Current Status Data with two Competing risks and Missing failure types: A Non-Parametric Approach

Technical Report No. ASU/2019/1
Dated 04 September, 2019

Tamalika Koley

and

Anup Dewanji

Applied Statistics Unit
Indian Statistical Institute
Kolkata 700 108

Current Status Data with two Competing risks and Missing failure types: A Non-Parametric Approach

Tamalika Koley

email: tamalikakoley@gmail.com

Anup Dewanji

email: dewanjia@isical.ac.in

Abstract

Parametric analysis of current status data with two competing risks and a general pattern of missing failure types is studied by Koley and Dewanji (2018a). In this work, we have tried to focus on non-parametric maximum likelihood estimation of sub-distribution functions based on current status data with two competing risks under the assumption of time independent missing probabilities. Depending on the support of monitoring time, different methods for maximum likelihood estimation are proposed. Simulation studies are carried out in order to investigate the finite sample properties of maximum likelihood estimators. Finally, the methods are illustrated through some real life data sets.

Keywords— Current status data, masking probabilities, monitoring time, maximum likelihood estimation

1 Introduction

Censoring is very common in survival analysis where the information about the survival time is incomplete. Interval censoring occurs when the exact value of the failure time is not observed but is only known to lie within an interval. Current status data is an extreme form of interval censoring in which each individual under study is observed only once and the status (failure/success of the event) is noted at the observation (monitoring) time. The individuals experiencing the event are exposed to the risk of more than one causes called competing risks (See Kalbfleisch and Prentice (2002)). If an individual is observed to experience the event, then the corresponding cause resulting the event is also observed. This type of data is called current status data with competing risks.

Several works on current status data with competing risks are already present in the literature focusing on non-parametric estimation. See for example, Hudgens et al. (2001), Jewell et al. (2003), Jewell and Kalbfleisch (2004), Maathuis (2006) and many others. In all the above cases, they have considered the situation when complete information on the cause of failure is available. However, in competing risks data, missing or uncertain failure causes is a common phenomenon due to inadequacy in the diagnostic mechanism. Dewanji and Sengupta (2003) considered non-parametric estimation and proposed an EM algorithm under MAR assumption, and a Nelson-Aalen type estimator when additional information is available, for estimation of different cause-specific hazards based on a general missing pattern. In this missing pattern, if a true cause of failure is not observed, then one observes a set of possible causes containing the true cause. This type of uncertainity regarding the cause of failure arises in carcinogenicity studies where the true cause of death of an individual suffering from cancer cannot be diagnosed. In the context of current status data with two competing risks and above-mentioned missing failure type, parametric analysis has been considered by Koley and Dewanji (2018a).

In this work, we consider non-parametric maximum likelihood estimation of current status data with two competing risks and the above-mentioned missing failure types. We consider non-parametric maximum likelihood estimation of sub-distribution functions under two different situations depending on the assumption made on the monitoring time. In the first situation, monitoring times are fixed which can be extended to random monitoring time taking finite, discrete values. The second situation is the more general one, where monitoring time for each individual is random.

The data is described along with the construction of corresponding likelihood function in Section 2. In Section 3, non-parametric maximum likelihood estimation with fixed monitoring times is discussed while Section 4 deals with the general random monitoring time. Section 5 presents several simulation studies to investigate the finite sample properties of the proposed estimators while Section 6 illustrates the estimation procedures through the analyses of two real life data sets. Finally, Section 7 ends with some concluding remarks.

2 The Data and the Likelihood

Suppose there are n individuals under study and the event of interest is exposed to the risk of $m(=2)$ competing risks. Let T be the random variable denoting the failure time and $J \in \{1,2\}$ be the true cause of failure which are not observed directly. Also, let X denote the monitoring time random variable which is assumed to be independent of the random vector (T,J). If the failure happens, that is, $T \leq X$, then one observes $G \in \{\{1\},\{2\},\{1,2\}\}$, the set of possible causes containing the true cause, and, if $T > X$, then $G = \phi$. Clearly, if G is a singleton set, then $G = J$, that is, true cause is observed, whereas, $G = \{1,2\}$ represents complete missing. The support of G is given by $\mathcal{G} = \{\phi,\{1\},\{2\},\{1,2\}\}$. The observation on each individual consists of (x,g), a realization of the random vector (X,G). Define, the conditional probability of observing the set g of possible failure causes, given the true cause $J = j$ and all other related information, as

$$p_{gj}(x,t) = P\left[G = g \mid t = T < X = x, J = j\right],$$

for $g \in \mathcal{G} \setminus \phi$ and $j \in g$. If $j \notin g$, then this conditional probability is 0, since we assume that there is no misspecification in the data in the sense that the observed set of possible causes does contain the true cause always. This gives $\sum_{g \ni j} p_{gj}(x,t) = 1$, for a fixed j. These conditional probabilities are often referred to as masking probabilities with the true cause being masked (See Basu (2009)). Throughout this work, as in Koley and Dewanji (2018a), the masking probabilities are taken to be time independent, that is, $p_{gj}(x,t) = p_{gj}$, for $t \leq x, j = 1,2$ and $g \ni j$. Write $p_j = p_{\{j\}j}$; then, clearly, $p_{\{1,2\}j} = 1 - p_j$, for $j = 1,2$. For $g \neq \phi$,

$$\begin{aligned}
P\left[T \leq x, G = g\right] &= \sum_{j \in g} P\left[T \leq X = x, G = g, J = j\right] h(x) \\
&= \sum_{j \in g} \left[\int_0^x p_{gj} f_j(t) \mathrm{d}t\right] h(x) \\
&= \left[\sum_{j \in g} p_{gj} F_j(x)\right] h(x), \qquad (2.1)
\end{aligned}$$

where $f_j(\cdot)$ and $F_j(\cdot)$ are, respectively, the sub-density and sub-distribution functions due to cause j, for $j = 1,2$, and $h(\cdot)$ is the density function of X. For $g = \phi$,

$$\begin{aligned}
P\left[T > X = x\right] &= \left[\int_x^\infty f(t) \mathrm{d}t\right] h(x) \\
&= S(x) h(x), \qquad (2.2)
\end{aligned}$$

where $S(\cdot)$ is the survival function and $f(\cdot) = f_1(\cdot) + f_2(\cdot)$ is the density function of the failure time T. Note that, for a given x, $F(x) = 1 - S(x)$ is the overall distribution function. The data $\{(x_i, g_i),\ i = 1, \cdots, n\}$ are n iid realizations from the common density function given by

$$f^*(x, g) = \begin{cases} p_j F_j(x) h(x), & \text{if } T \le X = x, g = \{j\},\ j = 1, 2, \\ \Big[(1 - p_1)F_1(x) + (1 - p_2)F_2(x)\Big] h(x), & \text{if } T \le X = x, g = \{1, 2\}, \\ S(x)h(x), & \text{if } T > X = x, g = \phi, \end{cases} \tag{2.3}$$

with respect to the dominating measure $H \times \mu$, where $H(\cdot)$ is the distribution function of X and μ is the counting measure. The likelihood of the data can be written as the product of these density functions given by $\prod_{i=1}^n f^*(x_i, g_i)$.

3 Fixed monitoring times

Let $0 < \tau_1 < \cdots < \tau_K$ be the distinct, finite, fixed monitoring time points with n_i number of individuals being observed at τ_i. So the total number of individuals under study is $n = \sum_{i=1}^K n_i$. Among these n_i individuals observed at τ_i, suppose d_{gi} are observed to have failed with g as the observed set of possible causes, for $g \in \mathcal{G} \setminus \phi$ and $d_i = \sum_{g \in \mathcal{G} \setminus \phi} d_{gi}$ determines the total number of observed failures at τ_i. The remaining $(n_i - d_i)$ are censored at τ_i. Note that, at each monitoring time τ_i, there are n_i independent realizations from the identical distribution given by $f_i^*(g) = f^*(\tau_i, g \mid X = \tau_i)$. The likelihood function is, therefore, given by

$$\begin{aligned} &= \prod_{i=1}^K \{p_1 F_1(\tau_i)\}^{d_{\{1\}i}} \{p_2 F_2(\tau_i)\}^{d_{\{2\}i}} \{(1 - p_1)F_1(\tau_i) + (1 - p_2)F_2(\tau_i)\}^{d_{\{1,2\}i}} S(\tau_i)^{n_i - d_i}. \\ &= \prod_{i=1}^K \prod_{l=1}^{n_i} f_i^*(g_{il}), \end{aligned} \tag{3.1}$$

where g_{il} is the observed value of G for the lth individual monitored at τ_i, for $l = 1, \cdots, n_i$, $i = 1, \cdots, K$.

3.1 Identifiability

Let $\underset{\sim}{\zeta}$ denote the vector of parameters to be estimated, that is, p_1, p_2 and $F_j(\tau_i)$'s, for $j = 1, 2$ and $i = 1, \cdots, K$. To check the identifiability of the likelihood, suppose $\underset{\sim}{\zeta}^{(1)}$ and

$\underset{\sim}{\zeta}^{(2)}$ be two values of the parameter vector and consider the equality

$$\log L(\underset{\sim}{\zeta}^{(1)}) = \log L(\underset{\sim}{\zeta}^{(2)}), \tag{3.2}$$

for all possible data frequencies, where $L(\underset{\sim}{\zeta})$ is the likelihood (3.1) depending on the parameter $\underset{\sim}{\zeta}$. If the probabilities p_1 and p_2 are known beforehand, then $\underset{\sim}{\zeta}$ consists of only the $F_j(\tau_i)$'s, for $j = 1, 2$ and $i = 1, \cdots, K$. Equating the coefficients of $d_{\{1\}i}$ and $d_{\{2\}i}$ from both sides of (3.2), we have $F_1^{(1)}(\tau_i) = F_1^{(2)}(\tau_i)$, $F_2^{(1)}(\tau_i) = F_2^{(2)}(\tau_i)$, respectively. This holds for all $i = 1, \cdots, K$, and hence, $\underset{\sim}{\zeta}^{(1)} = \underset{\sim}{\zeta}^{(2)}$. Thus, the likelihood is identifiable.

Now suppose both p_1 and p_2 are unknown and are to be estimated. Then $\underset{\sim}{\zeta}$ consists of these masking probabilities along with $F_j(\tau_i)$'s, for $j = 1, 2$ and $i = 1, \cdots, K$. Note that there is identifiability when, if the equality (3.2) holds for two choices of the parameter vector $\underset{\sim}{\zeta}^{(1)}$ and $\underset{\sim}{\zeta}^{(2)}$, we have $\underset{\sim}{\zeta}^{(1)} = \underset{\sim}{\zeta}^{(2)}$. Equating the coefficients of $d_{\{1\}i}, d_{\{2\}i}$ and $(n_i - d_i)$ from both sides of (3.2), we have

$$p_1^{(1)} F_1^{(1)}(\tau_i) = p_1^{(2)} F_1^{(2)}(\tau_i),$$

$$p_2^{(1)} F_2^{(1)}(\tau_i) = p_2^{(2)} F_2^{(2)}(\tau_i),$$

and

$$F_1^{(1)}(\tau_i) + F_2^{(1)}(\tau_i) = F_1^{(2)}(\tau_i) + F_2^{(2)}(\tau_i),$$

for all $i = 1, \cdots, K$. The first two equations give

$$\frac{p_1^{(1)} F_1^{(1)}(\tau_i)}{p_2^{(1)} F_2^{(1)}(\tau_i)} = \frac{p_1^{(2)} F_1^{(2)}(\tau_i)}{p_2^{(2)} F_2^{(2)}(\tau_i)} = K^*(\tau_i), \quad \text{say}, \tag{3.3}$$

possibly a function of $p_1^{(1)}, p_1^{(2)}, p_2^{(1)}, p_2^{(2)}, F_1^{(1)}(\tau_i), F_1^{(2)}(\tau_i), F_2^{(1)}(\tau_i)$ and $F_2^{(2)}(\tau_i)$, for $i = 1, \cdots, K$. This implies

$$F_1^{(1)}(\tau_i) = K^*(\tau_i) \frac{p_2^{(1)}}{p_1^{(1)}} F_2^{(1)}(\tau_i),$$

$$F_1^{(2)}(\tau_i) = K^*(\tau_i) \frac{p_2^{(2)}}{p_1^{(2)}} F_2^{(2)}(\tau_i),$$

and

$$(1 + K^*(\tau_i) \frac{p_2^{(1)}}{p_1^{(1)}}) F_2^{(1)}(\tau_i) = (1 + K^*(\tau_i) \frac{p_2^{(2)}}{p_1^{(2)}}) F_2^{(2)}(\tau_i), \tag{3.4}$$

for all $i = 1, \cdots, K$. Using (3.4) and $p_2^{(1)} F_2^{(1)}(\tau_i) = p_2^{(2)} F_2^{(2)}(\tau_i)$, we have

$$\frac{1 + K^*(\tau_i)\dfrac{p_2^{(1)}}{p_1^{(1)}}}{1 + K^*(\tau_i)\dfrac{p_2^{(2)}}{p_1^{(2)}}} = \frac{p_2^{(1)}}{p_2^{(2)}}, \quad \text{assuming} \quad F_2^{(2)}(\tau_i) \neq 0, \quad \text{for some } i.$$

This implies,

$$K^*(\tau_i) = \frac{p_1^{(2)} p_1^{(1)} (p_2^{(2)} - p_2^{(1)})}{p_2^{(2)} p_2^{(1)} (p_1^{(2)} - p_1^{(1)})},$$

for some i. If the likelihood is identifiable, we have $p_j^{(1)} = p_j^{(2)}$, for $j = 1, 2$ and, hence, $K^*(\tau_i)$ will be of the form $0/0$, which contradicts (3.3), that shows $K^*(\tau_i) > 0$, for some i. So, $p_j^{(1)} \neq p_j^{(2)}$, for $j = 1, 2$, which gives $\underset{\sim}{\zeta}^{(1)} \neq \underset{\sim}{\zeta}^{(2)}$. Thus, the likelihood is not identifiable. This non-identifiability result is in line with that in the parametric analysis of the same problem in Section 3 of Koley and Dewanji (2018b). The following theorem establishes some relationship between the masking probabilities for identifiability of the likelihood.

Theorem 3.1. *If $p_2 = c\ p_1$, where $c > 0$ is some known constant, then the likelihood is identifiable.*

Proof. Note that the parameter vector $\underset{\sim}{\zeta}$ now consists of p_1 and the $F_j(\tau_i)$'s. To check the identifiability, let $\underset{\sim}{\zeta}^{(1)}$ and $\underset{\sim}{\zeta}^{(2)}$ be two values of the parameter vector and consider the equality (3.2). Equating the coefficients of $d_{\{1\}i}, d_{\{2\}i}$ and $(n_i - d_i)$ from both sides of (3.2), adding both sides of the first two equalities and then comparing with the third, we get $p_1^{(1)} = p_1^{(2)}$. Using this, it is easy to show that $F_1^{(1)}(\tau_i) = F_1^{(2)}(\tau_i)$ and $F_2^{(1)}(\tau_i) = F_2^{(2)}(\tau_i)$, by equating the coefficients of $d_{\{1\}i}$ and $d_{\{2\}i}$, respectively. This holds for all $i = 1, \cdots, K$ and hence the likelihood is identifiable. □

3.2 Estimability

When both p_1 and p_2 are known, the hessian matrix, H is a block diagonal matrix given by

$$\text{H=} \begin{pmatrix} H_{11} & 0 & 0 & \cdots & 0 \\ 0 & H_{22} & 0 & \cdots & 0 \\ 0 & 0 & H_{33} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & H_{KK} \end{pmatrix},$$

where the i^{th} block H_{ii} of dimension $m \times m$ corresponds to the i^{th} monitoring time point τ_i, for $i = 1, \cdots, K$. The $(j, j^{'})$th element of H_{ii} is given by

$$-\frac{\partial^2}{\partial F_j(\tau_i)\partial F_{j'}(\tau_i)} \log L(\underset{\sim}{F}) = \begin{cases} \dfrac{d_{\{j\}i}}{F_j(\tau_i)^2} + \dfrac{d_{\{1,2\}i}(1-p_j)^2}{\{(1-p_1)F_1(\tau_i) + (1-p_2)F_2(\tau_i)\}^2} + \dfrac{(n_i - d_i)}{S(\tau_i)^2}, & \text{if } j^{'} = j \\ \\ \dfrac{d_{\{1,2\}i}(1-p_j)(1-p_{j'})}{\{(1-p_1)F_1(\tau_i) + (1-p_2)F_2(\tau_i)\}^2} + \dfrac{(n_i - d_i)}{S(\tau_i)^2}, & \text{if } j^{'} \neq j, \end{cases}$$

for $j, j^{'} = 1, 2$. To check the positive definiteness of a matrix it is sufficient to prove all principal minors of the matrix are positive. It is clear that the first principal minor of the matrix H_{ii} is $\dfrac{d_{\{1\}i}}{F_1(\tau_i)^2} + \dfrac{d_{\{1,2\}i}(1-p_1)^2}{\{(1-p_1)F_1(\tau_i) + (1-p_2)F_2(\tau_i)\}^2} + \dfrac{(n_i - d_i)}{S(\tau_i)^2} > 0$. The other principal minor (the last one) of H_{ii} is $\mid H_{ii} \mid$. Note that

$$\mid H_{ii} \mid = \frac{d_{\{1\}i}}{F_1(\tau_i)^2}\left\{\frac{d_{\{2\}i}}{F_2(\tau_i)^2} + \frac{d_{\{1,2\}i}(1 - c\, p_1)^2}{\{(1-p_1)F_1(\tau_i) + (1 - c\, p_1)F_2(\tau_i)\}^2} + \frac{n_i - d_i}{S(\tau_i)^2}\right\}$$

$$+ \left\{\frac{d_{\{1,2\}i}(1-p_1)^2}{\{(1-p_1)F_1(\tau_i) + (1 - c\, p_1)F_2(\tau_i)\}^2} + \frac{n_i - d_i}{S(\tau_i)^2}\right\}\frac{d_{\{2\}i}}{F_2(\tau_i)^2}$$

$$+ \frac{n_i - d_i}{S(\tau_i)^2}\frac{d_{\{1,2\}i}}{\{(1-p_1)F_1(\tau_i) + (1 - c\, p_1)F_2(\tau_i)\}^2}(p_1 - c\, p_1)^2 > 0.$$

This shows the matrix H_{ii} is positive definite, for all $i = 1, \cdots, K$ and, hence, the matrix H is positive definite.

In the following, we develop a sufficient condition for positive definiteness of the hessian matrix for the likelihood function (3.1) with $p_2 = c\, p_1$ for some known $c > 0$. The observed information matrix denoted by H^*, say, is given by

$$H^* = \left[\begin{array}{c|c} H & \underset{\sim}{z} \\ \hline \underset{\sim}{z}^T & -\dfrac{\partial^2 l(\underset{\sim}{\zeta})}{\partial p_1^2} \end{array}\right],$$

where the matrix H is similar to the hessian matrix for known p_1 and p_2 with $p_2 = c\, p_1$ and $\underset{\sim}{z}$ is the vector with $((j-1) * k + i)$th element given by $-\dfrac{\partial^2}{\partial F_j(\tau_i)\partial p_1} \log L(\underset{\sim}{\zeta})$, for $j = 1, 2$ $i = 1, \cdots, K$. Following the similar argument for known p_1 and p_2, it is sufficient to prove that all principal minors of the matrix are positive. Hence, since H is positive definite, it is only required to show that the determinant of the matrix H^* is positive. Using the formula for the determinant of a partitioned matrix,

$$\mid H^* \mid = \mid H \mid (-\frac{\partial^2 l(\underset{\sim}{\zeta})}{\partial p_1^2} - \underset{\sim}{z}^T H^{-1} \underset{\sim}{z}),$$

where

$$-\frac{\partial^2 l(\underset{\sim}{\zeta})}{\partial p_1^2} = \sum_{i=1}^{K} \left[\frac{d_{\{1\}}i}{p_1^2} + \frac{d_{\{2\}}i}{p_1^2} + \frac{d_{\{1,2\}i}}{\{(1-p_1)F_1(\tau_i) + (1-c\ p_1)F_2(\tau_i)\}^2} \right] > 0.$$

For a given i, let us denote $\{(1-p_1)F_1(\tau_i) + (1-c\ p_1)F_2(\tau_i)\}^2$ by D_i^*. Then, the inverse of H_{ii} is given by

$$H_{ii}^{-1} = \frac{1}{\mid H_{ii} \mid} \begin{bmatrix} \frac{d_{\{2\}i}}{F_2(\tau_i)^2} + \frac{d_{\{1,2\}i}(1-c\ p_1)^2}{D_i^*} + \frac{n_i - d_i}{S(\tau_i)^2} & -\frac{d_{\{1,2\}i}(1-p_1)(1-c\ p_1)}{D_i^*} - \frac{n_i - d_i}{S(\tau_i)^2} \\ -\frac{d_{\{1,2\}i}(1-p_1)(1-c\ p_1)}{D_i^*} - \frac{n_i - d_i}{S(\tau_i)^2} & \frac{d_{\{1\}i}}{F_1(\tau_i)^2} + \frac{d_{\{1,2\}i}(1-c\ p_1)^2}{D_i^*} + \frac{n_i - d_i}{S(\tau_i)^2} \end{bmatrix}.$$

To show $\mid H^* \mid > 0$, it is only required to show $(-\frac{\partial^2 l(\underset{\sim}{\zeta})}{\partial p_1^2} - \underset{\sim}{z}^T H^{-1} \underset{\sim}{z}) > 0$, since H is positive definite. Note that

$$-\frac{\partial^2 l(\underset{\sim}{\zeta})}{\partial p_1^2} - \underset{\sim}{z}^T H^{-1} \underset{\sim}{z}$$

$$= \sum_{i=1}^{K} \left[\frac{d_{\{1\}i}}{p_1^2} + \frac{d_{\{2\}i}}{c\ p_1^2} + \frac{d_{\{1,2\}i}(F_1(\tau_i) + c\ F_2(\tau_i))^2}{D_i^*} \right]$$

$$- \sum_{i=1}^{K} \frac{(c-1)^2}{\mid H_{ii} \mid D_i^{*2}} \Bigg[d_{\{2\}i} + F_2(\tau_i)^2 \left\{ \frac{d_{\{1,2\}i}(1-c\ p_1)^2}{D_i^*} + \frac{n_i - d_i}{S(\tau_i)^2} \right\}$$

$$+ 2F_1(\tau_i)F_2(\tau_i) \left\{ \frac{d_{\{1,2\}i}(1-p_1)(1-c\ p_1)}{D_i^*} + \frac{n_i - d_i}{S(\tau_i)^2} \right\} + d_{\{1\}i}$$

$$+ F_1(\tau_i)^2 \left\{ \frac{d_{\{1,2\}i}(1-p_1)^2}{D_i^*} + \frac{n_i - d_i}{S(\tau_i)^2} \right\} \Bigg]$$

$$= \sum_{i=1}^{K} \Bigg[d_{\{1\}i} \Bigg\{ \frac{1}{p_1^2} - \frac{(c-1)^2}{\mid H_{ii} \mid D_i^{*2}} \Bigg\} + d_{\{2\}i} \Bigg\{ \frac{1}{c\ p_1^2} - \frac{(c-1)^2}{\mid H_{ii} \mid D_i^{*2}} \Bigg\} + \frac{d_{\{1,2\}i}}{D_i^*} (F_1(\tau_i) + c\ F_2(\tau_i))^2 \Bigg]$$

$$- \sum_{i=1}^{K} \frac{(c-1)^2}{\mid H_{ii} \mid D_i^{*2}} \Bigg[d_{\{1,2\}i} + \frac{(n_i - d_i)(1 - S(\tau_i))^2}{S(\tau_i)^2} \Bigg]. \tag{3.5}$$

Clearly, a sufficient condition for the matrix H^* to be positive definte is that the above expression in (3.5) is positive. In particular, for the MAR assumption (that is, $c = 1$), this expression reduces to

$$\sum_{i=1}^{K} d_{\{1\}i} \frac{1}{p_1^2} + d_{\{2\}i} \frac{1}{c\ p_1^2} + \frac{d_{\{1,2\}i}}{D_i^*} (F_1(\tau_i) + \ F_2(\tau_i))^2 > 0.$$

3.3 Maximum Likelihood Estimation

Theorem 3.1 provides a sufficient modeling assumption for the identifiability of the model parameters. Note that $0 < p_1, p_2 < 1$ implies $0 < p_1 < \min(1, 1/c)$, since $p_2 = c\ p_1$. The MLEs of p_1 and the $F_j(\tau_i)$'s can be obtained by maximizing the likelihood (3.1) with p_2 replaced by $c\ p_1$ and subject to $0 < p_1 < \min(1, 1/c)$. It is clear from (3.1) that the sub-distribution functions are estimable only at the monitoring time points τ_i, for $i = 1, \cdots, K$. Due to the inherent isotonic constraints in the sub-distribution functions $F_j(\cdot)$'s , for $j = 1, 2$, it becomes a constrained optimization problem. To transform it into an unconstrained one, we reparametrize the model in terms of 'interval hazards' by constructing the intervals $I_i = (\tau_{i-1}, \tau_i]$, with $\tau_0 = 0$ and defining interval hazards λ_{ji}'s, for $j = 1, 2,\ i = 1, \cdots, K$, as discussed in Koley and Dewanji (2018b). The likelihood (3.1) under $p_2 = c\ p_1$ can be rewritten in terms of these λ_{ji}'s as

$$L(p_1, \underset{\sim}{\lambda}) \propto \prod_{i=1}^{K} \Bigg[\prod_{j=1}^{2} \Bigg\{ p_j \sum_{l=1}^{i} \lambda_{jl} \prod_{l'<l} (1 - \lambda_{l'}) \Bigg\}^{d_{\{j\}i}} \Bigg\{ \sum_{j=1}^{2} (1 - p_j) \sum_{l=1}^{i} \lambda_{jl} \prod_{l'<l} (1 - \lambda_{l'}) \Bigg\}^{d_{\{1,2\}i}}$$
$$\times \Bigg\{ \prod_{l=1}^{i} (1 - \lambda_l) \Bigg\}^{n_i - d_i} \Bigg], \tag{3.6}$$

where $\underset{\sim}{\lambda}$ denotes the vector of λ_{ji}'s and $p_2 = c\ p_1$. Since $\lambda_{ji} = \dfrac{F_j(\tau_i) - F_j(\tau_{i-1})}{1 - F(\tau_i)}$ with $F_j(\tau_0) = 0$, for $j = 1, 2$ and $i = 1, \cdots, K$, the estimability of λ_{ji}'s follow from that of the $F_j(\tau_i)$'s. Maximization of the likelihood (3.6) turns out to be simpler in the absence of the isotonic constraints. To obtain the MLEs of the model parameters $(\underset{\sim}{\lambda}, p_1)$, numerical maximization of the log-likelihood function, from (3.6), is carried out using the R function *optim()*.

For known p_1 and p_2, the parameters to be estimated are only the $F_j(\tau_i)$'s, or equivalently, the interval hazards λ_{ji}'s, for $j = 1, 2$ and $i = 1, \cdots, K$. The MLE of the vector of model parameters, $\underset{\sim}{\lambda}$, is obtained by numerical maximization of the log-likelihood function, from (3.6), with known p_1 and p_2, by using the R function *optim()*.

3.4 Asymptotic Results

In order to study the asymptotic behaviour of the MLEs, note that the vector $\underset{\sim}{\zeta}$ of parameters now consists of p_1 and $\underset{\sim}{\lambda}$. At each monitoring time τ_i, there are n_i independent realizations coming from the identical density function $f_i^*(g)$, as remarked at the beginning of Section 3. Let us write this as $f_i^*(g; \underset{\sim}{\zeta})$ to make the dependence on $\underset{\sim}{\zeta}$ explicit. The likelihood function is the product of these density functions at each τ_i, as given in (3.1) and (3.5). This representation is useful in establishing the asymptotic properties of the MLE $\hat{\underset{\sim}{\zeta}} = (\hat{p}_1, \hat{\underset{\sim}{\lambda}})$ of $\underset{\sim}{\zeta} = (p_1, \underset{\sim}{\lambda})$ using the results of Lehmann and Casella (1998, p463-465). Consider the following properties:

P1 The true value of the parameter vector $\underset{\sim}{\zeta}$ denoted by $\underset{\sim}{\zeta_0} = (p_1^{(0)}, \underset{\sim}{\lambda}^{(0)})$ lies in an open set since $0 < p_1 < 1$.

P2 Suppose $\underset{\sim}{\zeta}' = (p_1', \underset{\sim}{\lambda}')$ and $\underset{\sim}{\zeta}'' = (p_1'', \underset{\sim}{\lambda}'')$ with $\underset{\sim}{\zeta}' \neq \underset{\sim}{\zeta}''$ be two values of the parameter vector $\underset{\sim}{\zeta}$. Then, the corresponding density functions $f_i^*(g; \underset{\sim}{\zeta}')$ and $f_i^*(g; \underset{\sim}{\zeta}'')$ are not equal for some g and i. The likelihood functions being the product of these densities, we have $L(\underset{\sim}{\zeta}') \neq L(\underset{\sim}{\zeta}'')$.

P3 For each $i = 1, 2, \cdots, K$, $E_{i, \underset{\sim}{\zeta}}\left[\frac{\partial}{\partial(\underset{\sim}{\zeta})} \log f_i^*(g; \underset{\sim}{\zeta})\right] = 0$ and

$$E_{i, \underset{\sim}{\zeta}}\left[-\frac{\partial^2}{\partial \underset{\sim}{\zeta} \partial \underset{\sim}{\zeta}'} \log f_i^*(g; \underset{\sim}{\zeta})\right] = E_{i, \underset{\sim}{\zeta}}\left[\left(\frac{\partial}{\partial \underset{\sim}{\zeta}} \log f_i^*(g; \underset{\sim}{\zeta})\right)^2\right] = \mathcal{I}_i(\underset{\sim}{\zeta}),$$ say, is assumed to be positive definite.

P4 For a given i, the density function $f_i^*(g; \underset{\sim}{\zeta})$ is linear in p_1 and quadratic in each component of $\underset{\sim}{\lambda}$. Hence, it is continuous in each component of the parameter vector $\underset{\sim}{\zeta}$ admitting all third order partial derivatives, which are bounded by functions with finite expectations.

Theorem 3.2. *Under the assumption* $\frac{n_i}{n} \to w_i$*, where* w_i*'s are positive constants for all* $i = 1, 2, \cdots, k$ *with* $\sum_{i=1}^{k} w_i = 1$*, and the properties* ***P1-P4****, we have*

1. $\hat{\underset{\sim}{\zeta}} \xrightarrow{P} \underset{\sim}{\zeta}_0$,

2. $\sqrt{n}(\hat{\underset{\sim}{\zeta}} - \underset{\sim}{\zeta}_0)$ *is asymptotically a normal random vector with mean vector* $\underset{\sim}{0}$ *and variance-covariance matrix* $\left[\sum_{i=1}^{k} w_i \mathcal{I}_i(\underset{\sim}{\zeta}_0)\right]^{-1}$.

Proof. The proof directly follows from Theorem 7.5.2 (See Lehmann and Casella, 1998, p463-465) using **P1-P4.** □

Since $\hat{\underset{\sim}{\zeta}}$ is a consistent estimator of $\underset{\sim}{\zeta}$, from the first part of Theorem 3.2, using invariance property of a consistent estimator and the Weak Law of Large Numbers, $-\frac{1}{n}\frac{\partial^2}{\partial \underset{\sim}{\zeta} \partial \underset{\sim}{\zeta}'} \log L(\underset{\sim}{\zeta}) = -\sum_{i=1}^{K} \frac{n_i}{n} \frac{1}{n_i} \sum_{l=1}^{n_i} \frac{\partial^2}{\partial \underset{\sim}{\zeta} \partial \underset{\sim}{\zeta}'} \log f_i^*(g_{il}; \underset{\sim}{\zeta})$, evaluated at $\underset{\sim}{\zeta} = \hat{\underset{\sim}{\zeta}}$, can be taken as a consistent estimator of $\sum_{i=1}^{K} w_i \mathcal{I}_i(\underset{\sim}{\zeta}_0)$.

When p_1 and p_2 are known, same asymptotic results, as in Theorem 3.2, follow by using the similar arguments.

To estimate the variance of $\hat{F}_j(\tau_i)$, for $j = 1, 2$ and $i = 1, \cdots, K$, note that $F_j(\tau_i) = \sum_{l=1}^{i} \lambda_{jl} \prod_{l'=1}^{l-1} (1 - \lambda_{l'}) = g_{ji}(\underset{\sim}{\lambda})$, say, which is a continuous function of $\underset{\sim}{\lambda}$. Therefore, using the delta method, the estimated variance of $\hat{F}_j(\tau_i)$ can be obtained by using the estimated variance-covariance matrix of $\hat{\underset{\sim}{\lambda}}$.

4 Random monitoring time

Recall that, in general, X is the random variable denoting the monitoring time for each individual. Note that the current status data at the monitoring time for each individual gives an interval within which the unobserved failure time for that individual lies. Let A_i denote this interval for the i^{th} individual. Then A_i is either of the form $(0, x_i)$ or (x_i, ∞), where x_i is the monitoring time point for the i^{th} individual, for $i = 1, \cdots, n$. Following Hudgens et al. (2001) (See also Turnbull (1976) and Gentleman and Geyer (1994)), the non-parametric maximum likelihood estimators of the two sub-distribution functions, based on the likelihood $\prod_{i=1}^{n} f^*(x_i, g_i)$ of Section 2, have mass concentrated on the union of m_j closed and disjoint intervals given by $C_j = \bigcup_{l=1}^{m_j} [q_{jl}, p_{jl}] = \bigcup_{l=1}^{m_j} C_{jl}$, say, the construction of which is discussed below.

For a fixed j, let $D_j = \{i \colon g_i \ni j\}$ and C the set of censored individuals. We assume that the set D_j is non-empty, for $j = 1, 2$; otherwise, $F_j(\cdot)$ cannot be estimated. Define $L_j = \{x_i \colon i \in C\} \bigcup \{0\}$ and $R_j = \{x_i \colon i \in D_j\} \bigcup \{\infty\}$ depending on the end points of sets

A_i, for $i = 1, \cdots, n$. The closed, disjoint intervals $C_{jl}(l = 1, \cdots, m_j)$ are constructed in such a way that the left and right end points of the intervals belong to L_j and R_j, respectively, and no member of L_j or R_j lies within the intervals. For $l = 1, \cdots, m_j$ and $j = 1, 2$, define

$$\alpha_{ijl} = \begin{cases} 1, & \text{if } [q_{jl}, p_{jl}] \subseteq A_i, \\ 0, & \text{otherwise,} \end{cases}$$

for all $i \in D_j$. Write $\phi_{jl} = P\left[T \in [q_{jl}, p_{jl}]\right] = F_j(p_{jl}) - F_j(q_{jl})$. It is clear that the likelihood can be written in terms of the ϕ_{jl}'s and, hence, it is independent of the behavior of $F_j(\cdot)$ within the intervals C_{jl}, for $l = 1, \cdots, m_j,\ j = 1, 2$. Define $\underset{\sim}{\phi}$ to be the vector of all the ϕ_{jl}'s. Then, the likelihood can be written in terms of p_1, p_2 and $\underset{\sim}{\phi}$ as

$$L_I(\underset{\sim}{\phi}, p_1, p_2) = \left\{ \prod_{i:\, i \in D} \sum_{j \in g_i} p_{g_i j} \sum_{l=1}^{m_j} \alpha_{ijl} \phi_{jl} \right\} \left\{ \prod_{i:\, i \in C} \sum_{j=1}^{2} \sum_{l=1}^{m_j} \alpha_{ijl} \phi_{jl} \right\}, \tag{4.1}$$

where $p_{\{j\}j} = p_j$ and $p_{\{1,2\}j} = 1 - p_j$, for $j = 1, 2$, $\phi_{jl} \geq 0$, $\sum_{j=1}^{2} \sum_{l=1}^{m_j} \phi_{jl} = 1$ and $D = D_1 \bigcup D_2$ is the set of all individuals observed to have failed. The estimates of p_1, p_2 and $\underset{\sim}{\phi}$, obtained by maximizing this likelihood, are termed as non-parametric maximum likelihood estimates (NPMLEs). We consider two situations: (i) p_1, p_2 both known, and (ii) $p_2 = c\, p_1$, for some known constant $c > 0$.

The MLE of the parameter vector $\underset{\sim}{\phi}$ with both p_1 and p_2 known is obtained by maximizing the likelihood (4.1) using the EM algorithm (See Dempster et al. (1977)) as described below. Let us define the indicator function

$$\mathcal{I}_{ijl} = \mathbb{I}\left[T_i \in C_{jl}, J_i = j\right],$$

for $i = 1, \cdots, n,\ l = 1, \cdots, m_j$ and $j = 1, 2$. The value of this indicator function is unknown in the observed data as the exact failure time is never observed and the exact cause is not always observed. However, in the complete data version, it is assumed to be known. The complete data likelihood is therefore given by

$$L_C(\underset{\sim}{\phi}) = \left\{ \prod_{i:\, i \in D} \prod_{j \in g_i} \prod_{l=1}^{m_j} \{p_{g_i j} \phi_{jl}\}^{\mathcal{I}_{ijl}} \right\} \times \left\{ \prod_{i:\, i \in C} \prod_{j=1}^{2} \prod_{l=1}^{m_j} \{\phi_{jl}\}^{\mathcal{I}_{ijl}} \right\}$$

At each iteration, the E-step computes the conditional expectation of the complete data log-likelihood, given the observed data and the estimates of $\underset{\sim}{\phi}$ obtained from the previous iteration. Let the estimate of $\underset{\sim}{\phi}$ at the uth iteration be denoted by $\underset{\sim}{\phi}^{(u)}$. Then, at the

$(u+1)$st iteration, we have

$$E\left[\mathcal{I}_{ijl} \mid \text{observed data}, \underset{\sim}{\phi}^{(u)}\right] = I_{ijl}(\underset{\sim}{\phi}^{(u)}), \quad \text{say.}$$

Note that

$$I_{ijl}(\underset{\sim}{\phi}^{(u)}) = \begin{cases} P\left[T \in C_{jl}, J_i = j \mid T \in A_i, G = g_i; \underset{\sim}{\phi}^{(u)}\right], & \text{if } g_i \neq \phi \\ \\ P\left[T \in C_{jl}, J_i = j \mid T \in A_i; \underset{\sim}{\phi}^{(u)}\right], & \text{if } g_i = \phi \end{cases}$$

$$= \begin{cases} \dfrac{p_{g_i j}\alpha_{ijl}\phi_{jl}^{(u)}}{\sum_{j' \in g_i} \sum_{l'=1}^{m_{j'}} p_{g_i j'} \alpha_{ij'l'} \phi_{j'l'}^{(u)}}, & \text{if } g_i \neq \phi \\ \\ \dfrac{\alpha_{ijl}\phi_{jk}^{(u)}}{\sum_{j'=1}^{2} \sum_{l'=1}^{m_{j'}} \alpha_{ij'l'} \phi_{j'l'}^{(u)}}, & \text{if } g_i = \phi. \end{cases}$$

Clearly, the expectation of the complete data log-likelihood can be obtained by replacing the unobserved quantities (the $\mathcal{I}_{ijl}$'s) with their conditional expectations (the $I_{ijl}(\underset{\sim}{\phi}^{(u)})$'s), as obtained in the uth iteration of the E-step. The M-step maximizes this expected complete data log-likelihood to obtain the next improved estimate of ϕ_{jl} as

$$\phi_{jl}^{(u+1)} = \frac{\sum_{i:\, i \in D, g_i \ni j} I_{ijl}(\underset{\sim}{\phi}^{(u)}) + \sum_{i \in C} I_{ijl}(\underset{\sim}{\phi}^{(u)})}{\sum_{j=1}^{m} \sum_{l=1}^{m_j} \left[\sum_{i:\, i \in D, g_i \ni j} I_{ijl}(\underset{\sim}{\phi}^{(u)}) + \sum_{i \in C} I_{ijl}(\underset{\sim}{\phi}^{(u)})\right]},$$

for all j, l. The EM algorithm iterates between E- and M- steps until convergence.

The MLEs of the parameter vector $\underset{\sim}{\phi}$ and p_1, for $p_2 = c\, p_1$, $c > 0$ known, are obtained in the similar manner. As before, let us denote the estimates of $\underset{\sim}{\phi}$ and p_1 at the uth iteration by $\underset{\sim}{\phi}^{(u)}$ and $p_1^{(u)}$, respectively. Then, at the $(u+1)$st iteration, we have

$$E\left[\mathcal{I}_{ijl} \mid \underset{\sim}{\phi}^{(u)}, p_1^{(u)}\right] = I_{ijl}^{1}(\underset{\sim}{\phi}^{(u)}, p_1^{(u)}), \quad \text{say, where,}$$

$$I_{ijl}^1(\underset{\sim}{\phi}^{(u)}, p_1^{(u)}) = \begin{cases} P\left[T \in C_{jl}, J_i = j \mid T \in A_i, G = g_i; \underset{\sim}{\phi}^{(u)}, p_1^{(u)}\right], & \text{if } g_i \neq \phi \\ \\ P\left[T \in C_{jl}, J_i = j \mid T \in A_i; \underset{\sim}{\phi}^{(u)}\right], & \text{if } g_i = \phi \end{cases}$$

$$= \begin{cases} \dfrac{p_{g_i j}^{(u)} \alpha_{ijl} \phi_{jl}^{(u)}}{\sum_{j' \in g_i} \sum_{l'=1}^{m_{j'}} p_{g_i j'}^{(u)} \alpha_{ij'l'} \phi_{j'l'}^{(u)}}, & \text{if } g_i \neq \phi \\ \\ \dfrac{\alpha_{ijl} \phi_{jk}^{(u)}}{\sum_{j'=1}^{2} \sum_{l'=1}^{m_{j'}} \alpha_{ij'l'} \phi_{j'l'}^{(u)}}, & \text{if } g_i = \phi. \end{cases}$$

Similarly, the expectation of the complete data log-likelihood is obtained by replacing the unobservable quantities ($\mathcal{I}_{ijl}$'s) by their conditional expectations (I_{ijl}^1) as obtained in the E-step. The M-step maximizes this expected complete data log-likelihood to obtain the next improved estimate of ϕ_{jl} as

$$\phi_{jl}^{(u+1)} = \frac{\sum_{i:\, i \in D, g_i \ni j} I_{ijl}(\underset{\sim}{\phi}^{(u)}, p_1^{(u)}) + \sum_{i \in C} I_{ijl}(\underset{\sim}{\phi}^{(u)}, p_1^{(u)})}{\sum_{j=1}^{m} \sum_{l=1}^{m_j} \left[\sum_{i:\, i \in D, g_i \ni j} I_{ijl}(\underset{\sim}{\phi}^{(u)}, p_1^{(u)}) + \sum_{i \in C} I_{ijl}(\underset{\sim}{\phi}^{(u)}, p_1^{(u)})\right]},$$

for all j, l. The next improved estimate $p_1^{(u+1)}$ of p_1 is obtained by solving the following numerical equation

$$\sum_{j=1}^{2} \frac{\sum_{i:\, i \in D, g_i = \{j\}} \sum_{l=1}^{m_j} I_{ijl}^1(\underset{\sim}{\phi}^{(u)}, p_1^{(u)})}{p_j^{(u+1)}} - \sum_{i:\, i \in D, g_i = \{1,2\}} \left[\frac{\sum_{l=1}^{m_1} I_{i1l}^1(\underset{\sim}{\phi}^{(u)})}{1 - p_1^{(u+1)}} + \frac{\sum_{l=1}^{m_2} c I_{i2l}^1(\underset{\sim}{\phi}^{(u)})}{1 - c p_1^{(u+1)}}\right] = 0.$$

The EM algorithm iterates between E- and M- steps until convergence.

The NPMLE of the sub-distribution function $F_j(\cdot)$ is obtained (in both the cases) from that of $\underset{\sim}{\phi}$ as

$$\hat{F}_j(x) = \begin{cases} 0, & \text{if } x \leq p_{j1} \\ \sum_{l=1}^{k} \hat{\phi}_{jl}, & \text{if } p_{jk} < x \leq p_{j\overline{k+1}}; 1 \leq k < (m_j - 1) \\ \sum_{l=1}^{m_j} \hat{\phi}_{jl}, & \text{if } x > p_{jm_j}, \end{cases} \tag{4.2}$$

for $j = 1, 2$.

From Theorem 1 of Dempster et al. (1977), it follows that the observed likelihood function evaluated at the improved estimates at successive EM iterations is an increasing sequence. The observed likelihood function being a product of probabilities is bounded above by 1. So, this sequence of the observed likelihood functions converges. Note that the mapping $\underset{\sim}{\phi}^{(u)} \rightarrow \underset{\sim}{\phi}^{(u+1)}$, or $(p_1^{(u)}, \underset{\sim}{\phi}^{(u)}) \rightarrow (p_1^{(u+1)}, \underset{\sim}{\phi}^{(u+1)})$, is continuous. Then, following Theorem 2 of Wu (1983), it is clear that the sequence converges monotonically to a local maxima.

Now, consider the matrix $A = (\alpha^*_{ijl})$ of dimension $n \times \sum_{j=1}^2 m_j$, where

$$\alpha^*_{ijl} = \begin{cases} p_{g_i j}\alpha_{ijl}, & \text{if } g_i \in G \setminus \phi \\ \alpha_{ijl}, & \text{if } g_i = \phi, \end{cases}$$

for $i = 1, \cdots, n,\ j = 1, 2$ and $l = 1, \cdots, m_j$.

Theorem 4.1. *For known p_1 and p_2, a sufficient condition for the uniqueness of the maximum likelihood estimate $\hat{\underset{\sim}{\phi}}$ is that the matrix A is of full column rank.*

Proof. Reusing the notation of Section 2, let us define H to be the $\sum_{j=1}^2 m_j \times \sum_{j=1}^2 m_j$ hessian matrix of the observed likelihood function (4.1) with known p_1 and p_2. Clearly, $H = A^T D A$, where D is a $n \times n$ diagonal matrix with diagonal entries

$$\begin{cases} \left(\sum_{j\in g_i} \sum_{l=1}^{m_j} p_{g_i j}\alpha_{ijk}\phi_{jk}\right)^{-2}, & \text{if } g_i \in G \setminus \phi \\ \left(\sum_{j=1}^{2} \sum_{k=1}^{m_j} \alpha_{ijk}\phi_{jk}\right)^{-2}, & \text{if } g_i = \phi\ , \end{cases}$$

Clearly, the matrix D is positive definite and, so, H is non-negative definite. However, if the matrix A is of full column rank, then H is positive definite. Hence, the observed likelihood has a unique maxima and so the local maxima is also the global one. □

Theorem 4.2. *For $p_2 = cp_1$, $c > 0$ known, if the matrix A is of full column rank and $-\frac{\partial^2 l(\underset{\sim}{\phi}, p_1)}{\partial p_1^2} - C^T H^{-1} C > 0$, where C is the vector of dimension $(\sum_{j=1}^2 m_j + 1) \times 1$ having the $((j-1)m_j + l)^{th}$ entry as $-\frac{\partial l(\underset{\sim}{\phi}, p_1)}{\partial \phi_{jl}\partial p_1}$, for $l = 1, \cdots, m_j$ and $j = 1, 2$, then the maximum likelihood estimates of $\underset{\sim}{\phi}$ and p_1 are unique.*

Proof. The hessian matrix H^1 of the likelihood (4.1), with $p_2 = c\ p_1$, $c > 0$ known, is of dimension $(\sum_{j=1}^2 m_j + 1) \times (\sum_{j=1}^2 m_j + 1)$ and can be written in terms of the matrix H, as defined for the case of known p_1, p_2, as

$$H^1 = \left[\begin{array}{c|c} H & C \\ \hline C^T & -\frac{\partial^2 l(\underset{\sim}{\phi}, p_1)}{\partial p_1^2} \end{array}\right].$$

From the definition of block matrix,

$$| H^1 |=| H || \frac{\partial^2 l(\underset{\sim}{\phi}, p_1)}{\partial p_1^2} - C^T H^{-1} C |,$$

where H^{-1} is the inverse of H. Note that, using the assumptions, H is positive definite and we also have $| H^1 |> 0$. Hence, the local maxima is also the global maxima. □

It is well-known from the convergence results of Groeneboom et al. (2008b,a) that the asymptotic distribution of the estimators is non-standard and not simply a normal distribution. Hence, the standard errors of the estimates cannot possibly be based on the observed information matrix. We, therefore, resort to using the bootstrap resampling technique, as discussed below, to obtain the standard errors. After estimating the quantities of interest as mentioned above, based on the observed data $\{(x_i, g_i \; i = 1, \cdots, n\}$, we repeatedly draw a number of, say B, bootstrap samples with replacement from the original sample (that is, the observed data set) each of the same size as the original sample. For each bootstrap sample, we obtain the estimates of the quantities of interest, termed as bootstrap estimates. The sample standard deviation of these estimates across the B bootstrap samples gives the estimate of the corresponding standard error. This method is commonly referred to as non-parametric bootstrap sampling in the literature.

5 Simulation Studies

In this section, several simulation studies have been carried out to study the finite sample behaviour of the NPMLEs obtained in Sections 3 and 4. We describe the simulation procedures along with the results for fixed and random monitoring times in the following two subsections, respectively.

5.1 Fixed monitoring time

In all the simulation studies, the failure time distribution is taken to be exponential with rate 1 and the two types of failure occurring with rate ratio 6 : 4. We first simulate n observations on the failure time T from the assumed distribution. We take $K = 5$ and the five fixed monitoring time points $(\tau_1, \cdots, \tau_5)$ are chosen as the $10^{th}, 25^{th}, 50^{th}, 75^{th}$ and 90^{th} quantiles of the true exponential failure time distribution with rate 1, which are $0.105, 0.288, 0.693, 1.386, 2.303$, respectively. The first n_1 observations are monitored at time τ_1, the next n_2 observations at time τ_2, and so on, such that $n_1 + \cdots + n_5 = n$. If an individual monitored at time τ_i, say, is observed to have failed, then the observed set of possible causes of failure is generated with probability

$$P\left[G=g \mid T \leq \tau_i\right]=\frac{\sum_{j \in g} p_{gj} F_j(\tau_i)}{F(\tau_i)}, \quad \text{for } g \in \{\{1\},\{2\},\{1,2\}\},$$

for $i = 1, \cdots, K = 5$. This leads to a simulated data set $\{(x_i, g_i); i = 1, \cdots, n\}$, where g is taken as the empty set ϕ if the individual is censored. We carry out 10000 simulations to obtain 10000 such data sets. For each simulated data set, we obtain the maximum likelihood estimates of the two sub-distribution functions at the five monitoring times (τ_i's), and that of the probability p_1 in the second case, along with the corresponding standard errors using the method of Section 3. Then, from the 10000 such estimates, we estimate the corresponding bias by subtracting the average of the estimates from its true value. The average of the standard errors over 10000 simulated data sets, denoted by ASE, are also obtained. Sample standard errors, denoted by SSE, are computed as the square root of the sample variances of the estimates over 10000 simulations. Finally, the approximate 95% confidence intervals for each of the $F_j(\tau_i)$'s are constructed by using normal approximation of the estimators. The cover percentage is estimated as the proportion of times these intervals contain the true value among 10000 simulations and is denoted by CP.

The simulation is carried out with three different choices of n_i's, assumed to be equal, as $50, 150$ and 250. First, we consider $(p_1, p_2) = (0.9, 0.8)$ and (0.7,0.6) which are assumed to be known in the analysis (See Table 1). Then, we take $p_1 = 0.9$, $p_2 = cp_1$ with $c = 0.8$ and 1, and only c is assumed known in the analysis (See Tables 2 and 3, respectively). In the first case with known p_1 and p_2 (Table 1), the results on Bias, SSE, ASE and CP indicate more precise and efficient estimates with increasing sample size, as expected, providing some evidence in favour of asymptotic normality. Note that, with the decrease in $p_j (j = 1, 2)$ values, the standard errors of the estimates increase, as expected, since smaller p_j value means more number of observations with missing causes. For the second case (Tables 2 and 3), the results on Bias, SSE, ASE and CP are again qualitatively similar with increasing sample size when the correct value of c is used for the analysis. However, the results are seen to be senstitive to the misspecifiaction of the value of c used in the analysis, as expected, affecting both bias and CP, as is evident with large sample size. We also notice that, if the value of c used in the analysis is larger (smaller) than the true value, then $F_1(\tau_i))$ is over (under)-estimated while $F_2(\tau_i))$ is under (over)-estimated. A possible reasoning for this is as follows. If the value of c used in the analysis is smaller than the true one, there is a tendency to over-estimate p_1 to counter-balance the opposite force on p_2. This results in less representation of type 1 failures among the observations with missing failure type which, in turn, under-estimates $F_1(\cdot)$ and over-estimates $F_2(\cdot)$. Similarly, if the value of c used in the analysis is larger than the true one, p_1 is under-estimated and, following the same argument, $F_1(\cdot)$'s are over-estimated and $F_2(\cdot)$'s are under-estimated. The standard errors are also seen to follow the similar pattern.

5.2 Random monitoring time

As before, we simulate n observations from the assumed failure time distribution which is taken to be exponential with mean rate $\lambda = 1$ and two types of failures occurring with rate ratio 6: 4. Let us define Dom X to be the domain of the monitoring time random variable X. The monitoring time distribution is also taken to be exponential but with rate 0.8. If an individual monitored at $x \in$Dom X is observed to have failed, then the corresponding observed set of possible causes is assigned with probability

$$P[G = g \mid T \leq x] = \frac{\sum_{j \in g} p_{gj} F_j(x)}{F(x)},$$

for $g \in \mathcal{G} \setminus \phi = \{\{1\}, \{2\}, \{1,2\}\}$. Thus we obtain a simulated dataset $\{(x_i, g_i) \colon i = 1, \cdots, n\}$, where $g = \phi$ if $T > X = x$. Repeat this 10000 times to get 10000 such simulated data sets.

We fix five time points as the $10^{th}, 25^{th}, 50^{th}, 75^{th}$ and 90^{th} quantiles of the assumed failure time distribution (that is, the exponential distribution with rate 1). For each simulated dataset, using the methods of Section 4, the maximum likelihood estimates of the sub-distribution functions at these fixed time points, and that of the probability p_1 in the second case, are obtained. The corresponding standard errors of the estimators are obtained using the bootstrap resampling technique as discussed at the end of Section 4. The average of these standard errors over 10000 simulations, denoted by ASE, is also computed. Finally, sample standard errors over the 10000 simulations, denoted by SSE, are obtained and compared with ASE. In all the simulation studies B is taken to be 500 and the results are presented in Table 4-6 for known (p_1, p_2) and $p_2 = c\, p_1$ with $c = 0.8$ and 1, respectively.

Three different choices of sample size $n = 50, 150$ and 250 are considered to study the estimates with increasing sample size in all the simulation studies. We consider the following different choices of the masking probabilities to describe the different cases: (i) $(p_1, p_2) = (0.9, 0.8)$ and $(0.7, 0.6)$ and both are assumed to be known in the analysis (See Table 4), and then (ii) $p_1 = 0.9$, $p_2 = c\, p_1$ with $c = 0.8, 1$ and only c is assumed known in the analysis (See Tables 5 and 6). As expected, bias, SSE and ASE decrease with increase in sample size. For case (i), similar to the fixed monitoring case of Section 5.1, the number of missing observations increases with the decrease in (p_1, p_2), resulting in the increase of the standard errors of the estimators. For case (ii), the results are again seen to be sensitive to the misspecification of the value of c and the similar interpretation follows as that of the fixed monitoring case. The CP values (not reported) obtained by assuming normal distribution of the estimators does not however converge to 0.95 with increasing sample size. This is expected in view of the convergence results of Groeneboom et al. (2008a,b) which essentially state that the asymptotic distribution of the NPMLEs are not simply normal.

Table 1: Simulation results on $\left\{\hat{F}_j(\tau_i), j = 1, 2,\ i = 1, \cdots, 5\right\}$ under exponential(1) failure time distribution with two types (6: 4 rate ratio) and known (p_1, p_2) for five fixed monitoring times τ_i's. The five τ_i's are 0.105, 0.288, 0.693, 1.386 and 2.303, respectively.

		Cause 1				Cause 2			
(p_1, p_2)	n_i	$10 \times \lvert\text{Bias}\rvert$	ASE	SSE	CP	$10 \times \lvert\text{Bias}\rvert$	ASE	SSE	CP
(0.9,0.8)	50	0.006	0.047	0.033	0.925	0.025	0.053	0.028	0.993
		0.022	0.053	0.052	0.937	0.008	0.046	0.042	0.928
		0.017	0.067	0.064	0.942	0.008	0.060	0.054	0.938
		0.037	0.074	0.063	0.953	0.032	0.069	0.059	0.971
		0.040	0.076	0.062	0.980	0.052	0.074	0.059	0.983
	150	0.003	0.020	0.019	0.931	0.003	0.017	0.016	0.917
		0.010	0.030	0.031	0.940	0.005	0.026	0.026	0.942
		0.015	0.039	0.039	0.947	0.007	0.035	0.035	0.941
		0.013	0.043	0.043	0.980	0.012	0.040	0.037	0.965
		0.019	0.043	0.040	0.962	0.008	0.042	0.039	0.969
	250	0.001	0.016	0.016	0.945	1.01×10^{-3}	0.013	0.013	0.920
		0.006	0.023	0.022	0.955	2.33×10^{-3}	0.020	0.020	0.954
		0.008	0.030	0.031	0.950	7.87×10^{-4}	0.027	0.027	0.954
		0.009	0.033	0.033	0.948	0.007	0.031	0.030	0.955
		0.007	0.034	0.032	0.960	0.007	0.033	0.032	0.957
(0.7,0.6)	50	0.025	0.054	0.038	0.926	0.040	0.054	0.031	1.000
		0.076	0.058	0.053	0.937	0.015	0.053	0.045	0.972
		0.022	0.074	0.067	0.943	0.013	0.067	0.062	0.953
		0.039	0.083	0.068	0.983	0.035	0.079	0.060	0.985
		0.040	0.086	0.071	0.980	0.042	0.084	0.069	0.983
	150	0.007	0.022	0.022	0.957	0.015	0.019	0.018	0.928
		0.062	0.033	0.033	0.945	0.006	0.029	0.028	0.931
		0.017	0.043	0.043	0.945	0.010	0.039	0.037	0.952
		0.017	0.048	0.044	0.969	0.016	0.045	0.040	0.968
		0.033	0.049	0.044	0.975	0.031	0.048	0.043	0.970
	250	0.003	0.017	0.017	0.943	0.005	0.014	0.014	0.939
		0.017	0.025	0.025	0.945	0.003	0.022	0.022	0.936
		0.009	0.033	0.034	0.954	0.005	0.030	0.030	0.950
		0.010	0.037	0.035	0.968	0.008	0.035	0.032	0.964
		0.011	0.038	0.036	0.966	0.009	0.037	0.035	0.961

Table 2: Simulation results on $\left\{\hat{F}_j(\tau_i), j=1,2,\ i=1,\cdots,5\right\}$ and $\hat{p}_1$ under exponential(1) failure time distribution with two types (6: 4 rate ratio), $p_1 = 0.9$, $p_2 = c\ p_1$ and known $c = 0.8$ for five fixed monitoring times τ_i's. The five τ_i's are 0.105, 0.288, 0.693, 1.386 and 2.303, respectively.

		Cause 1				Cause 2			
Value of c used in the analysis	n_i	10× Bias	ASE	SSE	CP	10× Bias	ASE	SSE	CP
0.8	50	1.03×10^{-2}	0.047	0.032	0.920	-0.032	0.056	0.028	0.999
		1.52×10^{-2}	0.053	0.050	0.935	-0.006	0.046	0.041	0.957
		-1.54×10^{-2}	0.069	0.065	0.944	0.035	0.064	0.054	0.963
		8.02×10^{-2}	0.075	0.063	0.974	0.033	0.070	0.056	0.976
		-1.25×10^{-2}	0.076	0.066	0.974	0.026	0.074	0.062	0.982
$\hat{p}_1 = 0.899$, ASE= 0.040, SSE= 0.031, CP=0.934									
1	50	0.045	0.048	0.035	0.933	-0.060	0.056	0.027	0.989
		0.104	0.054	0.052	0.962	-0.118	0.046	0.037	0.949
		0.266	0.070	0.068	0.941	-0.274	0.057	0.051	0.914
		0.382	0.076	0.064	0.955	-0.409	0.067	0.054	0.934
		0.403	0.077	0.067	0.945	-0.391	0.072	0.060	0.951
$\hat{p}_1 = 0.829$, ASE= 0.034, SSE=0.030 , CP=0.435									
0.8	150	-1.02×10^{-3}	0.020	0.019	0.925	0.003	0.017	0.016	0.926
		-2.91×10^{-3}	0.030	0.030	0.945	-0.005	0.026	0.026	0.943
		9.18×10^{-3}	0.040	0.039	0.954 -	0.013	0.035	0.034	0.958
		-3.35×10^{-2}	0.043	0.042	0.960	0.026	0.040	0.037	0.970
		1.15×10^{-2}	0.044	0.041	0.971	0.022	0.043	0.040	0.963
$\hat{p}_1 = 0.900$, ASE= 0.022, SSE= 0.021, CP=0.942									
1	150	0.052	0.021	0.021	0.953	-0.050	0.017	0.015	0.885
		0.122	0.031	0.031	0.944	-0.122	0.024	0.023	0.886
		0.251	0.040	0.039	0.913	-0.259	0.033	0.031	0.871
		0.403	0.044	0.043	0.849	-0.410	0.039	0.035	0.813
		0.444	0.045	0.041	0.843	-0.441	0.041	0.037	0.833
$\hat{p}_1 = 0.827$, ASE= 0.020, SSE= 0.019, CP=0.02									
0.8	250	9.49×10^{-4}	0.015	0.015	0.930	-0.002	0.013	0.013	0.931
		1.40×10^{-3}	0.024	0.024	0.949	0.004	0.020	0.020	0.947
		-1.67×10^{-3}	0.031	0.031	0.949	-0.010	0.027	0.026	0.948
		1.38×10^{-2}	0.034	0.033	0.959	0.007	0.031	0.031	0.954
		-2.91×10^{-3}	0.034	0.032	0.962	0.007	0.033	0.031	0.960
$\hat{p}_1 = 0.900$, ASE= 0.017, SSE= 0.016, CP=0.944									
1	250	0.060	0.016	0.017	0.936	-0.052	0.012	0.011	0.883
		0.134	0.024	0.025	0.919	-0.013	0.019	0.019	0.863
		0.262	0.032	0.032	0.866	-0.254	0.026	0.026	0.808
		0.399	0.035	0.033	0.790	-0.400	0.030	0.028	0.731
		0.451	0.035	0.032	0.733	-0.449	0.032	0.029	0.727
$\hat{p}_1 = 0.828$, ASE= 0.015, SSE=0.015 , CP=0.002									

Table 3: Simulation results on $\left\{\hat{F}_j(\tau_i), j=1,2,\ i=1,\cdots,5\right\}$ and $\hat{p}_1$ under exponential(1) failure time distribution with two types (6 : 4 rate ratio), $p_1 = 0.9$, $p_2 = c\ p_1$ and known $c = 1$ for five fixed monitoring times τ_i's. The five τ_i's are 0.105, 0.288, 0.693, 1.386 and 2.303, respectively.

		Cause 1				Cause 2			
Value of c in the analysis	n_i	10× Bias	ASE	SSE	CP	10× Bias	ASE	SSE	CP
1	50	0.061	0.044	0.034	0.891	-0.017	0.056	0.028	0.997
		0.127	0.052	0.048	0.945	-0.021	0.046	0.040	0.927
		-0.213	0.066	0.063	0.942	0.040	0.058	0.054	0.955
		-0.371	0.073	0.066	0.973	0.041	0.067	0.057	0.968
		0.428	0.074	0.065	0.974	-0.058	0.072	0.060	0.986
$\hat{p}_1 = 0.901$, ASE= 0.027, SSE= 0.029, CP=0.937									
0.8	50	-0.069	0.047	0.031	0.845	0.051	0.056	0.028	0.993
		-0.156	0.049	0.047	0.931	0.158	0.046	0.043	0.952
		-0.298	0.064	0.061	0.922	0.211	0.060	0.056	0.956
		-0.447	0.071	0.061	0.936	0.333	0.068	0.058	0.968
		-0.542	0.072	0.060	0.939	0.457	0.077	0.060	0.999
$\hat{p}_1 = 0.982$, ASE= 0.033, SSE= 0.029, CP= 0.209									
1	150	0.004	0.019	0.020	0.917	0.006	0.017	0.016	0.913
		-0.013	0.030	0.029	0.954	0.017	0.026	0.025	0.944
		0.019	0.039	0.038	0.949	-0.021	0.034	0.033	0.946
		-0.024	0.042	0.040	0.968	0.018	0.039	0.036	0.956
		0.011	0.043	0.038	0.973	-0.010	0.041	0.037	0.972
$\hat{p}_1 = 0.900$, ASE= 0.016, SSE= 0.015, CP=0.938									
0.8	150	-0.055	0.019	0.018	0.920	0.045	0.017	0.016	0.954
		-0.052	0.019	0.018	0.920	0.049	0.017	0.016	0.954
		-0.124	0.028	0.028	0.896	0.126	0.026	0.025	0.935
		-0.274	0.037	0.035	0.886	0.269	0.035	0.034	0.907
		-0.384	0.041	0.039	0.851	0.387	0.040	0.037	0.843
		-0.474	0.041	0.039	0.801	0.484	0.042	0.038	0.797
$\hat{p}_1 = 0.987$, ASE= 0.018, SSE= 0.014, CP= 0.003									
1	250	0.003	0.015	0.015	0.934	-0.004	0.013	0.013	0.928
		0.009	0.023	0.023	0.948	-0.008	0.020	0.020	0.950
		- 0.012	0.030	0.031	0.951	0.002	0.026	0.027	0.947
		0.003	0.033	0.032	0.950	0.017	0.030	0.030	0.952
		-0.006	0.033	0.032	0.954	-0.003	0.032	0.030	0.962
$\hat{p}_1 = 0.9$, ASE= 0.012, SSE= 0.012, CP=0.949									
0.8	250	-0.061	0.014	0.014	0.885	0.049	0.013	0.013	0.952
		-0.288	0.022	0.022	0.860	0.129	0.020	0.020	0.929
		-0.279	0.028	0.028	0.815	0.275	0.027	0.027	0.847
		-0.402	0.031	0.032	0.747	0.382	0.030	0.030	0.760
		-0.479	0.032	0.032	0.687	0.484	0.033	0.032	0.676
$\hat{p}_1 = 0.988$, ASE= 0.014, SSE= 0.012, CP= 0.001									

Table 4: Simulation results on $\left\{\hat{F}_j(t_i), j = 1, 2\ i = 1, \cdots, 5\right\}$ under exponential(1) failure time distribution with two types (6 : 4 rate ratio) and known (p_1, p_2) for random monitoring time. The five t_i's are 0.105, 0.288, 0.693, 1.386 and 2.303, respectively.

		Cause 1			Cause 2		
(p_1, p_2)	n_i	$10 \times \lvert\text{Bias}\rvert$	ASE	SSE	$10 \times \lvert\text{Bias}\rvert$	ASE	SSE
(0.9,0.8)	50	0.039	0.021	0.059	0.026	0.011	0.044
		0.036	0.067	0.114	0.040	0.042	0.078
		0.018	0.121	0.134	0.023	0.089	0.120
		0.012	0.123	0.127	0.011	0.105	0.120
		0.011	0.110	0.130	0.029	0.111	0.128
	150	0.035	0.016	0.046	0.022	0.010	0.034
		0.014	0.066	0.071	0.024	0.041	0.053
		0.005	0.083	0.081	0.013	0.062	0.074
		0.011	0.074	0.078	0.006	0.069	0.074
		0.004	0.073	0.086	0.016	0.070	0.085
	250	0.027	0.013	0.033	0.019	0.010	0.020
		0.009	0.057	0.060	0.016	0.040	0.048
		0.001	0.070	0.073	0.008	0.053	0.052
		0.009	0.060	0.062	0.002	0.054	0.058
		0.004	0.058	0.063	0.003	0.058	0.066
(0.7,0.6)	50	0.040	0.021	0.077	0.032	0.015	0.044
		0.036	0.068	0.131	0.054	0.042	0.079
		0.020	0.125	0.152	0.029	0.094	0.114
		0.030	0.127	0.137	0.011	0.117	0.125
		0.021	0.124	0.135	0.029	0.119	0.131
	150	0.036	0.020	0.054	0.026	0.011	0.035
		0.026	0.066	0.083	0.025	0.041	0.066
		0.006	0.084	0.094	0.017	0.067	0.075
		0.017	0.080	0.086	0.007	0.074	0.076
		0.019	0.078	0.089	0.009	0.065	0.072
	250	0.029	0.020	0.046	0.024	0.010	0.030
		0.024	0.058	0.062	0.021	0.040	0.055
		0.010	0.070	0.075	0.009	0.056	0.057
		0.015	0.067	0.068	0.002	0.058	0.059
		0.009	0.065	0.072	0.004	0.063	0.069

Table 5: Simulation results on $\left\{\hat{F}_j(t_i), j=1,2,\ i=1,\cdots,5\right\}$ and $\hat{p}_1$ under exponential(1) failure time distribution with the two types (6 : 4 rate ratio) and $p_2 = c\ p_1$ (known c=0.8) for random monitoring time. The five t_i's are 0.105, 0.288, 0.693, 1.386 and 2.303, respectively.

		Cause 1			Cause 2		
Value of c in the analysis	n_i	Bias	ASE	SSE	Bias×10	ASE	SSE
		0.043	0.025	0.054	0.031	0.016	0.033
		0.054	0.063	0.116	0.046	0.046	0.073
0.8	50	0.016	0.121	0.114	0.021	0.096	0.106
		0.017	0.120	0.137	0.021	0.114	0.119
		0.015	0.117	0.132	0.034	0.115	0.120
Bias($\hat{p}_1$) = 0.057, ASE= 0.057, SSE= 0.049							
		0.043	0.013	0.058	0.030	0.009	0.022
		0.060	0.064	0.117	0.044	0.026	0.065
1	50	0.054	0.122	0.141	0.020	0.070	0.104
		0.040	0.120	0.146	0.020	0.104	0.110
		0.054	0.118	0.156	0.032	0.112	0.099
$\hat{p}_1$ = 0.876, ASE= 0.055, SSE=0.053							
		0.030	0.020	0.047	0.027	0.008	0.020
		0.026	0.060	0.079	0.025	0.038	0.055
0.8	150	0.009	0.076	0.081	0.011	0.066	0.067
		0.015	0.078	0.071	0.013	0.068	0.071
		0.014	0.075	0.072	0.031	0.074	0.070
Bias($\hat{p}_1$) = 0.048, ASE= 0.035, SSE= 0.034							
		0.027	0.020	0.052	0.032	0.005	0.024
		0.062	0.066	0.081	0.037	0.039	0.054
1	150	0.023	0.081	0.083	0.044	0.056	0.067
		0.045	0.076	0.090	0.046	0.062	0.061
		0.063	0.074	0.082	0.049	0.069	0.067
$\hat{p}_1$ = 0.879, ASE= 0.031, SSE= 0.029							
		0.023	0.019	0.045	0.015	0.007	0.011
		0.016	0.056	0.056	0.022	0.037	0.048
0.8	250	0.008	0.063	0.063	0.002	0.055	0.056
		0.002	0.061	0.068	0.006	0.060	0.056
		0.003	0.062	0.063	0.018	0.059	0.057
Bias($\hat{p}_1$) = 0.042, ASE= 0.027, SSE= 0.026							
		0.029	0.020	0.042	0.027	0.004	0.011
		0.001	0.057	0.069	0.030	0.036	0.047
1	250	0.030	0.070	0.078	0.033	0.048	0.055
		0.053	0.065	0.064	0.036	0.050	0.054
		0.064	0.062	0.066	0.044	0.055	0.051
$\hat{p}_1$ = 0.873, ASE= 0.025, SSE=0.025							

Table 6: Simulation results on $\left\{\hat{F}_j(t_i), j=1,2,\ i=1,\cdots,5\right\}$ and $\hat{p}_1$ under exponential(1) failure time distribution with the two types (6 : 4 rate ratio) and $p_2 = c\ p_1$ (known c=1) for random monitoring time. The five t_i's are 0.105, 0.288, 0.693, 1.386 and 2.303, respectively.

		Cause 1			Cause 2		
Value of c in the analysis	n_i	Bias	ASE	SSE	Bias×10	ASE	SSE
1	50	0.040	0.021	0.061	0.035	0.029	0.034
		0.044	0.067	0.114	0.050	0.043	0.078
		0.025	0.114	0.127	0.027	0.089	0.116
		0.019	0.110	0.121	0.016	0.112	0.106
		0.023	0.108	0.122	0.054	0.111	0.126
Bias($\hat{p}_1$) = 0.030, ASE= 0.039, SSE= 0.040							
0.8	50	0.041	0.014	0.051	0.038	0.043	0.042
		0.057	0.058	0.105	0.054	0.055	0.083
		0.043	0.112	0.117	0.018	0.094	0.117
		0.048	0.105	0.117	0.036	0.116	0.107
		0.029	0.108	0.105	0.066	0.117	0.128
$\hat{p}_1$ = 0.975, ASE= 0.050, SSE=0.031							
1	150	0.037	0.016	0.044	0.022	0.013	0.033
		0.032	0.057	0.084	0.027	0.042	0.055
		0.007	0.077	0.083	0.016	0.062	0.064
		0.009	0.075	0.089	0.007	0.068	0.068
		0.017	0.075	0.087	0.015	0.070	0.073
Bias($\hat{p}_1$) = 0.029, ASE= 0.023, SSE= 0.022							
0.8	150	0.044	0.014	0.036	0.022	0.013	0.037
		0.037	0.064	0.068	0.026	0.042	0.065
		0.013	0.076	0.083	0.018	0.070	0.068
		0.079	0.071	0.075	0.032	0.069	0.072
		0.037	0.073	0.078	0.053	0.071	0.073
$\hat{p}_1$ = 0.990, ASE= 0.030, SSE= 0.028							
1	250	0.027	0.013	0.043	0.021	0.012	0.020
		0.002	0.056	0.070	0.017	0.034	0.049
		0.001	0.065	0.057	0.006	0.049	0.063
		0.001	0.064	0.070	0.002	0.054	0.057
		0.006	0.062	0.063	0.009	0.058	0.058
Bias($\hat{p}_1$) = 0.024, ASE= 0.018, SSE= 0.018							
0.8	250	0.029	0.022	0.046	0.021	0.013	0.029
		0.018	0.053	0.063	0.018	0.040	0.054
		0.028	0.060	0.052	0.026	0.055	0.063
		0.033	0.058	0.058	0.045	0.056	0.059
		0.056	0.058	0.060	0.066	0.058	0.067
$\hat{p}_1$ = 0.995, ASE= 0.018, SSE=0.019							

6 Data Analysis

In this section, the proposed methods are illustrated using two real data sets. First, we consider the menopause data (See Section 6.1 of Koley and Dewanji (2018b)) and induce missing type at random to study the effect of different extent of missingness. Next, we consider a data set on hearing loss collected from Ali Yavar Jung National Institute of Speech and Hearing Disabilities, Eastern Regional Center (See Banik et al. (2018)). In this data, there are some observations with missing hearing type.

6.1 Menopause Data

This data has all the failure type information available. We induce some missing type randomly using different choices of the masking probabilities $(p_1, p_2) = (0.8, 0.8)$ and (0.9,0.8). At any monitoring time τ_i, the observed set of possible causes g for each failure is imputed using the binomial distribution as discussed in Section 6.1 of Koley and Dewanji (2018a). First we discard the missing observations and analyze the data based on only the complete part of the data set. Then, the full data including the imputed missing observations is analysed using the methods of Section 3. This procedure is repeated 50 times and the average of the estimates and the corresponding standard errors are noted. The results are presented in Tables 7 and 8 for $(p_1, p_2) = (0.8, 0.8)$ and (0.9,0.8), respectively, with the standard errors in parentheses. These results may be compared with those of the original data analysis without imputation (See Table 7 of Koley and Dewanji (2018b)). When the imputed missing observations are ignored, the estimates seem to be biased, as expected. On the other hand, the estimates are similar to the original data analysis when the imputed missing observations are considered for analysis, specially when the values of p_1 and p_2 used in the analysis coincide with the ones used for imputation. Also, the standard errors seem to be little smaller when $(p_1, p_2) = (0.9, 0.8)$ than those for $(p_1, p_2) = (0.8, 0.8)$, as expected, indicating some loss of efficiency due to higher probability of missingness.

6.2 Hearing Loss Data

Next, let us consider the data on hearing loss. We split the range of monitoring times into five sub-intervals as $(0, 10], (10, 30], (30, 50], (50, 70]$ and $(70, 90]$ for the ease of illustration. To illustrate the method discussed in Section 3, the five fixed monitoring time points are taken to be 10,30,50,70 and 90, the right end times of the five intervals. We estimate the two sub-distribution functions at these monitoring time points and the results are presented in Table 9. The results seem to be mildly sensitive to the choice of p_1 and p_2 when these are assumed known and also to the choice of c when it is assumed known. Considering the

AIC values, the assumption in the top panel with $(p_1, p_2) = (0.9, 0.8)$ seems to give the best fit to the data with the five fixed monitoring time points out of the four choices.

For the illustration of the method discussed in Section 4, we construct the disjoint intervals C_{jk}, for $k = 1, \cdots, m_j$ and $j = 1, 2$. The values of m_1, m_2 are obtained as 93 and 31, respectively. After construction of these intervals, the non-parametric maximum likelihood estimates of the sub-distribution functions are computed at the time points $10, 30, 50, 70$ and 90. The corresponding standard errors of the estimates are also obtained by using the bootstrap method. We consider the following cases one by one and analyze the data. First, we take $(p_1, p_2) = (0.9, 0.8)$ and $(0.7, 0.6)$ and both are assumed known in the analysis. Then, we consider $p_2 = cp_1$ with $c = 0.8, 1$ and only c is assumed known in the analyses. All the results are presented in Table 10. As before, the results seem to be mildly sensitive to the choice of p_1 and p_2 values in the first case and to the choice of c value in the second case.

Table 7: Analysis of Menopause Data with imputed missing types and $(p_1, p_2) = (0.8, 0.8)$ using the method of Section 3. The standard errors are given in parentheses.

	Type of Analysis									
	Ignoring Missing data		$p_1 = 0.9, p_2 = 0.8$ known		$p_1 = 0.8, p_2 = 0.8$ known		$p_2 = cp_1, c = 0.9$ known		$p_2 = cp_1, c = 1$ known	
τ	$\hat{F}_1$	$\hat{F}_2$	$\hat{F}_1$	$\hat{F}_2$	$\hat{F}_1$	$\hat{F}_2$	$\hat{F}_1$	$\hat{F}_2$	$\hat{F}_1$	$\hat{F}_2$
27.5	0.009	0	0.011	0	0.011	0	0.011	0	0.011	0
	(0.005)	(0.051)	(0.006)	(0.031)	(0.006)	(0.031)	(0.006)	(0.030)	(0.006)	(0.031)
32.5	0.048	0	0.058	0	0.058	0	0.058	0	0.058	0
	(0.012)	(0.053)	(0.014)	(0.030)	(0.014)	(0.030)	(0.014)	(0.021)	(0.014)	(0.030)
35.5	0.057	0	0.068	0	0.068	0	0.068	0	0.068	0
	(0.024)	(0.107)	(0.028)	(0.064)	(0.028)	(0.065)	(0.028)	(0.048)	(0.028)	(0.065)
36.5	0.057	0	0.068	0	0.068	0	0.068	0	0.068	0
	(0.027)	(0.108)	(0.032)	(0.087)	(0.032)	(0.087)	(0.032)	(0.077)	(0.032)	(0.087)
37.5	0.067	0.008	0.085	0.011	0.085	0.011	0.086	0.011	0.086	0.011
	(0.043)	(0.061)	(0.042)	(0.042)	(0.042)	(0.040)	(0.043)	(0.031)	(0.042)	(0.040)
38.5	0.090	0.014	0.108	0.024	0.113	0.019	0.110	0.021	0.113	0.019
	(0.030)	(0.012)	(0.033)	(0.018)	(0.033)	(0.015)	(0.034)	(0.016)	(0.033)	(0.015)
39.5	0.090	0.015	0.108	0.024	0.113	0.019	0.110	0.022	0.113	0.019
	(0.030)	(0.021)	(0.032)	(0.024)	(0.033)	(0.021)	(0.033)	(0.021)	(0.033)	(0.021)
40.5	0.092	0.015	0.108	0.024	0.113	0.019	0.110	0.022	0.113	0.019
	(0.035)	(0.035)	(0.038)	(0.038)	(0.039)	(0.034)	(0.039)	(0.032)	(0.039)	(0.034)
41.5	0.094	0.016	0.109	0.025	0.114	0.020	0.111	0.022	0.113	0.020
	(0.038)	(0.038)	(0.042)	(0.037)	(0.043)	(0.033)	(0.043)	(0.031)	(0.043)	(0.033)
42.5	0.141	0.042	0.159	0.059	0.167	0.051	0.162	0.056	0.167	0.051
	(0.037)	(0.023)	(0.039)	(0.031)	(0.040)	(0.027)	(0.040)	(0.028)	(0.040)	(0.027)
43.5	0.141	0.050	0.159	0.065	0.167	0.051	0.162	0.062	0.167	0.057
	(0.043)	(0.025)	(0.045)	(0.031)	(0.047)	(0.028)	(0.047)	(0.029)	(0.047)	(0.028)
44.5	0.144	0.050	0.159	0.065	0.167	0.051	0.162	0.062	0.167	0.057
	(0.046)	(0.029)	(0.048)	(0.037)	(0.050)	(0.033)	(0.049)	(0.034)	(0.050)	(0.033)
45.5	0.174	0.101	0.188	0.134	0.203	0.119	0.194	0.128	0.203	0.119
	(0.039)	(0.031)	(0.041)	(0.038)	(0.043)	(0.035)	(0.042)	(0.036)	(0.043)	(0.035)
46.5	0.188	0.119	0.202	0.152	0.219	0.137	0.208	0.146	0.219	0.137
	(0.046)	(0.039)	(0.048)	(0.047)	(0.050)	(0.043)	(0.051)	(0.045)	(0.050)	(0.043)
47.5	0.207	0.182	0.214	0.233	0.234	0.213	0.221	0.226	0.234	0.213
	(0.049)	(0.047)	(0.050)	(0.054)	(0.052)	(0.051)	(0.051)	(0.052)	(0.052)	(0.051)
48.5	0.210	0.203	0.216	0.246	0.237	0.227	0.222	0.239	0.237	0.227
	(0.047)	(0.047)	(0.048)	(0.053)	(0.051)	(0.050)	(0.050)	(0.051)	(0.051)	(0.050)
49.5	0.222	0.278	0.220	0.345	0.242	0.321	0.227	0.337	0.242	0.321
	(0.051)	(0.060)	(0.051)	(0.065)	(0.054)	(0.063)	(0.052)	(0.063)	(0.054)	(0.063)
50.5	0.222	0.414	0.220	0.469	0.243	0.447	0.227	0.462	0.243	0.447
	(0.052)	(0.063)	(0.051)	(0.063)	(0.055)	(0.063)	(0.053)	(0.062)	(0.055)	(0.063)
51.5	0.222	0.486	0.220	0.529	0.243	0.508	0.227	0.523	0.243	0.508
	(0.064)	(0.068)	(0.064)	(0.068)	(0.067)	(0.069)	(0.065)	(0.068)	(0.067)	(0.069)
52.5	0.267	0.550	0.252	0.595	0.277	0.570	0.260	0.588	0.277	0.570
	(0.064)	(0.074)	(0.062)	(0.071)	(0.066)	(0.073)	(0.063)	(0.071)	(0.066)	(0.073)
53.5	0.268	0.572	0.253	0.615	0.278	0.589	0.261	0.607	0.278	0.589
	(0.061)	(0.066)	(0.059)	(0.065)	(0.063)	(0.066)	(0.060)	(0.075)	(0.063)	(0.066)
54.5	0.314	0.583	0.287	0.630	0.317	0.601	0.296	0.621	0.317	0.601
	(0.071)	(0.077)	(0.067)	(0.073)	(0.071)	(0.076)	(0.068)	(0.073)	(0.071)	(0.076)

Continued on next page

Table 7 – continued from previous page

	Type of Analysis									
	Ignoring Missing data		$p_1 = 0.9, p_2 = 0.8$ known		$p_1 = 0.8, p_2 = 0.8$ known		$p_2 = cp_1, c = 0.9$ known		$p_2 = cp_1, c = 1$ known	
τ	$\hat{F}_1$	$\hat{F}_2$	$\hat{F}_1$	$\hat{F}_2$	$\hat{F}_1$	$\hat{F}_2$	$\hat{F}_1$	$\hat{F}_2$	$\hat{F}_1$	$\hat{F}_2$
55.5	0.314	0.651	0.287	0.687	0.316	0.657	0.296	0.678	0.317	0.657
	(0.071)	(0.076)	(0.067)	(0.070)	(0.071)	(0.074)	(0.068)	(0.071)	(0.071)	(0.074)
56.5	0.314	0.663	0.287	0.694	0.314	0.664	0.296	0.685	0.314	0.665
	(0.076)	(0.077)	(0.073)	(0.073)	(0.076)	(0.077)	(0.074)	(0.074)	(0.076)	(0.077)
57.5	0.314	0.665	0.287	0.696	0.314	0.666	0.296	0.686	0.314	0.666
	(0.077)	(0.076)	(0.072)	(0.073)	(0.076)	(0.076)	(0.073)	(0.074)	(0.076)	(0.076)
58.5	0.318	0.682	0.289	0.717	0.314	0.682	0.299	0.701	0.314	0.680
	(0.095)	(0.134)	(0.084)	(0.129)	(0.093)	(0.125)	(0.088)	(0.126)	(0.093)	(0.125)
							$\hat{p}_1$	0.846	$\hat{p}_1$	0.8
								(0.016)		(0.015)

Table 8: Analysis of Menopause Data with imputed missing types and $(p_1, p_2) = (0.9, 0.8)$ using the method of Section 3. The standard errors are given in parentheses.

	Type of Analysis									
	Ignoring Missing data		$p_1 = 0.9, p_2 = 0.8$ known		$p_1 = 0.8, p_2 = 0.8$ known		$p_2 = cp_1, c = 0.9$ known		$p_2 = cp_1, c = 1$ known	
τ	$\hat{F}_1$	$\hat{F}_2$	$\hat{F}_1$	$\hat{F}_2$	$\hat{F}_1$	$\hat{F}_2$	$\hat{F}_1$	$\hat{F}_2$	$\hat{F}_1$	$\hat{F}_2$
27.5	0.010	0	0.011	0	0.011	0	0.011	0	0.011	0
	(0.005)	(0.051)	(0.006)	(0.036)	(0.006)	(0.036)	(0.006)	(0.036)	(0.006)	(0.038)
32.5	0.052	0	0.059	0	0.059	0	0.059	0	0.059	0
	(0.012)	(0.053)	(0.013)	(0.023)	(0.013)	(0.023)	(0.013)	(0.024)	(0.014)	(0.035)
35.5	0.062	0	0.068	0	0.068	0	0.068	0	0.068	0
	(0.025)	(0.107)	(0.027)	(0.058)	(0.027)	(0.058)	(0.027)	(0.060)	(0.027)	(0.076)
36.5	0.062	0	0.068	0	0.068	0	0.068	0	0.068	0
	(0.029)	(0.108)	(0.032)	(0.084)	(0.032)	(0.084)	(0.032)	(0.085)	(0.032)	(0.094)
37.5	0.079	0.014	0.083	0.015	0.083	0.015	0.083	0.015	0.083	0.015
	(0.035)	(0.028)	(0.036)	(0.021)	(0.036)	(0.020)	(0.036)	(0.021)	(0.036)	(0.024)
38.5	0.104	0.016	0.114	0.018	0.112	0.012	0.114	0.018	0.112	0.012
	(0.031)	(0.012)	(0.033)	(0.015)	(0.033)	(0.013)	(0.033)	(0.015)	(0.033)	(0.014)
39.5	0.104	0.017	0.114	0.018	0.112	0.012	0.114	0.018	0.112	0.012
	(0.031)	(0.012)	(0.033)	(0.014)	(0.033)	(0.013)	(0.033)	(0.014)	(0.033)	(0.013)
40.5	0.104	0.017	0.114	0.018	0.112	0.012	0.114	0.018	0.112	0.012
	(0.038)	(0.039)	(0.040)	(0.027)	(0.040)	(0.025)	(0.040)	(0.027)	(0.040)	(0.032)
41.5	0.106	0.017	0.113	0.018	0.115	0.020	0.113	0.015	0.115	0.020
	(0.039)	(0.029)	(0.041)	(0.025)	(0.041)	(0.024)	(0.041)	(0.025)	(0.041)	(0.027)
42.5	0.156	0.032	0.171	0.045	0.176	0.038	0.172	0.044	0.176	0.038
	(0.038)	(0.020)	(0.041)	(0.026)	(0.041)	(0.023)	(0.041)	(0.026)	(0.040)	(0.024)
43.5	0.156	0.044	0.171	0.054	0.176	0.049	0.172	0.053	0.176	0.049
	(0.045)	(0.023)	(0.047)	(0.026)	(0.048)	(0.024)	(0.047)	(0.026)	(0.047)	(0.025)
44.5	0.156	0.045	0.171	0.054	0.176	0.049	0.172	0.054	0.176	0.049
	(0.047)	(0.028)	(0.050)	(0.032)	(0.051)	(0.029)	(0.050)	(0.032)	(0.050)	(0.030)
45.5	0.194	0.099	0.206	0.116	0.214	0.108	0.207	0.115	0.214	0.108
	(0.040)	(0.031)	(0.042)	(0.034)	(0.043)	(0.032)	(0.042)	(0.034)	(0.042)	(0.033)
46.5	0.202	0.120	0.214	0.143	0.224	0.133	0.215	0.142	0.224	0.133
	(0.047)	(0.038)	(0.049)	(0.043)	(0.051)	(0.040)	(0.052)	(0.044)	(0.050)	(0.041)
47.5	0.226	0.181	0.234	0.215	0.247	0.201	0.235	0.213	0.247	0.201
	(0.050)	(0.045)	(0.051)	(0.050)	(0.053)	(0.047)	(0.051)	(0.050)	(0.052)	(0.049)
48.5	0.232	0.191	0.237	0.225	0.253	0.210	0.239	0.223	0.253	0.210
	(0.049)	(0.046)	(0.050)	(0.050)	(0.052)	(0.048)	(0.050)	(0.050)	(0.051)	(0.049)
49.5	0.241	0.277	0.240	0.319	0.258	0.302	0.242	0.317	0.258	0.302
	(0.051)	(0.059)	(0.051)	(0.061)	(0.054)	(0.060)	(0.051)	(0.061)	(0.053)	(0.061)
50.5	0.241	0.403	0.240	0.447	0.258	0.429	0.242	0.446	0.258	0.429
	(0.052)	(0.062)	(0.052)	(0.061)	(0.055)	(0.061)	(0.052)	(0.062)	(0.054)	(0.062)
51.5	0.241	0.480	0.240	0.513	0.258	0.496	0.242	0.511	0.258	0.496
	(0.064)	(0.066)	(0.064)	(0.066)	(0.066)	(0.066)	(0.065)	(0.066)	(0.066)	(0.067)
52.5	0.299	0.526	0.283	0.566	0.307	0.542	0.285	0.564	0.307	0.542
	(0.066)	(0.073)	(0.064)	(0.070)	(0.067)	(0.071)	(0.064)	(0.071)	(0.067)	(0.072)
53.5	0.299	0.543	0.285	0.583	0.309	0.558	0.287	0.580	0.309	0.558
	(0.062)	(0.067)	(0.060)	(0.064)	(0.063)	(0.065)	(0.060)	(0.064)	(0.063)	(0.066)
54.5	0.334	0.563	0.310	0.604	0.336	0.578	0.313	0.602	0.336	0.578
	(0.070)	(0.076)	(0.067)	(0.072)	(0.070)	(0.074)	(0.067)	(0.072)	(0.071)	(0.074)

Continued on next page

Table 8 – continued from previous page

	Type of Analysis									
	Ignoring Missing data		$p_1 = 0.9, p_2 = 0.8$ known		$p_1 = 0.8, p_2 = 0.8$ known		$p_2 = cp_1, c = 0.9$ known		$p_2 = cp_1, c = 1$ known	
τ	$\hat{F}_1$	$\hat{F}_2$	$\hat{F}_1$	$\hat{F}_2$	$\hat{F}_1$	$\hat{F}_2$	$\hat{F}_1$	$\hat{F}_2$	$\hat{F}_1$	$\hat{F}_2$
55.5	0.334	0.634	0.310	0.664	0.336	0.638	0.313	0.661	0.336	0.638
	(0.071)	(0.074)	(0.067)	(0.070)	(0.071)	(0.073)	(0.067)	(0.070)	(0.071)	(0.073)
56.5	0.334	0.643	0.310	0.671	0.336	0.645	0.313	0.668	0.336	0.645
	(0.076)	(0.077)	(0.073)	(0.074)	(0.076)	(0.076)	(0.074)	(0.074)	(0.077)	(0.077)
57.5	0.334	0.646	0.310	0.672	0.336	0.646	0.313	0.670	0.336	0.646
	(0.75)	(0.75)	(0.072)	(0.072)	(0.074)	(0.074)	(0.072)	(0.073)	(0.075)	(0.075)
58.5	0.341	0.659	0.314	0.686	0.342	0.658	0.317	0.683	0.342	0.658
	(0.099)	(0.133)	(0.090)	(0.124)	(0.098)	(0.119)	(0.091)	(0.124)	(0.095)	(0.124)
							$\hat{p}_1$	0.901	$\hat{p}_1$	0.858
								(0.015)		(0.013)

Table 9: Analysis of Hearing loss data with fixed monitoring times using the method of Section 3. The standard errors are given in parentheses.

Type of Analysis		MLE with standard error in parantheses						
		Monitoring time (τ)						
		10	30	50	70	90	p_1	AIC
$p_1 = 0.9$, $p_2 = 0.8$ known	F_1	0.451	0.780	0.780	0.780	0.858	-	1441.392
		(0.025)	(0.033)	(0.043)	(0.053)	(0.183)		
	F_2	0.018	0.035	0.142	0.142	0.142	-	
		(0.007)	(0.015)	(0.037)	(0.047)	(0.300)		
$p_1 = 0.7, p_2 = 0.6$ known	F_1	0.452	0.792	0.792	0.792	0.873	-	1556.82
		(0.025)	(0.032)	(0.042)	(0.050)	(0.186)		
	F_2	0.018	0.033	0.127	0.127	0.127		
		(0.007)	(0.014)	(0.033)	(0.044)	(0.415)		
$p_2 = cp_1$, $c = 0.8$ known	F_1	0.450	0.763	0.763	0.763	0.837	0.913	1561.274
		(0.025)	(0.035)	(0.053)	(0.054)	(0.663)	(0.054)	
	F_2	0.019	0.039	0.163	0.163	0.163		
		(0.007)	(0.017)	(0.049)	(0.050)	(0.239)		
$p_2 = cp_1$, $c = 1$ known	F_1	0.452	0.798	0.798	0.798	0.881	0.886	1558.953
		(0.025)	(0.031)	(0.041)	(0.049)	(0.188)	(0.014)	
	F_2	0.018	0.031	0.119	0.119	0.119		
		(0.007)	(0.014)	(0.032)	(0.042)	(0.534)		

Table 10: Analysis of hearing loss data with random monitoring time using the method of Section 4. The standard errors are given in parentheses.

Type of Analysis		MLE with standard error in parantheses					
		Time points					
		10	30	50	70	90	p_1
$p_1 = 0.9, p_2 = 0.8$	F_1	0.693	0.789	0.789	0.807	0.861	-
known		(0.055)	(0.029)	(0.029)	(0.033)	(0.029)	-
	F_2	0.023	0.138	0.139	0.139	0.139	-
		(0.009)	(0.037)	(0.029)	(0.029)	(0.029)	-
$p_1 = 0.7, p_2 = 0.6$	F_1	0.693	0.805	0.805	0.825	0.879	-
known		(0.045)	(0.023)	(0.023)	(0.025)	(0.018)	-
	F_2	0.023	0.121	0.121	0.121	0.121	-
		(0.013)	(0.025)	(0.019)	(0.018)	(0.018)	-
$p_2 = cp_1$,	F_1	0.687	0.769	0.769	0.776	0.832	0.919
$c = 0.8$ known		(0.044)	(0.022)	(0.022)	(0.024)	(0.023)	(0.013)
	F_2	0.025	0.157	0.167	0.168	0.168	
		(0.011)	(0.031)	(0.022)	(0.023)	(0.023)	
$p_2 = cp_1$	F_1	0.689	0.813	0.813	0.833	0.887	0.892
$c = 1$, known		(0.041)	(0.024)	(0.024)	(0.028)	(0.019)	(0.012)
	F_2	0.023	0.113	0.113	0.113	0.113	
		(0.008)	(0.024)	(0.018)	(0.019)	(0.019)	

7 Concluding Remarks

For a given j, the sub-distribution function $F_j(\cdot)$ is right continuous, non-decreasing with $\sum_j F_j(\infty) \leq 1$. Thus, as remarked before, estimating sub-distribution functions is a constrained optimization problem. For fixed monitoring times in Section 3, a simple re-parametrization has been considered to make it an unconstrained problem so that the estimation procedure is simple and easy to deal with. The proposed estimation technique is quite flexible and can be applied to other design scenarios, for example, when the number of individuals observed at each monitoring time point is random. Assuming the monitoring time to be independent of the failure time random variable T, the observations on n individuals are then independent realizations $\{(x_i, g_i),\ i = 1, \cdots, n\}$ from the common density

$$p^*(g,\tau) = \begin{cases} p_1 F_1(\tau) P\left[X=\tau\right], & \text{if } g = \{1,\} X = \tau \\ p_2 F_2(\tau) P\left[X=\tau\right], & \text{if } g = \{2,\} X = \tau \\ \left[(1-p_1)F_1(\tau) + (1-p_2)F_2(\tau)\right] P\left[X=\tau\right], & \text{if } g = \{1,2,\} X = \tau \\ S(\tau) P\left[X=\tau\right], & \text{if } g = \{\phi,\} X = \tau, \end{cases}$$

with τ taking K values $\tau_1, \cdots, \tau_K$. This density at $\tau = \tau_i$ is proportional to the density $f_i^*(\cdot)$, as defined at the begining of Section 3. Hence, the estimation procedure in this case remains the same as the one discussed in Section 3. It is easy to observe that the asymptotic properties of the MLEs are also similar in both the cases.

In Section 4, we have developed method to obtain the non-parametric maximum likelihood estimates of the sub-distribution functions with random monitoring time. Both simulation studies and real life data analysis show that the results are mildly sensitive to the choice of p_1 and p_2. The choice of (p_1, p_2) remains a challenging task. One remedy is to consider the AIC values and choose the one with the least AIC value.

Note that, in the fixed monitoring time approach for data with random monitoring time, as in the case of the hearing loss data, method of Section 3 is based on a summarized version of the original data. In practice, for convenience, the data with random monitoring time may be stored as a grouped data and the mid-points of the group intervals are treated as fixed monitoring time. This may have been the situation with the menopause data. Therefore, the precision of the estimates may be somewhat reduced. We notice this phenomenon in our analysis of the hearing loss data with random monitoring time and fixed monitoring time approach as well. Comparing the estimates in Table 9 for fixed monitoring time approach and those in Table 10 for random monitoring time approach, it is seen that they are quite similar except at the first monitoring time/ time point, where the summarization has been rather heavy. Also, note that the standard errors for the random monitoring time approach are uniformly smaller than those in the fixed monitoring time approach, as expected.

References

Banik, P., Jaiswal, J., Koley, T., and Chatterjee., I. (2018). *Explanatory analysis of tertiary audiological data.*, volume 1. Asia Pacific Journal of Research.

Basu, S. (2009). Inference about the masking probabilities in the competing risks model. *Communications in Statistics—Theory and Methods*, 38(16-17):2677–2690.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)*, pages 1–38.

Dewanji, A. and Sengupta, D. (2003). Estimation of competing risks with general missing pattern in failure types. *Biometrics*, 59(4):1063–1070.

Gentleman, R. and Geyer, C. J. (1994). Maximum likelihood for interval censored data: Consistency and computation. *Biometrika*, 81(3):618–623.

Groeneboom, P., Maathuis, M. H., and Wellner, J. A. (2008a). Current status data with competing risks: Limiting distribution of the mle. *Annals of statistics*, 36(3):1064.

Groeneboom, P., Maathuis, M. H., Wellner, J. A., et al. (2008b). Current status data with competing risks: consistency and rates of convergence of the mle. *The Annals of Statistics*, 36(3):1031–1063.

Hudgens, M. G., Satten, G. A., and Longini Jr, I. M. (2001). Nonparametric maximum likelihood estimation for competing risks survival data subject to interval censoring and truncation. *Biometrics*, 57(1):74–80.

Jewell, N. P. and Kalbfleisch, J. D. (2004). Maximum likelihood estimation of ordered multinomial parameters. *Biostatistics*, 5(2):291–306.

Jewell, N. P., Van Der Laan, M., and Henneman, T. (2003). Nonparametric estimation from current status data with competing risks. *Biometrika*, 90(1):183–197.

Kalbfleisch, J. D. and Prentice, R. L. (2002). *The statistical analysis of failure time data, second edition.* John Wiley & Sons.

Koley, T. and Dewanji, A. (2018a). Current status data with two competing risks and missing failure types. https://www.isical.ac.in/~asu/TR/TechRepASU201802.pdf.

Koley, T. and Dewanji, A. (2018b). Revisiting non-parametric maximum likelihood estimation of current status data with competing risks. *Sankhya B*.

Lehmann, E. C. and Casella, G. (1998). G.(1998). theory of point estimation.

Maathuis, M. H. (2006). *Nonparametric estimation for current status data with competing risks*. PhD thesis, University of Washington.

Turnbull, B. W. (1976). The empirical distribution function with arbitrarily grouped, censored and truncated data. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 290–295.

Wu, C. J. (1983). On the convergence properties of the em algorithm. *The Annals of statistics*, pages 95–103.