 = 57.90$, $\hat{\tau}_2 = 71.00$							
0.1009	(0.0346, 0.2095)	0.5650	(0.1918, 1.1379)	0.0459	(0.0256, 0.1224)	0.0199	(0.0103, 0.0699)
Censoring Scheme: $k = 30, R = \left(5, 0_{(28)}\right)$; $\hat{\alpha} = 4.22$, $\hat{\tau}_1 = 61.40$, $\hat{\tau}_2 = 76.00$							
0.1211	(0.0381, 0.2590)	0.7042	(0.2401, 1.5410)	0.0397	(0.0201, 0.1168)	0.0263	(0.0122, 0.0950)
Censoring Scheme: $k = 30, R = \left(0_{(14)}, 5, 0_{(14)}\right)$; $\hat{\alpha} = 4.74$, $\hat{\tau}_1 = 62.70$, $\hat{\tau}_2 = 76.00$							
0.1414	(0.0448, 0.2949)	0.6184	(0.1642, 1.4770)	0.0453	(0.0219, 0.1239)	0.0211	(0.0092, 0.0815)

Table 8: Group-I (gauge lengths of 20 mm)

1.312	1.314	1.479	1.552	1.700	1.803	1.861	1.865	1.944	1.958
1.966	1.997	2.006	2.021	2.027	2.055	2.063	2.098	2.140	2.179
2.224	2.240	2.253	2.270	2.272	2.274	2.301	2.301	2.359	2.382
2.382	2.426	2.434	2.435	2.478	2.490	2.511	2.514	2.535	2.554
2.566	2.570	2.586	2.629	2.633	2.642	2.648	2.684	2.697	2.726
2.770	2.773	2.800	2.809	2.818	2.821	2.848	2.880	2.954	3.012
3.067	3.084	3.090	3.096	3.128	3.233	3.433	3.585	3.585	

Table 9: Group-II (gauge lengths of 10 mm)

1.901	2.132	2.203	2.228	2.257	2.350	2.361	2.396	2.397	2.445
2.454	2.474	2.518	2.522	2.525	2.532	2.575	2.614	2.616	2.618
2.624	2.659	2.675	2.738	2.740	2.856	2.917	2.928	2.937	2.937
2.977	2.996	3.030	3.125	3.139	3.145	3.220	3.223	3.235	3.243
3.264	3.272	3.294	3.332	3.346	3.377	3.408	3.435	3.493	3.501
3.537	3.554	3.562	3.628	3.852	3.871	3.886	3.971	4.024	4.027
4.225	4.395	5.020							

Table 10: Generated balanced joint progressive censored samples from Data Set 2

Censoring Scheme		Joint Progressive Censored Samples									
$k=40, R=\left(0_{(38)}, 5\right)$	$\boldsymbol{W}$	1.312	1.314	1.479	1.552	1.700	1.803	1.861	1.865	1.901	1.944
		1.958	1.966	1.997	2.006	2.021	2.027	2.055	2.063	2.098	2.132
		2.179	2.203	2.224	2.228	2.240	2.253	2.270	2.272	2.274	2.301
		2.350	2.361	2.382	2.426	2.434	2.435	2.454	2.478	2.490	2.514
	$\boldsymbol{Z}$	1	1	1	1	1	1	1	1	0	1
		1	1	1	1	1	1	1	1	1	0
		1	0	1	0	1	1	1	1	1	1
		0	0	1	1	1	1	0	1	1	0
$k=40, R=\left(5, 0_{(38)}\right)$	$\boldsymbol{W}$	1.312	1.314	1.479	1.700	1.803	1.861	1.865	1.944	1.958	1.966
		1.997	2.006	2.021	2.027	2.055	2.063	2.132	2.140	2.179	2.203
		2.224	2.253	2.272	2.274	2.301	2.350	2.359	2.361	2.382	2.397
		2.426	2.434	2.435	2.445	2.490	2.511	2.518	2.522	2.525	2.566
	$\boldsymbol{Z}$	1	1	1	1	1	1	1	1	1	1
		1	1	1	1	1	1	0	1	1	0
		1	1	1	1	1	0	1	0	1	0
		1	1	1	0	1	1	0	0	0	0
$k=40, R=\left(0_{(19)}, 5, 0_{(19)}\right)$	$\boldsymbol{W}$	1.312	1.314	1.479	1.552	1.700	1.803	1.861	1.865	1.901	1.944
		1.958	1.966	1.997	2.006	2.021	2.027	2.055	2.063	2.098	2.132
		2.140	2.179	2.203	2.224	2.240	2.253	2.270	2.272	2.274	2.301
		2.350	2.361	2.382	2.397	2.426	2.434	2.435	2.445	2.454	2.478
	$\boldsymbol{Z}$	1	1	1	1	1	1	1	1	0	1
		1	1	1	1	1	1	1	1	1	0
		1	1	0	1	1	1	1	1	1	1
		0	0	1	0	1	1	1	0	0	0

Table 11: MLE and 90% CI of the parameters for $M = 1$ (Data Set 2)

Parameters			
θ_1		θ_2	
MLE	90% CI	MLE	90% CI
Censoring Scheme: $k = 40, R = \left(0_{(38)}, 5\right)$			
0.2043	(0.1479, 0.2784)	0.0565	(0.0334, 0.1027)
Censoring Scheme: $k = 40, R = \left(5, 0_{(38)}\right)$			
0.1945	(0.1411, 0.2622)	0.0720	(0.0455, 0.1240)
Censoring Scheme: $k = 40, R = \left(0_{(19)}, 5, 0_{(19)}\right)$			
0.2017	(0.1520, 0.2731)	0.0648	(0.0389, 0.1133)

Table 12: MLE and 90% CI of the parameters for $M = 2$ (Data Set 2)

Parameters					
c_1		θ_1		θ_2	
MLE	90% CI	MLE	90% CI	MLE	90% CI
Censoring Scheme: $k = 40, R = \left(0_{(38)}, 5\right)$; $\hat{\alpha} = 25.0$, $\hat{\tau}_1 = 2.4173$					
0.0897	(0.0478, 0.1658)	1.9075	(1.1772, 3.2384)	0.5374	(0.2719, 1.1599)
Censoring Scheme: $k = 40, R = \left(5, 0_{(28)}\right)$; $\hat{\alpha} = 17.3$, $\hat{\tau}_1 = 2.4257$					
0.0770	(0.0429, 0.1263)	1.9300	(1.1957, 3.1114)	0.7360	(0.4309, 1.3761)
Censoring Scheme: $k = 40, R = \left(0_{(19)}, 5, 0_{(19)}\right)$; $\hat{\alpha} = 24.8$, $\hat{\tau}_1 = 2.3819$					
0.0694	(0.0359, 0.1255)	2.3510	(1.5139, 4.1340)	0.7745	(0.4074, 1.5609)

Table 13: MLE and 90% CI of the parameters for $M = 3$ (Data Set 2)

Parameters							
c_1		c_2		θ_1		θ_2	
MLE	90% CI	MLE	90% CI	MLE	90% CI	MLE	90% CI
Censoring Scheme: $k = 40, R = \left(0_{(38)}, 5\right)$; $\hat{\alpha} = 13.70$, $\hat{\tau}_1 = 2.26$, $\hat{\tau}_2 = 2.42$							
0.0664	(0.0305, 0.1101)	0.4233	(0.1609, 0.9079)	2.1082	(1.2272, 4.2553)	0.6000	(0.2940,1.5605)
Censoring Scheme: $k = 40, R = \left(5, 0_{(38)}\right)$; $\hat{\alpha} = 22.80$, $\hat{\tau}_1 = 2.41$, $\hat{\tau}_2 = 2.51$							
0.0543	(0.0276, 0.1017)	0.4268	(0.1393, 0.8149)	2.7464	(1.5062, 4.3446)	1.0481	(0.5471, 1.7664)
Censoring Scheme: $k = 40, R = \left(0_{(19)}, 5, 0_{(19)}\right)$; $\hat{\alpha} = 23.30$, $\hat{\tau}_1 = 2.32$, $\hat{\tau}_2 = 2.42$							
0.0442	(0.0209, 0.07855)	0.2835	(0.0651, 0.5741)	3.4965	(2.1657, 7.5354)	1.1586	(0.6450, 2.2557)

Table 14: AIC and BIC corresponding to different choices of M for Data Set 1

Censoring Scheme	$M=1$		$M=2$		$M=3$	
	AIC	BIC	AIC	BIC	AIC	BIC
$k=30, R=\left(0_{(28)},5\right)$	163.7279	166.5303	90.0830	94.2866	75.2178	80.8226
$k=30, R=\left(5,0_{(28)}\right)$	179.1262	181.9286	105.6228	109.8264	95.6585	101.2633
$k=30, R=\left(0_{(14)},5,0_{(14)}\right)$	168.9123	171.7146	97.2435	101.4471	83.3377	88.9425

Table 15: AIC and BIC corresponding to different choices of M for Data Set 2

Censoring Scheme	$M=1$		$M=2$		$M=3$	
	AIC	BIC	AIC	BIC	AIC	BIC
$k=40, R=\left(0_{(38)},5\right)$	-77.10769	-73.72993	-252.2475	-247.1809	-258.4348	-251.6793
$k=40, R=\left(5,0_{(38)}\right)$	-63.85576	-60.478	-244.1431	-239.0765	-270.3753	-263.6198
$k=40, R=\left(0_{(19)},5,0_{(19)}\right)$	-70.04863	-66.67087	-263.157	-258.0904	-293.0133	-286.2577

Table 16: Comparison of number of iterations corresponding to Data Set 1

Censoring Scheme					
$M=2$					
$k=30, R=\left(0_{(28)},5\right)$		$k=30, R=\left(5,0_{(28)}\right)$		$k=20, R=\left(0_{(14)},5,0_{(14)}\right)$	
ECM	Newton-Raphson	ECM	Newton-Raphson	ECM	Newton-Raphson
37	56	37	35	44	57
$M=3$					
$k=30, R=\left(0_{(28)},5\right)$		$k=30, R=\left(5,0_{(28)}\right)$		$k=20, R=\left(0_{(14)},5,0_{(14)}\right)$	
ECM	Newton-Raphson	ECM	Newton-Raphson	ECM	Newton-Raphson
87	76	101	62	124	60

Table 17: Comparison of number of iterations corresponding to Data Set 2

Censoring Scheme					
$M = 2$					
$k = 40, R = \left(0_{(38)}, 5\right)$		$k = 40, R = \left(5, 0_{(38)}\right)$		$k = 40, R = \left(0_{(19)}, 5, 0_{(19)}\right)$	
ECM	Newton-Raphson	ECM	Newton-Raphson	ECM	Newton-Raphson
83	25	68	22	101	25
$M = 3$					
$k = 30, R = \left(0_{(28)}, 5\right)$		$k = 30, R = \left(5, 0_{(28)}\right)$		$k = 20, R = \left(0_{(14)}, 5, 0_{(14)}\right)$	
ECM	Newton-Raphson	ECM	Newton-Raphson	ECM	Newton-Raphson
73	30	168	36	165	35

Table 18: Maximum log-likelihood values for the Weibull and proposed models (Data Set 1)

Censoring Scheme	Model			
	Weibull	PHC ($M = 1$)	PHC ($M = 2$)	PHC ($M = 3$)
$k = 30, R = \left(0_{(28)}, 5\right)$	-161.7307	-187.0149	-136.7767	-128.3037
$k = 30, R = \left(5, 0_{(28)}\right)$	-159.4166	-181.773	-136.5024	-130.5514
$k = 30, R = \left(0_{(14)}, 5, 0_{(14)}\right)$	-159.9552	-184.867	-136.1077	-128.1793

Table 19: Maximum log-likelihood values for the Weibull and proposed models (Data Set 2)

Censoring Scheme	Model			
	Weibull	PHC ($M = 1$)	PHC ($M = 2$)	PHC ($M = 3$)
$k = 40, R = \left(0_{(38)}, 5\right)$	-68.4181	-113.8005	-24.3002	-20.5261
$k = 40, R = \left(5, 0_{(38)}\right)$	-66.2690	-115.4201	-23.3526	-9.0590
$k = 40, R = \left(0_{(19)}, 5, 0_{(19)}\right)$	-64.6255	-114.2569	-15.7839	-0.6306

7. Conclusions

In this paper we have considered the classical inference of the unknown parameters of proportional hazard models for two sample problems and the data are obtained based on BJPC scheme. It is assumed that the base line distribution does not have any specific parametric form, rather it has piecewise constant hazard function. It makes the proposed model a very flexible one. The maximum likelihood estimators cannot be obtained in explicit forms, hence we have proposed ECM algorithm to compute the MLEs of the unknown parameters. Extensive simulations have been performed to show the effectiveness of the proposed algorithm, the results are quite along the expected line. Further we have discussed about the choice of the cut points of the base line hazard function based on non-homogeneous Poisson process model. It can be implemented very easily in practice. Two data sets have been analyzed and the results are quite satisfactory.

Acknowledgment

The authors are thankful to the Associate Editor and the anonymous reviewer for their helpful and constructive suggestions and comments which have led to a substantial improvement in the revised version of the manuscript.

References

Bader, M. and Priest, A. (1982). Statistical aspects of fibre and bundle strength in hybrid composites. *Progress in science and Engineering of Composites*, pages 1129–1136.

Balakrishnan, N. and Cramer, E. (2014). *The art of progressive censoring*. Springer, New York.

Balakrishnan, N., Koutras, M., Milienos, F., and Pal, S. (2016). Piecewise linear approximations for cure rate models and associated inferential issues. *Methodology and Computing in Applied Probability*, 18:937–966.

Dempster, A. (1977). Maximum likelihood estimation from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39:1–38.

Efron, B. (1982). *The jackknife, the bootstrap and other resampling plans*. SIAM.

Ibrahim, J. G., Chen, M.-H., and Sinha, D. (2001). *Bayesian Survival Analysis*. Springer, New York.

Kundu, D. and Gupta, R. D. (2006). Estimation of $P[Y < X]$ for Weibull distributions. *IEEE Trans. Reliab.*, 55(2):270–280.

Lee, P. N. (1979). *PSC-7-Mouse changeover experiment A.1.5.6. Statistical analysis of visible skin tumor data (Tobacco Advisory Council Document TA 51)*. Tobacco Research Council, London.

Meng, X.-L. and Rubin, D. B. (1993). Maximum likelihood estimation via the ECM algorithm: A general framework. *Biometrika*, 80(2):267–278.

Mondal, S. and Kundu, D. (2019a). Exact inference on multiple exponential populations under a joint type-II progressive censoring scheme. *Statistics*, 53(6):1329–1356.

Mondal, S. and Kundu, D. (2019b). A new two sample type-II progressive censoring scheme. *Communications in Statistics-Theory and Methods*, 48(10):2602–2618.

Mondal, S. and Kundu, D. (2020). Inference on weibull parameters under a balanced two-sample type ii progressive censoring scheme. *Quality and Reliability Engineering International*, 36(1):1–17.

Neal, R. M. and Hinton, G. E. (1998). A view of the EM algorithm that justifies incremental, sparse, and other variants. In *Learning in Graphical Models*, volume 89, pages 355–368. Springer Netherlands, Dordrecht.

Ng, H., Chan, P., and Balakrishnan, N. (2002). Estimation of parameters from progressively censored data using EM algorithm. *Computational Statistics & Data Analysis*, 39(4):371–386.

Rasouli, A. and Balakrishnan, N. (2010). Exact likelihood inference for two exponential populations under joint progressive type-II censoring. *Communications in Statistics-Theory and Methods*, 39(12):2172–2191.

Srivastava, J. (1987). More efficient and less time-consuming censoring designs for life testing. *Journal of Statistical Planning and Inference*, 16:389–413.

Sultana, F., Çetinkaya, Ç., and Kundu, D. (2024). Estimation of the stress-strength parameter under two-sample balanced progressive censoring scheme. *Journal of Statistical Computation and Simulation*, 94(6):1269–1299.