Statistical Modeling and Analysis of Competing Risk Lifetime Data

THESIS SUBMITTED FOR THE DEGREE OF

Doctor of Philosophy

in

Statistics

By

Chandra Prakash Yadav

Under the Supervision of

Dr. Mahaveer Singh Panwar

DEPARTMENT OF STATISTICS
INSTITUTE OF SCIENCE
BANARAS HINDU UNIVERSITY
VARANASI - 221005
INDIA

Enrolment No.: 401556 **March 2022**

Dedicated to

"My Parents and Elder Brother"

UNDERTAKING FROM CANDIDATE

I, hereby declare that I have completed the research work for the full time period prescribed under the clause VIII.1 of the Ph.D. ordinances of the Banaras Hindu University, Varanasi, India under the supervision of **Dr. Mahaveer Singh Panwar** and also I declare that the research work embodied in this thesis entitled "**Statistical Modeling and Analysis of Competing Risk Lifetime Data**" is my own research work.

I avail myself to responsibility such as an act will be taken on behalf of me, mistakes, errors of facts and misinterpretations are of course entirely my own.

Date:
Place: Varanasi

Signature of the Candidate
(**Chandra Prakash Yadav**)

Annexure-E
[Clause XIII. 2(b)(iii)]

CANDIDATE'S DECLARATION

I, **Chandra Prakash Yadav**, certify that the work embodied in this Ph.D. thesis is my own bonafide work carried out by me under the supervision of **Dr. Mahaveer Singh Panwar**, for a period of September 2017 to March 2022 at Banaras Hindu University. The matter embodied in this Ph.D. thesis has not been submitted for the award of any other degree/diploma.

I declare that I have faithfully acknowledged, given credit to and referred to the research workers wherever their works have been cited in the text and the body of the thesis. I further certify that I have not willfully lifted up some other's work, paragraph, text, data, results, etc. reported in the journals, books, magazines, reports, dissertations, theses, etc., or available at websites and included them in this Ph.D. thesis and cited as my own work.

Date:
Place: Varanasi

Signature of the Candidate
(**Chandra Prakash Yadav**)

.....

CERTIFICATE FROM THE SUPERVISOR

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Prof. S. K. Singh
(Head of the Department)

Dr. Mahaveer Singh Panwar
(Supervisor)

Annexure-F
[Clause XIII. 1(c) and XIII. 2(b) (iv)]

COURSE WORK COMPLETION CERTIFICATE

This is to certify that **Mr. Chandra Prakash Yadav**, a bonafide research scholar of this Department, has successfully completed the Ph.D. course work requirement, which is a part of his Ph.D. programme.

Date:
Place: Varanasi

Prof. S. K. Singh
(Head of the Department)

Annexure-F
[Clause XIII. 1(c) & XIII. 2(b) (iv)]

PRE-SUBMISSION SEMINAR COMPLETION CERTIFICATE

This is to certify that **Mr. Chandra Prakash Yadav**, a bonafide research scholar of this department, has successfully completed the pre-submission seminar requirement on **01-02-2022**, which is a part of his Ph.D. programme, on his thesis entitled **"Statistical Modeling and Analysis of Competing Risk Lifetime Data"**.

Date:
Place: Varanasi

Prof. S. K. Singh
(Head of the Department)

Annexure-G
[Clause XIII. 2(b) (v)]

COPYRIGHT TRANSFER CERTIFICATE

Title of the Thesis: **Statistical Modeling and Analysis of Competing Risk Lifetime Data**

Candidate's Name: **Chandra Prakash Yadav**

Copyright Transfer

The undersigned hereby assigns to the Banaras Hindu University all rights under copyright that may exist in and for the above thesis submitted for the award of the Ph.D. degree.

Signature of the Candidate

Note: However, the author may reproduce or authorize others to reproduce material extracted verbatim from the thesis or derivative of the thesis for author's personal use provided that the source and the University's copyright notice are indicated.

Acknowledgment

First of all, I bow my head with reverence to ***Lord Shiva*** *for his graciousness on oldest continually-inhabited places on earth and* ***Pt. Madan Mohan Malviya Ji****, the esteemed founder of such a beautiful university in Banaras. I am very lucky that I got an opportunity to pursue research in Statistics at Banaras Hindu University (BHU).*

The thesis in its current form, would not have been possible without the guidance and support of my supervisor ***Dr. Mahaveer Singh Panwar****. I feel immense pleasure in expressing my deep sense of respect and gratitude to him for his persistent guidance, affection and inspiration. His guidance and his endless encouragement throughout the span of this study will be cherished forever. His caring nature and amicable attitude always gave me the courage to overcome the hurdles throughout the journey. I am indebted to him more than he knows and I will be grateful to him throughout my life.*

I am grateful to ***Prof. S. K. Singh****, Head, Department of Statistics and* ***Prof. Brij Behari Khare****, Ex-Head, Department of Statistics, Institute of Science, Banaras Hindu University, for their keen interest and providing me all necessary departmental facilities during my research work. My special thanks to my Research Progress Committee,* ***Prof. Sanjeev Kumar****,* ***Prof. Rajesh Singh****, Department of Statistics, Institute of Science, Banaras Hindu University and* ***Dr. Jitendra Singh****, Department of Mathematics, Institute of Science, Banaras Hindu University, for their kind support and constructive suggestions during the research process.*

I take this opportunity to express my heartiest thanks to all of my respected teachers of the Department of Statistics, Banaras Hindu University for their generous help and kind support throughout my research period. I sincerely thank the non-teaching staff for their helping attitude at different stages of my research work.

I am thankful for ***Prof. Anup Dewanji****, Applied Statistics Unit, Indian Statistical Institute, Kolkata for the discussion in concept building and making doubts clear at the very beginning of my journey and* ***Prof. Debasis Sengupta****, Applied Statistics Unit, Indian Statistical Institute, Kolkata for providing me age at menarche data for research purposes. I wish to acknowledge my profound gratitude to* ***Dr. Himanshu Rai*** *for helping me to learn the basics of R programming and discussion at different stages of the formulation of the setup of the problem,* ***Ms. Rashmi Bhundel****, Assistant Professor, Department of Statistics, the University of Rajasthan for discussion and support in understanding the core concepts of Masking and Current Status data. I would like to give special thanks to my*

seniors ***Dr. Jitendra Kumar****, Directorate of Economics & Statistics, Planning Department, Delhi and* ***Dr. Sachin Chaudhary****, Assistant Professor, Department of Community Medicine, Government Medical College Kannauj, U.P.,* ***Dr. Amita Barnwal****, Assistant Professor, Department of Statistics, Patna Women's College for their support and motivation at the very beginning and later at various stages of the research journey. I would like to thank* ***Dr. Puneet Rakesh Gupta****, ICFAI University, Dehradun and* ***Dr. Abhinav Singh*** *for their motivation, guidance and support to start the research journey.*

I would like to express love and affection for my brother ***Mr. Vikas Barnwal*** *for his help, support and discussion during all stages of my research work. I feel pleasure to thank* ***Mr. Mridulesh Kumar Yadav*** *and* ***Mr. Vikas Baranwal*** *for their help, consistent support and discussion during various stages of this journey. I would like to express special thanks to* ***Mr. Tribhuvan Singh****,* ***Ms. Sonam Jaiswal****,* ***Mr. Pradeep Maurya****,* ***Dr. Pawan Chauhan*** *for their moral support and making the journey enjoyable. Moreover, I am thankful for my beloved juniors* ***Mr. Sanjiv Kumar*** *and* ***Mr. Avinash Kumar*** *Research Scholars, Department of Statistics, Banaras Hindu University for their multilevel support during the last stages of my research journey. Along with it, I am also thankful to all other research scholars of the Department of Statistics, Banaras Hindu University, for their wholehearted support and kind cooperation throughout the research period.*

My deepest thanks goes to my parents ***Smt. Susheela Devi & Shri. Ram Dev Yadav*** *for their unbounded love, sacrifice and moral support rendered throughout my study. They are always there for me in every step of my life. I am grateful to my loving elder brother* ***Mr. Ram Prakash Yadav****, sister-in-law* ***Mrs. Meera Devi*** *younger brother* ***Mr. Surya Prakash Yadav****, sisters* ***Mrs. Kamla****,* ***Mrs. Vimla*** *for pushing me up, motivating and cheering whenever I felt down. I am fortunate enough to have a nephew* ***Nitin*** *and three nieces who fulfilled lots of love and affection in my life.*

Last but not least, my special thanks to my close heart friends ***Mr. Abhishek Pandey****,* ***Mr. Vijay Gupta****,* ***Mr. Krishna Yadav****,* ***Mr. Dheeraj Verma*** *and* ***Mr. Brijesh Kasudhan*** *for their multilevel support from the beginning to the end of this academic endeavor. I feel it was impossible to imagine this path without your company.*

March, 2022 *Chandra Prakash Yadav*

Preface

Time to event data arises in many areas where an individual or a subject is followed over time. The main areas in which the time to event data arises can be seen commonly are public health, demography, epidemiology and economics. In cross-sectional studies, the individuals/subjects are monitored at a fixed or random time point to know the status of the event. The budget and time constraint are two important factors that are to be taken into account during these studies. These restrictions can generate censored or incomplete data. There are some well-known censoring schemes such as Type-I, Type-II, random, right, left and Interval which commonly arise in different situations. In the case of Interval censored data, individuals are visited at pre-specified monitoring points. But in the cross-sectional study, the individual is monitored only once at any particular time point. Further, if we are interested to know the status of events at the monitoring time, status quo or current status data is observed. These studies observe time to the event either in the form of left-censored or right-censored, which is considered to be a particular form of interval-censored data. Further in some studies, the individuals are asked to recall the event which had occurred in the past. These studies are known as recall-based studies in the literature. It may be the case that the individual may not able to report the time to event exactly and generated non-recall observations. The thesis is dedicated to modeling recall-based studies that arise in real-life situations. The information available on recall is utilized in the development of the mathematical model.

Chapter 1 contains the introduction, basic terms and concepts related to time to event data. A review of the literature on time to event studies is given. Competing risk data arises in the case when there is more than one risk acting simultaneously, and only one of them is responsible for the occurrence of an event. In such studies, time to event data with the associated cause of the event is observed. In some cases, the time to event is observed but the cause of occurrence of an event is not known exactly known as masked data. Computational techniques under classical as maximum likelihood principle, missing information principle for the calculation of point and interval estimates are discussed. Under

the Bayesian paradigm, choice of prior, and sample generation techniques as Markov Chain Monte Carlo (MCMC) are discussed.

In Chapter 2, we develop parametric models for the considered recall-based data. The chance of recalling an event for an individual depends on the elapsed time between monitoring time and time to event. Thus, the non-recall probabilities are modeled using this information. Time to event distribution is considered as Weibull distribution. In the classical framework, point estimates are obtained by using the Expectation-Maximization algorithm. The observed Fisher information matrix is constructed by using the missing information principle. In the Bayesian approach, point estimates are obtained under informative priors. The Highest Posterior Density intervals are calculated using Markov Chain Monte Carlo samples. To explore the choice of monitoring time patterns, some useful quantities are derived. The simulation study is done to access the performance of proposed estimators. The duration of breastfeeding for the recent child is estimated using the proposed methodology from NFHS-IV data.

In Chapter 3, we generalized the recall-based studies under competing risks setup. Here, time to event and non-recall probability are considered to be of exponential form. Also, we consider that more than one cause (factor) is simultaneously acting and one of them is responsible for the occurrence of the event of interest. An efficient approach for recall-based competing risk data is established through a competing risk setup. In the classical method, a nested Expectation-Maximization technique is worked out for the estimation purpose and for the observed Fisher information matrix, the missing information principle is utilized. In the Bayesian paradigm, point and interval estimates are obtained using the Gibbs sampling algorithm. An extensive simulation study is done to explore the different proportions of non-recall. Two patterns of monitoring time points are chosen to explore the results. The age at menarche is estimated considering first born and later born as competing causes for real data.

In Chapter 4, we deal with the recall-based problem under competing risk scenarios as in Chapter 3. To capture the flexibility of different nature of hazard function, Weibull distribution is assumed for latent time to event corresponding to different causes. Using

the frequentist approach, an efficient methodology based on the Expectation-Maximization algorithm is developed for point estimation. The observed Fisher information matrix is calculated by using the missing information principle. In the Bayesian framework, a three-stage nested Gibbs Sampling algorithm is proposed for sample generation from the conditional posteriors. For illustration purposes, an extensive simulation study is performed with varying sample sizes and with different proportions of non-recall and censored data. Finally, the age at menarche for the girls is estimated using proposed methodologies.

In most of the studies related to competing risks, the risks responsible for the occurrence of an individual are taken from the same family. But in real-life situations, it is not necessary and various risks which are responsible for the events may follow different distributions. In Chapter 5, we consider the time to event associated with different causes are not from the same families. Exponential and Rayleigh distributions are chosen for time to event associated with causes 1 and 2 respectively. The probability of non-recall is taken as exponentially distributed as a function of elapsed time between monitoring and time to event. Under the classical approach, for point estimation, we used an Expectation-Maximization approach. The missing information principle is utilized for the calculation of the observed Fisher information matrix. The study is extended to the Bayesian approaches by choosing suitable choices of the conjugate family of priors for exponential and Rayleigh distributions. An extensive simulation study is done to explore the impact of monitoring patterns and different proportions of non-recall.

Contents

List of Tables

List of Figures

List of Abbreviations

PDF:	Probability Density Function
CDF:	Cumulative Distribution Function
ML:	Maximum Likelihood
MSE:	Mean Square Error
AB:	Absolute Bias
AL:	Average Length
E-M:	Expectation Maximization
ACI:	Asymptotic Confidence Interval
BCI:	Bayesian Credible Interval
HPD:	Highest Posterior Density
M-H:	Metropolis-Hastings
MCMC:	Markov Chain Monte Carlo
CP:	Coverage Probability
SELF:	Squared Error Loss Function
TTT:	Total Time on Test
ETS:	Expected Time in Study
TTS:	Total Time in Study
ETS:	Expected Time in Study
ACF:	Auto Correlation Function
$\mathcal{E}$**:**	Exponential
$\mathcal{W}$**:**	Weibull
$\mathcal{U}$**:**	Uniform
$\mathcal{G}$**:**	Gamma
$\mathcal{R}$**:**	Rayleigh
$\mathcal{SQIG}$**:**	Square Root Inverse Gamma

Chapter 1

Introduction

As human beings, we are always concerned about lifetime. We draw various kinds of inferences on the basis of lifetime data and then utilize these results for future events. A lifetime experiment can be designed for different studies associated with time to event. In an experiment the lifetime can be defined as a failure time, survival time, diagnose time, exposure time, experience time, etc. of an event of interest. So the lifetime data occurs in a variety of fields such as biomedical research, medical health, demography, epidemiology, reliability engineering, survival analysis, biology, economics and industry. In lifetime data, the variable of interest is time to event of failure/death of the system/individual since follow-up starts. For example, in manufacturing industries, the mechanical items are put on the test and their failure times are the event of interest. Here in this case failure times of items can be considered lifetime data. In another example, demographers are interested in determining the duration of marriage in a region or a country. The marriages may end due to divorce or death etc. Here the lifetime data includes the duration of the marriage. Also in economics, the researchers are interested in the duration of unemployment of youths. Thus this duration time generates lifetime data. Similarly, in clinical studies, the time to treatment failure has an important role in determining the effectiveness of treatment. These are some basic examples of lifetime data in various fields.

The event of interest may be the death of kidney transplant patients (Klein and Moeschberger (2006)). It can also be relapse from remission of a disease, failure of a component, or breakdown of a system as mentioned in (Li et al. (2007)). The event of interest may also be onset

of menarche in young girls (Bergsten-Brucefors (1976), Chumlea et al. (2003), Salehabadi et al. (2015)). The onset of menopause in women is also an event of interest discussed in (MacMahon et al. (1966), Krailo and Pike (1983)), marriage or age of the first child (Allison (1982)), discontinuation of breastfeeding (Clements et al. (1997)), breast development (Aksglaede et al. (2009)) and age at first intercourse (Miller et al. (1997)) etc. The event of interest can be either in years, months, weeks, days and hours, etc. Data that arise in different fields can be named survival data, duration data, failure time data and lifetime data. The time to event stream is now very wide with the inclusion of reliability and survival modeling. In reliability analysis, mathematical modeling is done for the lifetime of components and systems in the field of industry and engineering. Survival analysis deals with the data that arise from medical and biological experiments.

There are some examples taken from the literature that illustrates the applicability of time to event data in different fields. An example is the investigation of a carcinogenic substance in which a dose of some substance is given to laboratory animals and then they are observed to see if tumors developed. Here the event of interest is the time to appearance of a tumor since the dose is given. In another example, Nelson (1972), considered a life test in which specimens of a type of electrical insulating fluid were subjected to constant voltage stress. Here, the event time is taken as the time until each specimen failed or broke down. In a clinical trial discussed by Gehan (1965), the drug 6-mercaptopurine (6-MP) was compared to placebo with respect to the ability to maintain remission in acute leukemia patients. In this case, the lifetime variable is the remission time for each patient which is used to compare the effect of two drugs on patients.

The time to event of interest has origin zero. It has an important place in studying the behaviour of different events through time. Thus one needs to analyze the lifetime data with special models that can deal with time variables. The basic purpose of lifetime data is to estimate the survival functions, compare treatment effects or survival curves and study the effect of the covariate on time to event.

1.1 Censoring Scheme

Observing the lifetime data for each individual in an experiment is not always possible or recommendable. So a sample in such experiments may have partial information about the lifetimes. This happens due to various reasons like time constraints, cost constraints or nature of the experiment, etc. Keeping this in view, a number of censoring schemes are introduced in the literature. Some frequently used censoring schemes for time to event studies are as follows

1.1.1 Type-I Censoring Scheme

Let someone be interested in the lifetime-based inferences of subjects in an experiment. So the experiment will continue up to the failure of all subjects. But it is possible that to wait till the failure of the last subject is not feasible. It may increase the experiment time as well as cost unnecessarily. So, in such cases, it is needed to restrict the duration of the experiment. For example, in a study, 200 mice were divided into four groups on the basis of dose levels. In this study, each mouse was followed until its death or the prefixed termination time of the study. The termination time chosen in this study reduces the time and cost of maintaining the mice Klein and Moeschberger (2006).

This censoring procedure is termed as Type-I censoring scheme in literature. In this type of censoring scheme, the maximum time of test or study is fixed in advance. Suppose n identical units are put on the test and the experiment is performed under the controlled environment which results in identically and independently failure time points. Let the termination time of the test be T_0 and up to the time T_0, k failure observed in the experiment. Then the remaining $n_c = n - k$ units at the termination point T_0 of the experiment will be considered as Type-I censored. In a Type-I censored experiment, the termination time for the experiment (T_0) is fixed, but the number of failures is random in nature Balakrishnan and Aggrawalla (2000).

1.1.2 Type-II Censoring Scheme

In a life-testing experiment, a sufficient number of failures is required for better estimation. But the Type-I censoring scheme does not assure to have at least r failures out of n units on test. In Type-II censoring scheme, the test is performed until a prefixed number of failures, say, $r(< n)$ is observed. This censoring scheme ensures that minimum r units are observed with exact failure time points. In this case, the time of experiment (T_0) is random while a number of failures (r) is fixed. For example, in testing the life of electric lamps, the experiment is stopped after observing the prefixed number of failures to reduce the wastage of lamps. In another example, a fixed number of animals are exposed to a certain drug to determine its effect. It is observed that some of the animals take time to react. The experiment is terminated after a number of animals react to the drug instead of waiting for all animals to react Gupta (1952).

1.1.3 Random Censoring Scheme

Another type of censoring scheme which can be seen mostly in clinical and medical studies is random censoring (Dietz et al. (2002), Lawless (2003), Deshpande and Purohit (2015)). Random censoring occurs when a few units under study are lost or removed from the running experiment before its failure. Generally, patients do not follow the complete course of study and withdraw or leave the study before the termination of the study. In reliability engineering, electrical or electronic devices such as bulbs on tests may break before their failure. In such cases, the exact survival time (or time to event of interest) of the subjects is unknown; therefore they are called randomly censored observations. Hence in random censoring, the number of complete (uncensored) observations r and the time for which the study lasts T_0 are random.

1.1.4 Progressive Censoring Scheme

In many experiments, the removal of some of the surviving units is performed progressively during the experiment unlike in conventional Type-I censoring and Type-II censoring

schemes. Few removals are done with the failure of units in the experiment. So in the large sample, censoring becomes progressive in terms of removals at prefixed time points Cohen (1963), Balakrishnan (2007). This is the notion of the progressive censoring scheme in which a prefix number of units are removed randomly from the experiment out of surviving units at random time points.

Consider an experiment with n units put on to test. Let $T_1, T_2, \ldots, T_n$ be random variables denoting the failure times of the n units. The experimenter prefixed the number of failures, say, m. Also consider that on the first failure, R_1 units out of remaining surviving units $n-1$ are removed randomly on $T_{1:m:n}$; the first failure time. On the second failure time $T_{2:m:n}$, R_2 removals are done randomly out of $n - R_1 - 2$. Proceeding in similar manner, on the m^{th} failure time $T_{m:m:n}$, all the remaining surviving units $n - \sum_{i=1}^{m-1} R_i - m$ are removed. This is the description of the Type-II progressive censoring scheme. Similarly, we can define Type-I progressive censoring scheme.

1.1.5 Hybrid and Progressive Hybrid Censoring Schemes

The mixture of Type-I censoring scheme and Type-II censoring scheme gives hybrid censoring proposed by Epstein (1954). In the hybrid censoring scheme, one can control the termination time to experiment as well as the maximum number of failures simultaneously. Further, the more time on test and non-availability of the desired number of failures lead to a more advanced censoring scheme known as progressively hybrid censoring scheme (PHCS) which add removal in hybrid censoring. Two kinds of PHCS known as Type-I PHCS and Type-II PHCS are proposed by Epstein (1954) and Childs et al. (2003) respectively. For more details about the Type-I PHCS, one may refer to Balakrishnan (2007).

1.1.6 Interval Censoring Schemes

In some studies, we can not put a subject on continuous observation. So in that case one may visit at regular intervals to get the time to event of the subject. Hence the actual failure time (T) lies in a time interval $(L, R]$ which is based on the previous and current time to visit. Interval censored can be converted into right and left censored data with $R = \infty$ and

$L = 0$ respectively. For more details about interval-censored data one can see Huang and Wellner (1997), Gail et al. (2007), Chen et al. (2012). Also, two types of interval censoring Type-I and Type-II appear in the literature.

1.1.7 Current Status Data

Current status data can be considered as a particular form of interval-censored data and also known as Type-I interval-censored data. In current status data, we are only able to know whether or not the failure occurred before the monitoring time (S). Hence in current status data, the exact failure time of the event of interest is not known. Current status data is generated when the actual failure time points (T) are observed in form of either left or right censored (i.e. $R = \infty$ or $L = 0$). In such studies, monitoring time (S) can be fixed or random. For n units/individuals in the study, we observe current status data as joint distribution (T, S), however, only $\{(\delta_i, S_i) : i = 1, 2, \cdots, n\}$ is observed. Here δ is an indicator variable and defined as $\delta = I(T \leq S)$. So it is a binary variable and can take values 0 and 1 if $T > S$ and $T \leq S$. There are very few studies in parametric cases with current status data. For detailed description about the current status data one may refer to Groeneboom and Wellner (1992), Jewell and van der Laan (2003) and Klein et al. (2016).

Earlier work on current stats data can be observed in the demography where age at weaning is considered as an event of interest can be seen in Trussell et al. (1992), Grummer-Strawn (1993). Breast-feeding data from Pakistan Fertility Survey (PFS) is discussed in the article Diamond et al. (1986). The information only on the status of breastfeeding of a child is collected at the time of monitoring. In such types of data inaccuracy and bias can be found due to self-reporting or recalling the time of specific events. Another study that got much attention is based on HIV infection data. This data is based on both partners' sexual relationships in the long term. It is assumed that each partnership consists of a primary infected individual, the index case, who became infected through an external source. The time from infection of this index case to infection of a susceptible partner is then of interest. A necessary assumption is that the susceptible partner has no other means of infection other than contact with the index case. Another example of current status data can be

seen in the age-incidence estimation of a non-fatal human disease. Age at the incidence of Hepatitis A based on a cross-sectional sample of a given population examined by an accurate diagnostic test given by Keiding (1991). The use of current status techniques in this setting is further examined by Keiding et al. (1996). Some of the studies in literature based on the development of and exploration of current status data are Sun and Kalbfleisch (1993) and Rossini and Tsiatis (1996). A study of current status data is based on tumorigenicity experiments (Dinse and Lagakos (1983),Dewanji and Kalbfleisch (1986)). In such studies, the tumor onset time is the main event of interest.

1.2 Recall-Based Data

Recall-based studies are generally seen in cross-sectional studies either in clinical trials, demography or in some other survey-based studies. In such studies at the monitoring time (S), the event of interest may already occur or not. If the event has already occurred then an individual is asked to recall the time to event (T). An individual may recall the time to event exactly, partially. It may also be a non-recall observation. Since these studies are based on retrieval of memory, they are named recall-based data.

Turnbull and Weiss (1978) discussed failure data generated from the study on the use of marijuana by high school students. In the data 191 California students were asked the question "When did you first use marijuana?". Some of the boys exactly remembered the exact age when they first used the marijuana, some of them had never used it at the date of the interview generating right-censored data. At last some of the boys who used marijuana but were not able to recall the exact age at the time of interview generating left censored data.

In another study, the information on variables related to age at menarche is collected. The data was used by Salehabadi et al. (2015) first to estimate the age at menarche through parametric modeling. The considered menarche data is a part of the project titled "Physical Growth, Body Composition and Nutritional Status of Bengali School aged Children, Adolescents and Young adults of Calcutta, India: Effects of Socioeconomic Factors on Secular

Trends". The data was collected by the Biological Anthropology Unit of the Indian Statistical Institute, Kolkata in duration $2005-11$. The original survey included approximately 4000 randomly selected individuals for the study. A subset of this data which includes 2195 girls aged between $7-21$ is used for study purposes. The girls were surveyed at any age and asked about the occurrence of age at menarche. The girls who had not experienced the event at interview time are treated as right-censored observations. If the girl had experienced the event, she was further asked to recall the time of occurrence of menarche. Some of the girls were able to recall the age at menarche exactly, some of them recalled it at month level and some of them at year levels and some were unable to recall anything. Also, the data contains the status of menarche, time to occurrence of menarche, birth order and some other socio-economic factors.

1.3 Competing Risk

The theory of competing risks is used in reliability theory and survival analysis when the occurrence of an event of interest is due to one of the causes acting on it. In survival analysis a patient is at risk of two or more than two mutually exclusive causes and the death of the patient may be due to any of these causes. In a similar fashion, the failure of a system or machine may be due to the failure of any of its components. In case of competing risks, the data can be observed in form of bivariate distribution i.e. time to event and corresponding cause of failure. The concepts and estimation under competing risks in reliability engineering and survival analysis along with real-life examples can be seen in details in Mann et al. (1974), Sinha (1986), Crowder (2001), Lawless (2003) and Lee and Wang (2003).

Let us discuss some real-life situations where competing risks data arises. It has abundant applicability in survival analysis and reliability engineering. In the case of bone marrow transplantation, the cause of death may be different kinds of infections (bacterial, viral, fungal) or death can be due to relapse from some other causes. In the case of cancer studies, the event of interest may be death from cancer and the death due to other causes (surgical

mortality, old age) can be competing risks. In another study based on HIV, the interest is in the time to HIV infection (the incubation time) and to compare the risk of disease by groups among injecting drug users, about 20 percent of the HIV infected individuals die before an AIDS diagnosis. Here, death before AIDS is a competing risk.

In some cases, the cause of failure is not observed fully due to time or cost constraints. Such situations in which we have incomplete data on the cause of failure generate masked data problems. The competing risks analysis with partial information on cause becomes more challenging to deal with. In a time to failure based data on computer component systems (Reiser et al. (1995)), where failure occurs due to malfunctioning of one of three components the motherboard, power supply and disc drive. Here the exact cause of failure is not known and generates masked data. ML estimators for two-component series systems with exponential as failure time is discussed by Miyakawa (1984). Generalizing the results from Miyakawa (1984), Usher and Hodgson (1988) introduced the concept of masked data under exact and partial masking. Mukhopadhyay and Basu (1993) applied E-M algorithm to solve the masked data problem for J-component series system considering the components' failure time as exponentially distributed. The alternate closed-form solution to masked data problem discussed by Usher and Hodgson (1988) is obtained by Lin and Guess (1994). An alternative ad-hoc iterative approach is developed to find ML estimators by Usher (1996) with Weibull as lifetime distribution under masked data setup. Analytical solution to three-component masked data problem under Bayesian approach is discussed by Flehinger et al. (2002a) with the series system having exponentially distributed component's failure. Sarhan (2003) derived ML and Bayes estimators of reliability functions of individual components for the series system with Pareto distribution as lifetime model. ML and Bayes estimators for a family of distribution are obtained by Singh and Tomer (2011) and Tomer et al. (2014). In a recent study, Rai et al. (2021), masked data is analyzed under classical and Bayesian approaches with Lindley distribution considering masking probabilities as cause and time-dependent.

1.3.1 Current Status data with Competing Risks

The current status data generally arises in cross-sectional studies. As discussed earlier that in current status data there is only one monitoring point for each individual in the study. Along with the monitoring point if the cause of occurrence is also observed then the current status under competing risks setup is formed. For example, Krailo and Pike (1983) discussed the menopause data on US women observed between $1960-62$. In the menopause data, age at menopause with the corresponding cause of menopause (natural/operative) is observed. In another cross-sectional example on HIV studies, the status of individuals is available with corresponding subtypes of HIV as competing causes.

There are several studies that are based on current status data with competing risks. Most of the work is based on non-parametric studies. Hudgens et al. (2001) derived the non-parametric MLE of cumulative incidence functions for competing risks survival data and applied the method to data from a cohort of injecting drug users in Thailand susceptible to infection from HIV-1 subtypes B and E. In the article, Jewell and van der Laan (2003) discussed current status data under competing risks situations along with some motivating examples. The thesis (Maathuis (2006)) discussed the data that arises in cross-sectional survival studies with more than one cause. HIV vaccine clinical trial data is analyzed under current status with competing risks data set up. The estimates of distribution functions were obtained with a non-parametric maximum likelihood estimator (MLE) and naive estimator and some asymptotic properties are also discussed.

Also, in some studies, along with monitoring time, the cause of the event is not known and generates incomplete data problems. Koley and Dewanji (2018) discussed parametric analysis of current status data under the general missing pattern of causes. Maximum likelihood is obtained and asymptotic properties of MLE are also discussed. In a study, Koley and Dewanji (2019a), non-parametric estimates for current status data under missing causes with independent missing probabilities were obtained with two assumptions when monitoring time is fixed and random. Asymptotic properties for non-parametric ML (NPML) estimator are also discussed. NPMLE of sub-distribution is obtained by Koley and Dewanji (2019b) by

utilizing the re-parametrization technique under the current status competing risks scenario. The E-M algorithm is used for the estimation of the sub-distribution function and some asymptotic results were established for ML estimators. Koley and Dewanji (2020a) give the ML estimate of parameters and sub-distribution functions under parametric and non-parametric approaches taking missing probabilities as time-dependent. Asymptotic results are also discussed for ML estimators. Finally, simulation studies and real data based on hearing loss (Banik et al. (2018)) are discussed for illustration purposes. In the study Koley and Dewanji (2020a), authors utilized some additional information generated either through validation sampling or through some prior on masking probabilities and obtained estimates under parametric and non-parametric approaches. In such studies sometimes the masking probabilities are assumed to be time-dependent and accordingly estimation procedure is developed.

1.4 Estimation Methods

In recall-based studies, parametric modeling of time to event data is done primarily by assuming a lifetime distribution. The main focus is on the unknown parameters of the lifetime distribution under consideration. The population under these studies consists of many unknown parameters depicting the characteristics of that population. It is necessary to make inferences about those parameters based on a sample selected from that population which is more or less a true representative of it in order to ascertain the population's hidden nature. There are primarily two estimation approaches available in statistical literature to make inferences about the unknown parameters of the population under study. The estimation is carried out using two approaches: the classical method of estimation and the Bayesian approach to estimation.

1.4.1 Maximum Likelihood Method

The maximum likelihood method is one of the most popular methods of estimation in the classical paradigm. In this method, the estimation procedure is carried out by optimizing a

function named likelihood function, say, $L(\theta \mid x)$; where x is observed random sample from a population characterized by θ which may be scalar or vector of parameters.

Now suppose that we have a random sample of size n, i.e. $x = (x_1, \ldots, x_n)$ from a population with density $f(x;\theta)$. Then the likelihood function $L(\theta \mid x)$ is a real-valued function of unknown parameter θ with domain Θ and range $\mathcal{R}^+$. In mathematical notation, the likelihood function at $x \in \mathcal{X}$ is the function $L(\theta \mid x) : \Theta \to \mathcal{R}^+$ given by

$$L(\theta \mid x) = \prod_{i=1}^{n} f(x_i;\theta); \quad \theta \in \Theta$$

The idea behind the maximum likelihood (ML) estimator is that one chooses that value of unknown parameter θ which maximizes the likelihood of observing the sample x from the density $f(x;\theta)$ in hand. It is clear that maximizing the likelihood function is quite tedious work as it is a product of n quantities $f(x_i;\theta)$; $i = 1, \ldots, n$. So our problem reduces to maximizing the corresponding log-likelihood function denoted by $\ln L(\theta \mid x)$ as it is a monotone increasing function and attains maximum value at the same point at which the likelihood function attains its maximum. Then the log-likelihood function at $x \in \mathcal{X}$ is given by taking the natural logarithm of likelihood function $L(\theta \mid x)$ i.e.

$$\ln L(\theta \mid x) = \ln f(x;\theta); \quad \theta \in \Theta$$

We say that the statistic $t = t(x) \in \Theta$ is the maximum likelihood estimator of θ if the maximum value of $\ln L(\theta \mid x)$ occurs at point $t(x)$. It is noted that the ML estimator is not always unique and when unique it will be a function of sufficient statistics of θ.

1.4.2 Bayesian Approach

Under the Bayesian paradigm of estimation, the quantity of interest, i.e. the unknown parameter, say, θ which has to be estimated is considered to be a random variable having its own probability distribution. Regardless of the fact that θ is considered to be a constant but unknown quantity in the classical paradigm, the subjective belief about it is being updated in light of observed samples. The likelihood function still plays an important role

in estimating the parameter θ in the Bayesian paradigm.

Let us assume that $\pi(\theta)$ depicts our prior belief about θ and $L(\theta \mid x)$ be the likelihood function of this unknown quantity for the given observed sample x. Now to update our prior belief, we use the renowned Bayes theorem. The updated belief is known as posterior density and we denote it by $p(\theta \mid x)$. The posterior density up to proportionality is given by

$$p(\theta \mid x) \propto \pi(\theta) L(\theta \mid x),$$

where on removing the proportionality sign, we get a normalizing constant which makes the posterior distribution proper. Also, any inference which has to be carried out for θ is done using the posterior density $p(\theta \mid x)$.

The choice of prior depends upon the knowledge available to us on the random quantity θ. Based on it, the whole Bayesian paradigm further can be bifurcated into two dimensions. One is an objective Bayesian method and the other one is a subjective Bayesian method of estimation. We will discuss these two one by one.

Non-informative Priors

Sometimes it is possible that the information on θ available to us is quite vague in nature and we don't want to make inferences using posterior which is going to be dominated by the prior. So in this case, we carry out the inference under the non-informative prior. The non-informative prior makes the posterior inference dominated by observed data only through the likelihood function and makes the prior impact least. This whole method of estimation falls under the objective Bayesian method. A non-informative prior for the one-dimensional parameter θ is given by

$$\pi(\theta) = 1; \quad \theta \in (0, 1).$$

Whenever our prior belief is under suspicion so that we cannot assign weight to any particular value of θ, in that case, we can use the uniform prior defined above. The uniform prior gives equal weight to all possible values of θ in $(0, 1)$. Thus making it a suitable choice for non-informative prior. In the objective Bayesian method, the observed sample plays an

important role in dominating the posterior inference through the likelihood function over the prior. The information about the unknown quantity θ in the observed sample is given by the Fisher information matrix. Keeping this in mind, Jeffrey's formulated a more strong non-informative prior which is known as Jeffreys prior after his name. The Jeffreys prior, say, $\pi(\theta)$ is given by

$$\pi(\theta) \propto \sqrt{I(\theta)},$$

where $I(\theta)$ is the Fisher information matrix for observed sample x. Thus Jeffreys utilizes information contained in the sample to derive the prior density and thus posterior density is dominated by observed data over the prior.

Informative Priors

In the other case, when the prior belief is not susceptible to any vagueness, this belief about θ can be formulated with any suitable probability distribution available in the statistical literature. Thus the inference is done under the informative prior. A kind of informative prior in which the posterior distribution belongs to the same family of distribution as the prior belongs is known as conjugate prior. In the case of informative priors, the parameters of the priors are known as hyper-parameters. One problem that arises while using the informative prior for the parameters, is how to choose the suitable values for the hyper-parameters of the prior distribution. There is vast literature available on choosing the hyper-parameters values for different distributions. In our work, we prefer moment matching criteria to choose the values whenever it is necessary to carry out the Bayesian estimation.

1.5 Computational Techniques

The model formulation and developing estimation procedure is limited to a theoretical perspective. When it comes to the computational part, one faces many problems arising due to incompleteness of data, intractable form of the likelihood function and so on. To tackle these, we use a number of computational techniques discussed below:

1.5.1 Expectation-Maximization Technique

The Expectation-Maximization ($E-M$) technique is one of the most useful techniques for enabling the estimation in case of missing data. The recall-based study always confronts the problem of missing data with respect to the time or the cause of the event. So we apply this technique to estimate the parameters more efficiently. The $E-M$ algorithm is first proposed by Dempster et al. (1977) in the literature to handle the missing data and shows its application in mixture modeling. It employs the fact that the missing data will have the same behavior if it was observed in the sample. The functionality of the algorithm is given by two steps: Expectation E- step and Maximization M- step.

Let us assume that a random sample on X is observed from the population having density $f(x;\theta)$. The random variable Z denotes the missing data in the sample. We will define the complete data vector as $Y=(X,Z)$. Now the problem of estimation of unknown parameter θ is handled by the Expectation-Maximization technique.

In the E- step, the expectation of complete log likelihood function is taken over the missing data variable Z for a given initial value of θ, say, θ_0, i.e.,

$$\mathcal{Q}(\theta \mid \theta_0) = \mathbb{E}_Z(\ln L(\theta \mid x,z) \mid x,\theta_0)$$

Further in the M step, the parameter θ is computed by maximizing the expected complete log likelihood function obtained in E- step. Let at the k^{th} iteration, the value of θ is θ^k, then the value of θ at $(k+1)^{th}$ iteration is given by

$$\theta^{(k+1)} = \underset{\theta}{\operatorname{argmax}}\, \mathcal{Q}(\theta \mid \theta^k).$$

Thus the $E-M$ algorithm iterates between these two steps to compute the unknown parameter θ until a convergence criteria is satisfied.

1.5.2 Missing Information Principle

EM algorithm discussed in the previous subsection does not provide any direct method to calculate variance-covariance method of ML estimators. Thus the observed fisher information matrix can be calculated using the missing information principle proposed by Louis (1982). Consider the missing data problem as discussed earlier. Suppose $f(x;\theta)$ be the observed data density, $g(x,z;\theta)$ be the complete data density and $h(z;\theta) = \frac{g(x,z;\theta)}{f(x;\theta)}$ be the missing data density. The observed missing information matrix denoted by $\mathcal{I}_{\mathcal{X}}(\theta)$ is given by

$$\mathcal{I}_{\mathcal{X}}(\theta) = \mathcal{I}_{\mathcal{Y}}(\theta) - \mathcal{I}_{\mathcal{Z}|\mathcal{X}}(\theta),$$

where $\mathcal{I}_{\mathcal{Y}}(\theta)$ be the fisher information matrix for complete data $Y = (X, Z)$ and $\mathcal{I}_{\mathcal{Z}|\mathcal{X}}(\theta)$ be the fisher information matrix for missing data Z.

1.5.3 Markov Chain Monte Carlo Technique

It is often the case in the Bayesian method of estimation that the samples can not be generated directly from the posterior density, say, $p(\theta \mid x)$. The fact that obtained posterior $p(\theta \mid x)$ is generally not in analytical form, makes one helpless to draw samples from the density. The simulation technique is needed to do so. Thus the posterior sample can be generated using Markov Chain Monte Carlo ($MCMC$) techniques proposed by Metropolis et al. (1953) and later on extended by Hastings (1970). $MCMC$ technique consists of a number of algorithms to draw samples from the target distribution. The term Monte Carlo refers to the fact that a population can be characterized by drawing a large number of samples from it and estimation is done using the sample. While the term Markov chain shows that a sequence of random samples is drawn from the population in a way that the current observed sample only depends on the previously observed sample, a property of the Markov chain. Thus MCMC technique of sample generation as a whole consists of steps of drawing a sequence of sample observations that hold the Markov property. The MCMC samples are very sensitive to the starting point chosen to initiate the chain and thus a sufficient number of sample observations initially obtained are discarded.

1.5.4 Metropolis-Hastings Algorithm

As in the case discussed earlier, the form of posterior density is not always in analytical form. It results in difficulty to directly draw samples from it. In this case, the Metropolis-Hastings $(M-H)$ algorithm, a kind of $MCMC$ technique, is found very useful. Because of a non-standard form of posterior density, the samples are generated from another density which covers the support of target density. This density is known as proposal density. A chain of random samples is drawn from the proposal density and each sample observation is then accepted or rejected with some probability. In the long run, the obtained sample observations which are accepted with some probability are treated as drawn from the target density after discarding the initial sample observations. The obtained sample observations are known as samples obtained after the burn-in period. Suppose the target density is $p(\theta)$ and the proposal density is $q(.)$. To obtain the samples from $p(\theta)$, the Metropolis-Hastings algorithm steps are given below:

Initialize the value of θ, say, θ^0. For $k = 1, 2, \ldots, N$

- Draw a sample observation θ from the proposal density $q(\theta \mid \theta^{k-1})$
- Compute the following probability

$$\gamma(\theta, \theta^{k-1}) = \min\left(\frac{p(\theta)}{p(\theta^{k-1})}.\frac{q(\theta^{k-1} \mid \theta)}{q(\theta \mid \theta^{k-1})}, 1\right)$$

- Accept θ with probability $\gamma(\theta, \theta^{k-1})$ and set θ^k=θ; otherwise set θ^k=θ^{k-1}

Iterating the steps of the algorithm, we have a sample $(\theta^1, \theta^2, \ldots, \theta^N)$ whose limiting distribution will be $p(\theta)$. Since the choice of initial value has a great impact on the generated chain so we discard the first M observations from the sample obtained. Thus the remaining sample observations $(\theta^{M+1}, \theta^{M+2}, \ldots, \theta^N)$ is considered to be drawn from $p(\theta)$.

1.5.5 Gibbs Sampling Algorithm

Gibbs Sampling algorithm is used when one has to draw samples from the joint density of more than one parameter in the Bayesian paradigm. Suppose $p(\theta, \lambda)$ be the joint density of θ

and λ but it is not possible to generate samples on (θ, λ) directly from the joint density. Gibbs sampling technique is a computer-simulated algorithm that enables one to draw samples from their conditional densities. Suppose it is possible to draw samples from conditional densities of θ and λ, say $p(\theta \mid \lambda)$ and $p(\lambda \mid \theta)$ respectively. The steps to generate sample observation on (θ, λ) is given below:

- Initialize the vector (θ, λ), say (θ^0, λ^0).

  For $k = 1, 2, \ldots, N$

- Draw a sample observation from the conditional density $p(\theta \mid \lambda^{k-1})$, say, θ^k.

- Further draw a sample observation from the conditional density $p(\lambda \mid \theta^k)$, say, λ^k.

On iterating the above step for N number of times, we have a sequence $\{\theta, \lambda\}_{k=1}^{N}$ of observations on (θ, λ).

1.5.6 Bayesian Credible Intervals

The Bayesian credible intervals (BCIs) for θ_k can be obtained from MCMC samples

$$\theta_k^{(1)}, \theta_k^{(2)}, \cdots, \theta_k^{(N)}$$

based on posterior density $p(\theta_k \mid x)$. We first arrange the samples into ascending order and denote them as $\theta_{k(1)}, \theta_{k(2)}, \cdots, \theta_{k(N)}$. Then using the method of Chen and Shao (1999), we can obtain the $100(1-2\alpha)\%$ Bayesian credible interval estimates as

$$\left(\theta_{k[\alpha N]}, \theta_{k[(1-\alpha)N]}\right)$$

where $\theta_{k[qN]}$ denotes the $[qN]^{th}$ smallest in the chain $\theta_{k(1)}, \theta_{k(2)}, \cdots, \theta_{k(N)}$.

1.5.7 Highest Posterior Density Intervals

The credible intervals having the shortest lengths is a better estimate for the interval estimates. Such intervals are named as highest posterior density (HPD) intervals. HPD intervals

have a property that the minimum density of any point within this interval is equal to or larger than the density of any point outside it. For the construction of HPD intervals, we obtain all $100(1-2\alpha)\%$ credible intervals, that is

$$\left(\theta_{(k)}, \theta_{(k+[(1-2\alpha)N])}\right);\ k = 1, 2, \cdots, [2\alpha N]$$

and then evaluate the corresponding lengths which are given by

$$l_k = \theta_{(k+[(1-2\alpha)N])} - \theta_{(k)}.$$

Finally, we pick the interval having a minimum length.

1.6 Real Data Fitting

In recall-based studies, our aim is to build a statistical model which nearly depicts the behavior of the real data in hand. After the model formulation, the next step is to find suitable data for illustration of the established model. Before choosing any real data, it is inevitable to validate the fitting of that data for the model. There are a number of techniques used in statistics for real data fitting. Some of them are the K-S test, Chi-Square goodness of fit test and so on. It is used mainly when one has counted data on variables under study. It can be used for discrete as well as for continuous data.

1.6.1 Chi-Square Test

Chi-Square Goodness of fit test is used in the case when one has to show that the data in hand can be assumed to come from the established model. Thus one tests the null hypothesis that "The data comes from the specified distribution" against the alternative hypothesis that "It does not come from the specified distribution". It must keep in mind that the expected frequency must be at least 5 for the chi-square approximation to be accurate.

The test statistic in this case is defined as

$$\chi^2 = \sum_{i=1}^{n} \frac{(o_i - e_i)^2}{e_i},$$

where o_i is the observed frequency in the data for i^{th} group and e_i be the expected frequency of that group estimated under the null hypothesis.

We reject the null hypothesis at a pre-specified level of significance α if the calculated value of test statistic χ^2 is greater than $\chi^2_{1-\alpha,n-1}$.

In our work, we have used a modified version of the Chi-Square test proposed by Hope (1968) and recently used by Koley and Dewanji (2021) to show the fitting of current status competing risk data. The method is based on simulation procedure and the suitability of assumed distribution to considered data can be checked based on p value. The calculation of the p value for the defined test statistic can be done as described ahead. The parameters of the model are first calculated using the proposed method and then the modified χ^2 statistic is calculated using observed data. Then a large number of datasets (say, 1000) is generated of the same sample size n with estimated model parameters and then we calculated modified chi-square statistics for each simulated data. The p value is calculated as the proportion of times the modified χ^2 statistic exceeds the observed value.

Chapter 2

A Latent Variable Approach for Modeling and Analysis of Recall-Based Time to Event Data

2.1 Introduction

Time to event studies arise in many fields such as epidemiology, public health, clinical research, demography, and engineering. In such studies, the event of interest may be death, relapse from remission, hospital admission/readmission, failure of a component, or breakdown of a system. For example, in demographic studies, the event of interest may be the onset of menarche in young girls, menopause in women, or the first birth. If the individuals are followed up to the occurrence of an event of interest, complete data can be observed. It is not possible to obtain complete data in all real-life situations. Particularly, in survival analysis and reliability theory, we are interested in time to event based on living and non-living units. In survival scenarios, the study is done on living individuals (animals), so there are certain complications that get involved at the level of collection of data with respect to time and the number of occurrences of events. The individuals are taken into an experiment and the experiment is terminated at a prefixed time or when the event of interest is experienced by a prefixed number of individuals. Thus, it is very unlikely to observe the exact time to event for each individual and this incompleteness in time to event observation results in censoring. For details one may refer to Lawless (2003) and Nelson (1972). There are various censoring patterns like left, right and interval that arise in such experiments.

Left censored data arises in the cases where the event is already experienced by the individual before entering into the study. Thus in left-censored data, the only information available is that the event had happened before any particular time. Whereas in the right-censored data there is the possibility that few individuals did not experience the event of interest up to a given point in time. For example; we refer to an example from the book Andersen et al. (2012) where the event of interest is to estimate the time to baboon's descend from the trees. Some baboons, sleeping in the trees, were followed up for a number of days. And they descend from the trees every day. On any day in the morning we arrive at 9 am and it is noted that some baboons already descend from the tree before our arrival. They generate the left-censored data as we have only information that the time to baboon's descend from the trees is less than the time of arrival. The right-censored data can be observed in the case when at the end of the experiment some units/individuals still do not experience the event of interest. The interval-censored data arises generally in the case of clinical studies, public health, and cross-sectional studies where continuous monitoring of the individuals is not feasible. In such studies, individuals are visited at regular time intervals and the status of events of interest is noted. De Gruttola and Lagakos (1989) discussed the case of acquired immune deficiency syndrome (AIDS) trials, where the event of interest is time to AIDS for the HIV-infected individuals. For the diagnosis of disease, the testing of blood is done periodically rather than continuous monitoring, so the data is observed in the form of interval-censored.

Current status data is a particular case of interval-censored observations in which time to event of interest is not observed exactly but it is either left-censored or right-censored. Each individual in the study is monitored at a specific time and status (occurrence/non-occurrence of the event) is observed. In particular, if the event of interest occurs before the monitoring time point then observation is left-censored while if the event has not yet occurred at the monitoring time then observation is called right-censored. These datasets are also studied with the name of status quo or dichotomous data. These types of data are easy to obtain in any study by simply observing the status of a individual at any particular time. So these datasets can be modeled by simply using binomial type likelihood or probit

models. These datasets are generally generated in prospective studies. An example that deals with the estimation of the distribution of age at weaning under the current status data setup discussed can be seen in the article Diamond et al. (1986), where the age of a child at weaning is considered as time to event and the age of the child as observation time. For more details about the current status data, one can see Groeneboom and Wellner (1992), Sun and Kalbfleisch (1993), Jewell and van der Laan (2003) and Klein et al. (2016). Where a detailed review of the literature and several examples of current status data is available.

On the other hand in several retrospective studies, the individuals are monitored at any time and the information about events of interest is collected. If the individual has experienced the event of interest at the monitoring point, then it may be the case that they are further asked to remember the exact time when they experienced the event. These studies are named recall-based studies in the literature. In such studies, the data can be observed based on the individual's responses or available past records. In real-life situations, it can be observed that the chance of recalling the time to event decreases with the passing of time. It means that an individual who has experienced the event recently has more chance of recalling the event than that of the individual who has experienced the event earlier. Survey sampling-based studies, demographic surveys, and clinical trials generate such types of datasets. Applications of these studies are found in estimating the age at menarche [Salehabadi et al. (2015)], age at menopause [Krailo and Pike (1983)], age at weaning [Diamond et al. (1986)], etc. For example, one can see the menarche data of the girls aged $7-21$ years discussed in the article Salehabadi et al. (2015). Initially, these data were studied under binary responses, status quo, interval censored, and non-informative censoring schemes. Some studies related to recall-based events are found relevant in the direction of modeling these datasets. In a recall-based study based on age at menarche, Koo and Rohan (1997) concluded when the data were grouped by the interval of recall, higher accuracy was observed with a shorter interval. In the study, Cooney et al. (2009), authors obtained time to pregnancy and checked its validity based on previous studies. Gillespie et al. (2006) studied the accuracy of recall-based breastfeeding data on women in Michigan and Nebraska (USA). The recall-based data can be treated as interval-censored

data and estimates are obtained using methods proposed by Turnbull (1976). But this method ignores the informative nature of recall-based data. Due to additional information in the data, censoring is informative in the sense that the chance of recalling the event decays as the difference between monitoring time and time to the event becomes large. Ignoring this information results in a large bias in estimates, so this information can be utilized in a model to improve the estimates. Utilizing this information, Salehabadi et al. (2015) proposed a parametric model for such data. The functional form of the chance of recalling the event was taken as exponential and piecewise constant functions. They concluded that utilizing the extra information improves the estimates. Further, Salehabadi and Sengupta (2016) developed non-parametric models for the recall-based studies. In a recent study, a detailed review for modeling cross-sectional current status data is given by Salehabadi and Sengupta (2018).

In the case of recall-based data, the individuals may be able to recall the event exactly or not. Generally, in such cases, there might arise three possibilities:

1. The event has occurred before the monitoring time and the individual is able to recall the exact time of the event, called recall,

2. The event has occurred before the monitoring time but the individual is not able to recall the exact time, called as non-recall,

3. The event has not occurred at the monitoring time point, say right-censored.

These three cases should be included in the construction of likelihood. The additional information can be utilized by a suitable choice of non-recall probability. Moving farther from time to event reduces the chance of recalling the event. So, it can be taken as a function of the difference between these two.

In this chapter, our aim is to develop a parametric model using a latent variable approach for recall-based data. An estimation method based on the Expectation-Maximization (E-M) algorithm has been proposed to deal with recall-based data in Section 2.2. The point estimator of the parameters is derived using the proposed method. The observed Fisher information matrix is calculated using the missing information principle in order to con-

struct the confidence intervals for unknown parameters. Section 2.3 deals with the Bayesian estimation under informative priors. Several statistics such as probability of experiencing an event, expected time in study, total time in study, expected number of recall and non-recall observations are estimated to explore the impact of monitoring time points in Section 2.4. A simulation study is carried out to illustrate the performance of proposed estimators in Section 2.5. The duration of breastfeeding for the recent children from NFHS-IV data is estimated using the proposed methodologies and reported in Section 2.6. Section 2.7 is devoted to the discussion and future scope of the problem.

2.2 Classical Estimation

2.2.1 Likelihood Construction

As in Section 2.1, we have described the basic structure of the current status data for recall-based studies. Now to construct the likelihood function we need to incorporate the three cases: recall, non-recall, and right-censored. Let the random monitoring time be denoted by S with distribution and density functions $G(.)$ and $g(.)$ respectively. Also, let T be a random variable for the onset of the event of interest with distribution and density functions $F(.)$ and $f(.)$, respectively. Here, it is assumed that T and S are independent variables. Let us consider that there are n individuals in the study and each individual is visited at prefixed monitoring time and the status of the event is noted. Here, it is noted that more than one individual can be monitored at the same time point. The time to event T is observed only if the individual can exactly recall the time to event of interest. In the case of non-recall, the exact time to onset of the event is not observed but it is known that it lies before the monitoring time. And in the right-censored case, the event has not yet been experienced by individuals at the monitoring time. Thus, the observed random vector can be denoted by $\underline{D} = (S, \delta, \epsilon, T\delta\epsilon)$, where $\delta = I(T \leq S)$ is an indicator variable attached with the onset of the event, which takes value 1 if the event of interest has occurred before the monitoring time and 0 otherwise, ϵ is another indicator variable, which takes value 1 if individual exactly recalls the time to event of interest and 0 otherwise. Let $D_i = (S_i, \delta_i, \epsilon_i, T_i\delta_i\epsilon_i);\ i = 1, 2, \cdots, n$

be the independent and identically distributed random vector attached with i^{th} individual and d_i denotes the observed value of D_i. Now, the likelihood for the observed data d and parameter vector θ can be defined as

$$L(\theta|d) = \prod_{i=1}^{n} \Big[f(t_i;\theta)\Big(1-\psi(s_i,t_i)\Big)\Big]^{\delta_i\epsilon_i} \Big[\int_0^{s_i} f(u;\theta)\psi(s_i,u)du \Big]^{\delta_i(1-\epsilon_i)} \Big[\bar{F}(s_i;\theta)\Big]^{(1-\delta_i)}, \tag{2.1}$$

where $\bar{F}(.)$ is survival function for time to event and $\psi(s,t)$ represents the non-recall probability. The suitable choice of the functional form of non-recall probability is an important issue. In literature, there are some functional forms considered for defining non-recall probability. Salehabadi et al. (2015) analyzed recall-based data using exponential and piecewise functional forms. In competing risks scenario, Sukumaran and Dewan (2019) studied recall-based data using the piecewise functional form for the non-recall probability. Further, Mirzaei Salehabadi et al. (2020) had used a multinomial logistic regression model for the choice of non-recall probabilities. In real-life situations, the chance of recalling an event for an individual depends on the elapsed time between monitoring time and the time of the event and decreases exponentially. So, the exponential form of non-recall probability is a suitable choice and considered for further analysis purposes.

If the i^{th} individual experiences the event but is unable to recall the time to event exactly then the non-recall probability with parameter λ takes the form

$$\psi(s_i,t_i) = 1-\exp\{-\lambda(s_i-t_i)\}; \quad \lambda > 0 \tag{2.2}$$

Alternatively, the recall probability can be written as $1-\psi(s_i,t_i) = \exp\{-\lambda(s_i-t_i)\}$. As the difference between time to event and monitoring time increases the chances of non-recall increases with rate λ.

For the given sample of size n, let n_r, n_{nr}, n_c denote the number of recalled, non-recalled and right-censored observations, respectively, where $n_r = \sum \delta_i\epsilon_i$, $n_{nr} = \sum \delta_i(1-\epsilon_i)$; $i = 1,2,\ldots,n$ and $n_c = n - n_r - n_{nr}$.

Now, putting the values of recall and non-recall probabilities in (2.2), the likelihood function can be written as

$$L(\theta,\lambda|d) = \prod_{i=1}^{n_r}\Big[f(t_i;\theta)\exp\{-\lambda(s_i-t_i)\}\Big]\prod_{i=1}^{n_{nr}}\Big[\int_0^{s_i} f(u;\theta)\Big(1-\exp\{-\lambda(s_i-u)\}\Big)du\Big]$$
$$\prod_{i=1}^{n_c}\Big[\bar{F}(s_i;\theta)\Big]. \tag{2.3}$$

The likelihood function (2.3) contains non-recall and right-censored observations and thus it is not in a convenient form to calculate the maximum likelihood estimators using usual approaches.

We can treat the problem as missing or incomplete and to deal with such situations the expectation-maximization (E-M) algorithm comes into role. The E-M algorithm which was proposed by Dempster et al. (1977) specially deals with missing data that arises in censoring, clustering and mixture modeling. In the case of non-recall, it is observed that the exact time to event (T) lies in the interval (0, S). Hence we introduce a latent variable T_l^* utilizing the equivalent quantity approach [see Tan (2007), Tan (2009), Wang (2016) and Fan et al. (2019)]. For the right-censored observations, we have information that the time to event is greater than monitoring time and thus another latent variable $T_r^* \in (S,\infty)$ is introduced. Now, with the help of these latent variables, the likelihood function (2.3) becomes

$$L(\theta,\lambda|d) = \prod_{i=1}^{n_r}\Big[f(t_i;\theta)\exp\{-\lambda(s_i-t_i)\}\Big]\prod_{i=1}^{n_{nr}}\Big[f(t_{li}^*;\theta)\Big(1-\exp\{-\lambda(s_i-t_{li}^*)\}\Big)\Big]$$
$$\prod_{i=1}^{n_c}\Big[f(t_{ri}^*;\theta)\Big]. \tag{2.4}$$

Further, to make the non-recall probability into convenient form, we introduce an exponential latent variable W_l^* having mean $1/\lambda$ and truncated at point $(S-T_l^*)$. We note that $P\Big(w_{li}^* < (s_i - t_{li}^*)\Big) = 1-\exp\{-\lambda(s_i-t_{li}^*)\}$, which is the term in likelihood. Thus, with the help of latent variable W_l^*, we can transform the term $1-\exp\{-\lambda(s_i-t_{li}^*)\}$ into $\lambda\exp\{\lambda w_{li}^*\}$. Let, the data vector corresponding to introduced latent variables is de-

noted by $\underline{D}^* = (T_l^*, W_l^*, T_r^*)$. For i^{th} individual the latent observation can be denoted as $D_i^* = (T_{li}^*, W_{li}^*, T_{ri}^*)$ and d_i^* denotes the observed value of D_i^*. Thus, the complete data vector is written as $(\underline{D}, \underline{D}^*)$. In light of complete data, the complete likelihood can be written as

$$L_c(\theta, \lambda | d, d^*) = \prod_{i=1}^{n_r} \Big[f(t_i; \theta) \exp\{-\lambda(s_i - t_i)\} \Big] \prod_{i=1}^{n_{nr}} \Big[f(t_{li}^*; \theta) \lambda \exp\{-\lambda w_{li}^*\} \Big] \prod_{i=1}^{n_c} \Big[f(t_{ri}^*; \theta) \Big]. \tag{2.5}$$

Next important step is the choice of time to event distribution. Weibull distribution is widely used lifetime distribution in survival analysis as it facilitates to model increasing, constant, and decreasing hazard rates when the shape parameter is greater than 1, equal to 1, and less than 1 respectively. Due to this flexibility, fundamental researches in different fields are done considering Weibull as time to event distribution. So, for further analysis purposes, we consider time to event distribution as Weibull, i.e. $T \sim \mathcal{W}(\alpha, \beta)$, where α and β denote the shape and scale parameters respectively. Let us redefine the parameter vector as $\Theta = (\alpha, \beta, \lambda)$. Under these assumptions, the complete likelihood becomes

$$L_c(\Theta | d, d^*) = \prod_{i=1}^{n_r} \Big[\alpha\beta t_i^{\alpha-1} \exp\{-\beta t_i^\alpha\} \exp\{-\lambda(s_i - t_i)\} \Big] \prod_{i=1}^{n_c} \Big[\alpha\beta {t_{ri}^*}^{\alpha-1} \exp\{-\beta {t_{ri}^*}^\alpha\} \Big] \prod_{i=1}^{n_{nr}} \Big[\alpha\beta {t_{li}^*}^{\alpha-1} \exp\{-\beta {t_{li}^*}^\alpha\} \lambda \exp\{-\lambda w_{li}^*\} \Big]. \tag{2.6}$$

Taking natural logarithm of (2.6) the complete log-likelihood function takes the form

$$\begin{aligned} l_c(\Theta | d, d^*) = {} & n \ln(\alpha) + n \ln(\beta) + n_{nr} \ln(\lambda) + (\alpha - 1) \sum_{i=1}^{n_r} \ln(t_i) - \beta \sum_{i=1}^{n_r} t_i^\alpha \\ & - \lambda \sum_{i=1}^{n_r} (s_i - t_i) + (\alpha - 1) \sum_{i=1}^{n_{nr}} \ln(t_{li}^*) - \beta \sum_{i=1}^{n_{nr}} {t_{li}^*}^\alpha - \lambda \sum_{i=1}^{n_{nr}} w_{li}^* \\ & + (\alpha - 1) \sum_{i=1}^{n_c} \ln(t_{ri}^*) - \beta \sum_{i=1}^{n_c} {t_{ri}^*}^\alpha. \end{aligned} \tag{2.7}$$

In order to apply the E step of E-M algorithm let us define the quantity $Q\Big(\Theta | \hat{\Theta}^{(m)}\Big) = E\Big[l_c(\Theta | d, d^*) | d, \hat{\Theta}^{(m)} \Big]$, where $\hat{\Theta}^{(m)}$ be the estimate of parameters at m^{th} iteration. The expectation is taken with respect to the introduced latent variables. The expression of

$Q\Big(\Theta|\hat{\Theta}^{(m)}\Big)$ can be written as follows:

$$\begin{aligned}Q\Big(\Theta|\hat{\Theta}^{(m)}\Big) &= n\ln(\alpha)+n\ln(\beta)+n_{nr}\ln(\lambda)+(\alpha-1)\sum_{i=1}^{n_r}\ln(t_i)-\beta\sum_{i=1}^{n_r}t_i^{\alpha}-\lambda\sum_{i=1}^{n_r}(s_i-t_i)\\ &+(\alpha-1)\sum_{i=1}^{n_{nr}}E\big[\ln(t_{li}^*)\mid t_{li}^*<s_i\big]-\beta\sum_{i=1}^{n_{nr}}E\big[{t_{li}^*}^{\alpha}\mid t_{li}^*<s_i\big]\\ &-\lambda\sum_{i=1}^{n_{nr}}E\big[w_{li}^*\mid w_{li}^*<(s_i-t_{li}^*)\big]+(\alpha-1)\sum_{i=1}^{n_c}E\big[\ln(t_{ri}^*)\mid t_{ri}^*>s_i\big]\\ &-\beta\sum_{i=1}^{n_c}E\big[{t_{ri}^*}^{\alpha}\mid t_{ri}^*>s_i\big]. \end{aligned}\tag{2.8}$$

Now for finding the estimators of unknown parameters, (2.8) is differentiated with respect to Θ, we get

$$\begin{aligned}\frac{\partial Q}{\partial\alpha} &= \frac{n}{\alpha}+\sum_{i}^{n_r}\ln t_i-\beta\sum_{i}^{n_r}t_i^{\alpha}\ln(t_i)+\sum_{i=1}^{n_{nr}}E\big[\ln(t_{li}^*)\mid t_{li}^*<s_i\big]-\beta\sum_{i=1}^{n_{nr}}E\big[{t_{li}^*}^{\alpha}\ln(t_{li}^*)\mid t_{li}^*<s_i\big]\\ &+\sum_{i=1}^{n_c}E\big[\ln(t_{ri}^*)\mid t_{ri}^*>s_i\big]-\beta\sum_{i=1}^{n_c}E\big[{t_{ri}^*}^{\alpha}\ln(t_{ri}^*)\mid t_{ri}^*>s_i\big],\end{aligned}\tag{2.9}$$

$$\frac{\partial Q}{\partial\beta} = \frac{n}{\beta}-\sum_{i=1}^{n_r}t_i^{\alpha}-\sum_{i=1}^{n_{nr}}E\big[{t_{li}^*}^{\alpha}\mid t_{li}^*<s_i\big]-\sum_{i=1}^{n_c}E\big[{t_{ri}^*}^{\alpha}\mid t_{ri}^*>s_i\big],\tag{2.10}$$

$$\frac{\partial Q}{\partial\lambda} = \frac{n_{nr}}{\lambda}-\sum_{i=1}^{n_r}(s_i-t_i)-\sum_{i=1}^{n_{nr}}E\big[w_{li}^*\mid w_{li}^*<(s_i-t_{li}^*)\big].\tag{2.11}$$

The expectation terms used in the E step of the E-M can be calculated by suitable choices of a, b, c & d using expression

$$E\Big[\tau^a\Big(\ln(\tau)\Big)^b\Big|\tau\in(c,d)\Big]=\frac{\int_c^d u^a\Big(\ln(u)\Big)^b f(u)du}{F(d)-F(c)}\tag{2.12}$$

The conditional densities of introduced latent variables can be obtained by using concept of truncation. Here, we define the conditional densities of these latent variables in general form for arbitrary values of parameters. The conditional density of a non-recall latent variate τ_l^* is given by

$$f_l\Big(t_l^*\big|t_l^*<s_i,\varrho,\varsigma\Big)=\frac{f(t_l^*;\varrho,\varsigma)}{1-\bar{F}(s_i;\varrho,\varsigma)}=\frac{\varrho\varsigma{t_l^*}^{\varrho-1}\exp\{-\varsigma{t_l^*}^{\varrho}\}}{1-\exp\{-\varsigma s_i^{\varrho}\}}.\tag{2.13}$$

The conditional density of a right censored latent variate t_r^* is given by

$$f_r\Big(t_r^*\big|t_r^* > s_i, \varrho, \varsigma\Big) = \frac{f(t_r^*; \varrho, \varsigma)}{\bar{F}(s_i; \varrho, \varsigma)} = \frac{\varrho\varsigma t_r^{*\varrho-1} \exp\{-\varsigma t_r^{*\varrho}\}}{\exp\{-\varsigma s_i^{\varrho}\}}. \tag{2.14}$$

Further, the conditional density of random variate w_l^* is given by

$$h_w\Big(w_l^*\big|w_l^* < (s_i - t_l^*), \omega\Big) = \frac{h(w_l^*; \omega)}{1 - \bar{H}(s_i - t_l^*; \omega)} = \frac{\omega \exp\{-\omega w_l^*\}}{1 - \exp\{-\omega(s_i - t_l^*)\}}. \tag{2.15}$$

Here ϱ and ς are the parameters of Weibull distribution whereas ω is parameter of exponential distribution.

In the M-step on the $(k+1)^{th}$ iteration of E-M algorithm, the value of expectation terms of latent variables can be computed using the conditional densities defined in (2.13) to (2.15) at obtained values of parameters from k^{th} step. Now, the expectation terms used in E-M algorithm and construction of Fisher information matrix are given as

$$\begin{aligned}
\xi_1\Big(t_{li}^*; \alpha, \beta\Big) &= E\Big[\ln(t_{li}^*)\big|t_{li}^* < s_i, \alpha, \beta\Big] = \frac{\int_0^{s_i} \ln(u)\alpha\beta u^{\alpha-1} \exp\{-\beta u^\alpha\}\, du}{1 - \exp\{-\beta s_i^\alpha\}} \\
&= \frac{I_1(s_i) - \ln(\beta)\Big(1 - \exp\{-\beta s_i^\alpha\}\Big)}{\alpha\Big(1 - \exp\{-\beta s_i^\alpha\}\Big)}, \\
\xi_2\Big(t_{li}^*; \alpha, \beta\Big) &= E\Big[t_{li}^{*\,\alpha}\big|t_{li}^* < s_i, \alpha, \beta\Big] = \frac{\int_0^{s_i} u^{2\alpha-1}\alpha\beta \exp\{-\beta u^\alpha\}\, du}{1 - \exp\{-\beta s_i^\alpha\}} \\
&= \frac{1 - (1 + \beta s_i^\alpha)\exp\{-\beta s_i^\alpha\}}{\beta\Big(1 - \exp\{-\beta s_i^\alpha\}\Big)}, \\
\xi_3\Big(t_{ri}^*; \alpha, \beta\Big) &= E\Big[\ln(t_{ri}^*)\big|t_{ri}^* > s_i, \alpha, \beta\Big] = \frac{\int_{s_i}^{\infty} \ln(u)\alpha\beta u^{\alpha-1} \exp\{-\beta u^\alpha\}\, du}{\exp\{-\beta s_i^\alpha\}} \\
&= \frac{I_2(s_i) - \ln(\beta)\Big(\exp\{-\beta s_i^\alpha\}\Big)}{\alpha \exp\{-\beta s_i^\alpha\}}, \\
\xi_4\Big(t_{ri}^*; \alpha, \beta\Big) &= E\Big[t_{ri}^{*\,\alpha}\big|t_{ri}^* > s_i, \alpha, \beta\Big] = \frac{\int_{s_i}^{\infty} u^{2\alpha-1}\alpha\beta \exp\{-\beta u^\alpha\}\, du}{\exp\{-\beta s_i^\alpha\}} = \frac{1}{\beta}\Big(1 + \beta s_i^\alpha\Big), \\
\xi_5\Big(w_{li}^*; \lambda\Big) &= E\Big[w_{li}^*\big|w_{li}^* < (s_i - t_{li}^*), \lambda\Big] = \frac{\int_0^{s_i - t_{li}^*} u\lambda \exp\{-\lambda u\}\, du}{1 - \exp\{-\lambda(s_i - t_{li}^*)\}} \\
&= \frac{1}{\lambda}\left[\frac{1 - \Big(1 + \lambda(s_i - t_{li}^*)\Big)\exp\{-\lambda(s_i - t_{li}^*)\}}{1 - \exp\{-\lambda(s_i - t_{li}^*)\}}\right],
\end{aligned}$$

$$\xi_6\Big(t^*_{li};\alpha,\beta\Big) = E\Big[{t^*_{li}}^{\alpha}\ln(t^*_{li})\big|t^*_{li}<s_i,\alpha,\beta\Big] = \frac{\int_0^{s_i}\ln(u)\alpha\beta u^{2\alpha-1}\exp\{-\beta u^{\alpha}\}\,du}{1-\exp\{-\beta s_i^{\alpha}\}}$$
$$= \frac{I_3(s_i)-\ln(\beta)\Big(1-(1+\beta s_i^{\alpha})\exp\{-\beta s_i^{\alpha}\}\Big)}{\alpha\beta\Big(1-\exp\{-\beta s_i^{\alpha}\}\Big)},$$

$$\xi_7\Big(t^*_{li};\alpha,\beta\Big) = E\Big[{t^*_{li}}^{\alpha}(\ln(t^*_{li}))^2\big|t^*_{li}<s_i,\alpha,\beta\Big] = \frac{\int_0^{s_i}\Big(\ln(u)\Big)^2\alpha\beta u^{2\alpha-1}\exp\{-\beta u^{\alpha}\}\,du}{1-\exp\{-\beta s_i^{\alpha}\}}$$
$$= \frac{I_4(s_i)+\Big(\ln(\beta)\Big)^2\Big(1-(1+\beta s_i^{\alpha})\exp\{-\beta s_i^{\alpha}\}\Big)-2\ln(\beta)I_3(s_i)}{\alpha\beta^2\Big(1-\exp\{-\beta s_i^{\alpha}\}\Big)},$$

$$\xi_8\Big(t^*_{ri};\alpha,\beta\Big) = E\Big[{t^*_{ri}}^{\alpha}\ln(t^*_{ri})\big|t^*_{ri}>s_i,\alpha,\beta\Big] = \frac{\int_{s_i}^{\infty}\ln(u)\alpha\beta u^{2\alpha-1}\exp\{-\beta u^{\alpha}\}\,du}{\exp\{-\beta s_i^{\alpha}\}}$$
$$= \frac{I_5(s_i)-\ln(\beta)\Big((1+\beta s_i^{\alpha})\exp\{-\beta s_i^{\alpha}\}\Big)}{\alpha\beta\exp\{-\beta s_i^{\alpha}\}},$$

$$\xi_9\Big(t^*_{ri};\alpha,\beta\Big) = E\Big[{t^*_{ri}}^{\alpha}\Big(\ln(t^*_{ri})\Big)^2\big|t^*_{ri}>s_i,\alpha,\beta\Big] = \frac{\int_{s_i}^{\infty}\Big(\ln(u)\Big)^2\alpha\beta u^{2\alpha-1}\exp\{-\beta u^{\alpha}\}\,du}{\exp\{-\beta s_i^{\alpha}\}}$$
$$= \frac{I_6(s_i)+\Big(\ln(\beta)\Big)^2\Big((1+\beta s_i^{\alpha})\exp\{-\beta s_i^{\alpha}\}\Big)-2\ln(\beta)I_5(s_i)}{\alpha\beta^2\exp\{-\beta s_i^{\alpha}\}},$$

where,

$$I_1(z)=\int_0^{\beta z^{\alpha}}\ln(z)\exp\{-z\}dz,\quad I_2(z)=\int_{\beta z^{\alpha}}^{\infty}\ln(z)\exp\{-z\}dz,$$
$$I_3(z)=\int_0^{\beta z^{\alpha}}z\ln(z)\exp\{-z\}dz,\quad I_4(z)=\int_0^{\beta z^{\alpha}}z\Big(\ln(z)\Big)^2\exp\{-z\}dz,$$
$$I_5(z)=\int_{\beta z^{\alpha}}^{\infty}z\ln(z)\exp\{-z\}dz,\quad I_6(z)=\int_{\beta z^{\alpha}}^{\infty}z\Big(\ln(z)\Big)^2\exp\{-z\}dz.$$

These integrals are solved using MCMC techniques. Let at the k^{th} iteration of the E-M algorithm, $\hat{\alpha}^{(k)}$, $\hat{\beta}^{(k)}$, and $\hat{\lambda}^{(k)}$ be the estimates of α, β and λ. Thus, in the M-step of the $(k+1)^{th}$ iteration of the E-M algorithm, the value of the $\hat{\alpha}^{(k+1)}$, $\hat{\beta}^{(k+1)}$, and $\hat{\lambda}^{(k+1)}$ is obtained by solving

$$\hat{\alpha}^{(k+1)}=\frac{n}{\left[\begin{array}{l}\hat{\beta}^{(k)}\sum_{i=1}^{n_r}{t_i}^{\hat{\alpha}^{(k)}}\ln(t_i)-\sum_{i=1}^{n_r}\ln(t_i)-\sum_{i=1}^{n_{nr}}\xi_1\Big(t^*_{li};\alpha^{(k)},\beta^{(k)}\Big)\\ \quad-\sum_{i=1}^{n_{nc}}\xi_3\Big(t^*_{ri};\alpha^{(k)},\beta^{(k)}\Big)+\hat{\beta}^{(k)}\Big\{\sum_{i=1}^{n_{nr}}\xi_6\Big(t^*_{li};\alpha^{(k)},\beta^{(k)}\Big)+\sum_{i=1}^{n_c}\xi_8\Big(t^*_{ri};\alpha^{(k)},\beta^{(k)}\Big)\Big\}\end{array}\right]},\qquad(2.16)$$

$$\hat{\beta}^{(k+1)} = \frac{n}{\left[\sum_{i=1}^{n_r} t_i^{\hat{\alpha}^{(k+1)}} + \sum_{i=1}^{n_{nr}} \xi_2\Big(t_{li}^*; \alpha^{(k)}, \beta^{(k)}\Big) + \sum_{i=1}^{n_c} \xi_4\Big(t_{ri}^*; \alpha^{(k)}, \beta^{(k)}\Big)\right]}, \tag{2.17}$$

$$\hat{\lambda}^{(k+1)} = \frac{n_{nr}}{\left[\sum_{i=1}^{n_r}(s_i - t_i) + \sum_{i=1}^{n_{nr}} \xi_5\Big(w_{li}^*; \lambda^{(k)}\Big)\right]}. \tag{2.18}$$

The estimates of parameters can be computed by an iterative procedure using (2.16), (2.17) and (2.18). The iterations can be terminated when $|\hat{\alpha}^{(k+1)} - \hat{\alpha}^{(k)}| + |\hat{\beta}^{(k+1)} - \hat{\beta}^{(k)}| + |\hat{\lambda}^{(k+1)} - \hat{\lambda}^{(k)}| < \epsilon$, where $\epsilon > 0$ is a sufficiently small real number.

2.2.2 Information Matrix

In the case of the recall-based data model, the usual information matrix is unable to fulfill the purpose. So, here we use the missing information principle proposed by Louis (1982) to calculate the information matrix. The following identity is used for the construction of the observed Fisher information matrix for Θ

$$I(\hat{\Theta}) = I_1 - I_2 I_2^T,$$

where matrix I_1 can be obtained by taking negative of the second derivative of 2.8 and I_2 is formed with help of the gradient vectors (2.9) to (2.11). The structure of observed Fisher information matrix $I(\hat{\Theta})$ is given by

$$I(\hat{\Theta}) = \begin{bmatrix} \frac{\partial^2 Q}{\partial \alpha^2} + \left(\frac{\partial Q}{\partial \alpha}\right)^2 & \frac{\partial^2 Q}{\partial \alpha \partial \beta} + \frac{\partial Q}{\partial \alpha}\frac{\partial Q}{\partial \beta} & \frac{\partial^2 Q}{\partial \alpha \partial \lambda} + \frac{\partial Q}{\partial \alpha}\frac{\partial Q}{\partial \lambda} \\ \frac{\partial^2 Q}{\partial \beta \partial \alpha} + \frac{\partial Q}{\partial \beta}\frac{\partial Q}{\partial \alpha} & \frac{\partial^2 Q}{\partial \beta^2} + \left(\frac{\partial Q}{\partial \beta}\right)^2 & \frac{\partial^2 Q}{\partial \beta \partial \lambda} + \frac{\partial Q}{\partial \beta}\frac{\partial Q}{\partial \lambda} \\ \frac{\partial^2 Q}{\partial \lambda \partial \alpha} + \frac{\partial Q}{\partial \lambda}\frac{\partial Q}{\partial \alpha} & \frac{\partial^2 Q}{\partial \lambda \partial \beta} + \frac{\partial Q}{\partial \lambda}\frac{\partial Q}{\partial \beta} & \frac{\partial^2 Q}{\partial \lambda^2} + \left(\frac{\partial Q}{\partial \lambda}\right)^2 \end{bmatrix}\Bigg|_{\Theta=\hat{\Theta}}$$

The expectations of second order derivatives with respect to parameters used in the information matrix are given below

$$\frac{\partial^2 Q}{\partial \alpha^2} = -\frac{n}{\alpha^2} - \sum_{i=1}^{n_r} t_i^{\alpha}\Big(\ln(t_i)\Big)^2 - \beta \sum_{i=1}^{n_{nr}} \xi_7\Big(t_{li}^*; \alpha, \beta\Big) - \beta \sum_{i=1}^{n_c} \xi_9\Big(t_{ri}^*; \alpha, \beta\Big),$$

$$\frac{\partial^2 Q}{\partial \beta^2} = -\frac{n}{\beta^2}, \quad \frac{\partial^2 Q}{\partial \lambda^2} = -\frac{n_{nr}}{\lambda^2},$$
$$\frac{\partial^2 Q}{\partial \beta \partial \alpha} = -\sum_{i=1}^{n_r} t_i^{\alpha} \ln(t_i) - \sum_{i=1}^{n_{nr}} \xi_6\Big(t_{li}^*; \alpha, \beta\Big) - \sum_{i=1}^{n_{nr}} \xi_8\Big(t_{ri}^*; \alpha, \beta\Big) = \frac{\partial^2 Q}{\partial \alpha \partial \beta}.$$

The variance-covariance matrix can be found by inverting the matrix $I(\hat{\Theta})$. The square root of diagonal elements of the variance-covariance matrix gives the standard error (SE) of $\hat{\Theta}$ which can be used in the construction of confidence interval for unknown parameters. Thus, a $100(1-\gamma)\%$ confidence interval for $\hat{\Theta}$ is given by $\hat{\Theta} \pm z_{\gamma/2} SE(\hat{\Theta})$. Here, $z_{\gamma/2}$ denotes the upper $100(\gamma/2)\%$ quantile of standard normal distribution.

2.3 Bayesian Estimation

In a classical framework, the parameters of the model are considered to be fixed but unknown. In the Bayesian paradigm, the unknown parameters are assumed to be random variables having certain distribution. This distribution expresses one's prior belief about the unknown parameters and is known as the prior distribution. The prior information is updated by using a posterior distribution which is formed with help of likelihood and prior distribution.

The posterior distribution of Θ, say $\Pi(\Theta|d)$, is written as

$$\Pi(\Theta|d) \propto L(\Theta|d)\pi(\Theta), \tag{2.19}$$

where $L(\Theta|d)$ is the likelihood function of Θ based on observed data d given in (2.3) and $\pi(\Theta)$ denotes the joint prior distribution of Θ. Thus, the expression of likelihood in (2.3) indicates that it is difficult to study the properties of $\Pi(\Theta|d)$ analytically. The purpose can be solved by generating samples from the posterior distribution, but at this stage, it also seems a tedious task. So, to overcome this issue, we try to augment the data d with help of latent variables to form the complete data. From the classical section, d^* is related to imputed time to event of the non-recall and right-censored observation. For details about the augmented approach, one can see Roy et al. (2017).

The discussed augmentation technique helps us to convert the problem of generating samples from $\Pi(\Theta|d)$ from $\Pi(\Theta|d, d^*)$, where symbol $\Pi(\Theta|d, d^*)$ represents the joint distribution of D, D^* and Θ. Thus, sample generation from the $\Pi(\Theta|d, d^*)$ can be done using the Gibbs sampling technique, which is completed in two steps: the first step includes generation of d^* and then in the second step we generate the observations from $\Pi(\Theta|d, d^*)$.

2.3.1 Choice of Priors and Posterior Analysis

The choice of the prior distribution is an important issue and should be dealt with carefully. The information about the prior may be obtained from an expert's opinion or from the literature. Prior based on an individual's belief is known as subjective prior. Since we have assumed the time to event distribution as Weibull, a suitable choice is done based on the literature review. As stated by Soland (1969), for two-parameter Weibull distribution, the conjugate family of prior distributions does not exist. So, some mixed types of priors are suggested. A discrete type prior to the shape parameter and continuous type prior to the scale parameter is suggested by Soland (1969). Later the use of continuous priors, uniform for the shape parameter and inverted gamma prior to the scale parameter is suggested by Tsokos (1972). In article Kundu (2008), for scale parameter a gamma prior is taken and for shape parameter, no specific prior is chosen, and it is assumed that the support of the shape parameter is $(0, \infty)$ having log-concave density. Banerjee and Kundu (2008) suggested gamma distribution as priors for the shape and scale parameters.

Let us assume that prior distributions for the parameters α, β and λ are gamma distributed, i.e. $\mathcal{G}(\mu_1, \nu_1)$, $\mathcal{G}(\mu_2, \nu_2)$ and $\mathcal{G}(\mu_3, \nu_3)$. Therefore, under independent assumption, we can write the joint prior density of α, β and λ as

$$\pi(\Theta) \propto \alpha^{\mu_1-1}\beta^{\mu_2-1}\lambda^{\mu_3-1}\exp\{-(\nu_1\alpha + \nu_2\beta + \nu_3\lambda)\}; \quad \mu_j, \nu_j > 0, j = 1, 2, 3. \tag{2.20}$$

Now the posterior density of Θ are obtained by multiplying complete likelihood in (2.6) and

joint prior density in (2.20) and written as

$$\Pi(\Theta|d,d^*) \propto \alpha^{n+\mu_1-1}\beta^{n+\mu_2-1}\lambda^{n_{nr}+\mu_3-1}\left(\prod_{i=1}^{n_r} t_i^{\alpha-1}\right)\left(\prod_{i=1}^{n_{nr}} t_{li}^{*\,\alpha-1}\right)\left(\prod_{i=1}^{n_c} t_{ri}^{*\,\alpha-1}\right)$$
$$\exp\left\{-\beta\Big(\sum_{i=1}^{n_r} t_i^{\alpha}+\sum_{i=1}^{n_{nr}} t_{li}^{*\,\alpha}+\sum_{i=1}^{n_c} t_{ri}^{*\,\alpha}\Big)\right\}\exp\left\{-\lambda\Big(\sum_{i=1}^{n_r}(s_i-t_i)+\sum_{i=1}^{n_{nr}} w_{li}^{*}\Big)\right\}$$
$$\exp\left\{-\Big(\nu_1\alpha+\nu_2\beta+\nu_3\lambda\Big)\right\}. \tag{2.21}$$

From (2.21), the full conditionals of α, β and λ can be written as

$$\pi_1(\alpha|\beta) \propto \alpha^{n+\mu_1-1}\left(\prod_{i=1}^{n_r} t_i^{\alpha-1}\right)\left(\prod_{i=1}^{n_{nr}} t_{li}^{*\,\alpha-1}\right)\left(\prod_{i=1}^{n_c} t_{ri}^{*\,\alpha-1}\right)\exp\left\{-\beta\Big(\sum_{i=1}^{n_r} t_i^{\alpha}+\sum_{i=1}^{n_{nr}} t_{li}^{*\,\alpha}+\sum_{i=1}^{n_c} t_{ri}^{*\,\alpha}+\nu_1\Big)\right\}, \tag{2.22}$$

$$\pi_2(\beta|\alpha) \sim \mathcal{G}\left(n+\mu_2,\ \sum_{i=1}^{n_r} t_i^{\alpha}+\sum_{i=1}^{n_{nr}} t_{li}^{*\,\alpha}+\sum_{i=1}^{n_c} t_{ri}^{*\,\alpha}+\nu_2\right), \tag{2.23}$$

$$\pi_3(\lambda) \sim \mathcal{G}\left(n_{nr}+\mu_3,\ \sum_{i=1}^{n_r}(s_i-t_i)+\sum_{i=1}^{n_{nr}} w_{li}^{*}+\nu_3\right). \tag{2.24}$$

For given values of hyper-parameters μ_j, ν_j $(j=1,2,3)$ and latent variables t_{li}^*, w_{li}^* & t_{ri}^*, the posterior samples are generated by using Markov Chain Monte Carlo (MCMC) techniques: Gibbs Sampling and Metropolis-Hastings (M-H) algorithms [Metropolis and Ulam (1949), Hastings (1970)].

2.3.2 Data Augmentation Algorithm

The sample generation using the data augmentation approach is presented in the form of an algorithm. The steps of algorithm are given below:

1. Set the initial values of parameter vector as $\Theta^{(0)} = \Big(\alpha^{(0)}, \beta^{(0)}, \lambda^{(0)}\Big)$.

2. Based on $\Theta^{(0)}$ and observed data d, generate values on latent variables t_{li}^* and t_{ri}^* using conditional density given in (2.13) and (2.14) as

$$t_{li}^* = \left[-\frac{1}{\beta}\ln\left\{1-u_i\Big(1-\exp\{-\beta s_i^{\alpha}\}\Big)\right\}\right]^{1/\alpha};\quad i=1,2,\ldots,n_{nr} \tag{2.25}$$

and

$$t^*_{ri} = \left[s_i^{\alpha} - \frac{1}{\beta} \ln (1-u) \right]^{1/\alpha} ; \quad i = 1, 2, \ldots, n_c, \tag{2.26}$$

where $u \sim \mathcal{U}(0,1)$.

3. Next, for given values of $\Theta^{(0)}$, d and t^*_{li}, generate observations on w^*_{li} using density (2.15) by

$$w^*_{li} = -\frac{1}{\lambda} \ln \left[1 - u_i \Big(1 - \exp\{-\lambda(s_i - t^*_{li})\} \Big) \right]; \quad i = 1, 2, \ldots, n_{nr}. \tag{2.27}$$

4. Given values of $\Theta^{(0)}$, d, t^*_{li}, t^*_{ri} μ_1 and ν_1, we generate samples on $\alpha^{(1)}$ using (2.22) using M-H algorithm under the the proposal as normal distribution.

5. For given values of d, t^*_{li}, t^*_{ri}, μ_2, ν_2 and $\alpha^{(1)}$, generate samples on $\beta^{(1)}$ using (2.23).

6. Finally, for given values of d, μ_3, ν_3 and w^*_{li}, generate samples on $\lambda^{(1)}$ using (2.24).

Now, the current state is $\Theta^{(1)} = \left(\alpha^{(1)}, \beta^{(1)}, \lambda^{(1)} \right)$. Iterating these steps, we get a long chain as $\left(\Theta^{(0)}, \Theta^{(1)}, \cdots, \Theta^{(N)} \right)$. Discarding a few initial observations as burn-in samples, checking the convergence of MCMC chains using cumsum and ACF plots and obtaining a proper thinning interval to avoid auto-correlation, we get the final reduced chain of size $N'(< N)$. All the posterior inferences are be done based on these MCMC samples. The highest posterior density intervals are obtained by using the method of Chen and Shao (1999).

2.4 Some Statistics Based on Monitoring Time

In the current status data problem, we expect that under different constraints like cost and time, the number of occurrences of events in an experiment should be sufficiently large. This can be achieved by choosing a suitable pattern of the monitoring time which will be able to capture the pattern of time to event distribution. As there is a possibility of non-recall observations in the study, we try to keep the monitoring time close to the time to event to

reduce the non-recall. We are interested in a monitoring pattern/distribution which provides the optimum number of occurrences of events. Thus the pattern/distribution of monitoring time plays an important role for inferential purposes in such studies. Here we have derived some statistics such as probability of experiencing an event, expected time in study (ETS), total time in study (TTS), expected number of recall and non-recall observations which are useful in choosing a suitable monitoring time distribution.

2.4.1 Probability of Experiencing an Event

We define the probability of experiencing an event by an individual as

$$\begin{aligned} p &= P\Big[T \leq S\Big] = \int_{-\infty}^{\infty} P\Big[T \leq S | S = s\Big] g_S(s) ds \\ &= \int_{-\infty}^{\infty} \left[\int_0^s f_T(t) dt\right] g_S(s) ds = \int_{-\infty}^{\infty} F_T(s) g_S(s) ds. \end{aligned} \tag{2.28}$$

For particular distribution of time to event and monitoring time, we can calculate the quantity defined above. We consider time to event following Weibull, i.e. $T \sim \mathcal{W}(\alpha, \beta)$ and the distribution of monitoring times as uniform and exponential.

First, let us consider distribution of monitoring time as uniform, i.e. $S \sim \mathcal{U}\big[\kappa, \tau\big]$ and using (2.28), the probability of experiencing an event is given as

$$\begin{aligned} p &= 1 - \frac{1}{(\tau - \kappa)} \int_{\kappa}^{\tau} \exp\{-\beta u^{\alpha}\} du \\ &= 1 - \frac{\beta^{\alpha}}{\alpha(\tau - \kappa)} \left[\Gamma\left(\frac{1}{\alpha}, \beta\kappa^{\alpha}\right) - \Gamma\left(\frac{1}{\alpha}, \beta\tau^{\alpha}\right)\right], \end{aligned} \tag{2.29}$$

where, $\Gamma(q, z) = \int_z^{\infty} u^{q-1} \exp\{-u\}\, du$ is an upper incomplete gamma function.

Next, let us consider monitoring pattern as exponential, i.e. $S \sim \mathcal{E}(\zeta)$. Using (2.28), the probability of experiencing an event is given as

$$\begin{aligned} p &= \int_0^{\infty} \zeta \Big[1 - \exp\{-\beta u^{\alpha}\}\Big] \exp\{-\zeta u\} du \\ &= 1 - \zeta \int_0^{\infty} \exp\{-\beta u^{\alpha} - \zeta u\} du. \end{aligned} \tag{2.30}$$

The numerical values of these probabilities can be calculated for the given values of parameters and sample size n.

2.4.2 Expected Time in Study

Researchers are always keen to know the expected time in the study in an experiment/study. It can play an important role in planning strategies of the experiment/study in advance. So, we calculated the quantity ETS for our setup. Since a monitoring time point is attached to each individual, here we are assuming that the observed data will be minimum of time to event (T) and monitoring time (S). That is for i^{th} individual in the study, we observe the data as $X_i = min(T_i, S_i)$. At any monitoring time, the ETS can be found with the help of the distribution function of $X = max(X_1, X_2, \cdots, X_n)$. The CDF of X is given by

$$F_X(x) = P(X \leq x) = P\Big[max(X_1, X_2, \cdots, X_n) \leq x\Big] = \Big[P(X_1 \leq x)\Big]^n; \quad x > 0.$$

Since $X_i,\ i = 1, 2, ..., n$ are $i.i.d.$ random variables, so we have

$$\begin{aligned} P\Big[X_i \leq x\Big] &= P\Big[min\,(T_i, S_i) \leq x\Big] = 1 - P\Big[min(T_i, S_i) > x\Big] \\ &= 1 - P\Big[T_i > x\Big]P\Big[S_i > x\Big] \qquad (T_i,\ S_i\ are\ independent) \\ &= 1 - \bar{F}_T(x)\bar{G}_S(x). \end{aligned}$$

Using this expression calculated just above, we get the CDF of X as follows

$$F_X(x) = \Big[1 - \bar{F}_T(x)\bar{G}_S(x)\Big]^n. \tag{2.31}$$

Now, the ETS can be calculated by using expression

$$ETS = E(X) = \int_0^\infty \Big[1 - F_X(x)\Big]dx. \tag{2.32}$$

Assuming same distributions for the time to event and monitoring patterns as in subsection 2.4.1, we can calculate the ETS using expression (2.32). In first case, the ETS for an

individual is given by

$$ETS = \int_{\kappa}^{\tau} \left[1 - \left\{1 - \frac{\tau - u}{\tau - \kappa} \exp\{-\beta u^{\alpha}\}\right\}^{n}\right] du. \tag{2.33}$$

In second case, the ETS can be calculated as

$$ETS = \int_{0}^{\infty} \left[1 - \left\{1 - \exp\{-\beta u^{\alpha} - \zeta u\}\right\}^{n}\right] du. \tag{2.34}$$

The numerical values of ETS can be calculated for the given values of the parameters and sample size n.

2.4.3 Total Time in Study

The concept of total time on test (TTT) given by Barlow and Campo (1975), plays a very important role in the area of reliability theory to identify the characteristics of failure rate. Subsequently, TTT is utilized in other areas such as economics, risk management and maintenance schedule. The TTT statistic is given by the total sum of time spent by each unit in the experiment (exact failure time or censored time). Utilizing the same concept for our setup, we defined the quantity TTS. For any individual inspected at any monitoring time point, there may be two possibilities, i.e. the individual may fall into the left-censored category or in the right-censored category. For right-censored observation, it is known that the individual has not experienced the event at the monitoring time. Further, if it belongs to left-censored, the time to event may belong to either recall or non-recall. If the individual is able to recall the time to the event, the exact time (T) is known but for the case of non-recall, we have introduced the latent time T_l^* earlier. Based on these quantities, we can calculate the TTS for an individual, and finally summing all the times, we get a value of TTS. For i^{th} individual, let us define the TTS as

$$(TTS)_i = min(T_i, T_{li}^*, S_i) \tag{2.35}$$

If there are n individuals in the study, then TTS can be given as

$$TTS = \sum_{i=1}^{n} (TTS)_i \tag{2.36}$$

2.4.4 Expected Number of Recall and Non-Recall

In this subsection, we find out the mathematical expressions for calculating the expected number of recall and non-recall observations for a given sample size n. For this, we concentrate on the indicator ϵ_i which takes values 1 and 0 corresponding to any individual, i.e. it seems like a Bernoulli random variate but with varying recall probabilities for each individual and classifies them into either recall or non-recall category. Therefore, the estimated number of recall observations based on sample size n can be found by multiplying n into the probability of recall, which is a random variable and a function of $(S-T)$. Hence, first, we proceed by finding the distribution of difference of $(S-T)$, and then based on it we will be able to find the estimated value of the probability of recall and then multiplying with n, we get the expected number of recall.

The expected value of exactly recalled observations can be given by $E\big(\epsilon\big) = E\big(\sum_{i=1}^{n} \epsilon_i\big) = \sum_{i=1}^{n} E(\epsilon_i) = n\hat{\psi}(z;\lambda)$.

Since the distribution of T and S are assumed to be independent, thus their joint density $\big($say, h(.)$\big)$ can be written as the product of two densities.

$$h(t,s) = f(t)g(s)$$

Now, let us make the transformation $Z = S - T$ which implies $S = T + Z$ and $T = t$. The Jacobian of transformation comes out to be 1. Now, writing the joint density in terms of transformed random variables and marginalizing with respect to T, we get the density of Z as

$$h(z) = \int_0^\infty f(t)g(t+z)dt$$

Now, the estimated value of recall probability can be calculated as

$$\hat{\psi}(z;\lambda) = E\left[\exp\{-\lambda z\}\right] = \int_0^\infty \exp\{-\lambda z\} h(z) dz \tag{2.37}$$

The expected value of recall and non-recall observations based on sample size n can be calculated by using the expressions $n\hat{\psi}(z;\lambda)$ and $n\left[1-\hat{\psi}(z;\lambda)\right]$. The estimated probability of recall is calculated using (2.37) under two cases as below:
First, let us assume that $T \sim \mathcal{W}(\alpha,\beta)$ and $S \sim \mathcal{U}\Big(\kappa,\tau\Big)$. Since T and S are independent, the joint density function in this case can be written as

$$h(t,s;\alpha,\beta,\kappa,\tau) = \frac{\alpha\beta}{\tau-\kappa} t^{(\alpha-1)} \exp\{-\beta t^\alpha\}$$

Now make the transformation $Z = (S-T)$ which implies $S = Z+T$ and $T = t$. The Jacobian of the transformation comes out to 1. The density of Z can be obtained by marginalizing with respect to T, can be written as

$$h(z;\alpha,\beta,\kappa,\tau) = \frac{\alpha\beta}{\tau-\kappa} \int_0^\infty t^{(\alpha-1)} \exp\{-\beta t^\alpha\} dt$$

The expected value of recall probability can be calculated by using the above density as

$$\hat{\psi}(z;\lambda) = E\left[\exp\{-\lambda z\}\right] = \int_0^\infty \exp\{-\lambda z\} h(z;\alpha,\beta,\kappa,\tau) dz$$

Next, let us assume $T \sim \mathcal{W}(\alpha,\beta)$ and $S \sim \mathcal{E}(\zeta)$. Assuming T and S as independent, the joint density can be written as

$$h(t,s;\alpha,\beta,\lambda) = \alpha\beta\zeta t^{(\alpha-1)} \exp\{-\beta t^\alpha - \zeta s\}$$

Now making the same transformation as earlier, $Z = (S-T)$ implies $S = Z+T$ and $T = t$. The Jacobian of the transformation comes out to 1. The density of Z can be obtained by

marginalizing with respect to T, can be written as

$$h(z;\alpha,\beta,\zeta) = \alpha\beta\zeta \exp\{-\zeta z\} \int_0^\infty t^{(\alpha-1)} \exp\{-\beta t^\alpha - \zeta t\} dt$$

The expected value of recall probability can be calculated by using the above density as

$$\hat{\psi}(z;\lambda) = E\left[\exp\{-\lambda z\}\right] = \int_0^\infty \exp\{-\lambda z\} h(z;\alpha,\beta,\zeta) dz$$

Finally, to calculate the derived statistics numerically, we take the values of $n = 50$, $\alpha = 1.15$ and $\beta = 0.85$. The symbol n_{nc} denotes the number of non-censored observations. To explore the monitoring time distribution, five choices of the monitoring patterns are considered as

S_1. $S \sim \mathcal{U}(Q_1, Q3)$,

S_2. $S \sim \mathcal{U}(Q_2, Q_3)$,

S_3. $S \sim \mathcal{U}(P_{75}, P_{90})$,

S_4. $S \sim \mathcal{E}\{Q_2\}$,

S_5. $S \sim \mathcal{E}\{Q_3\}$.

where Q_1, Q_2 & Q_3 are 1^{st}, 2^{nd} & 3^{rd} quartiles and P_{75} & P_{90} are 75^{th} & 90^{th} percentiles of time to event distribution T respectively. Further, to introduce the non-recall proportion in data, we have taken eight different values of the non-recall parameter λ.

Table 2.1: Calculated quantities under various patterns of monitoring time with various values non-recall parameter

					λ															
					0	0.10		0.20		0.40		0.55		0.65		0.75		1.00	
	p	n_{nc}	n_c	*ETS*	*TTS*	n_{nr}	*TTS*	n_{nr}	*TTS*	n_{nr}	*TTS*	n_{nr}	*TTS*	n_{nr}	*TTS*	n_{nr}	*TTS*	n_{nr}	*TTS*
S_1	0.67	33.32	16.68	1.45	33.02	1.51	34.80	2.74	36.41	5.24	39.31	6.76	41.21	7.82	42.60	8.58	43.47	10.77	46.44
S_2	0.74	36.95	13.05	1.48	38.03	1.97	40.06	3.94	42.32	6.98	45.02	9.10	47.15	10.36	48.50	11.52	49.63	14.22	52.42
S_3	0.88	44.08	5.92	2.25	47.82	4.50	48.93	8.46	49.73	14.69	50.65	18.53	51.48	20.56	51.73	22.51	52.10	26.06	52.72
S_4	0.50	25.01	24.99	2.45	29.49	2.71	29.66	4.89	29.89	8.25	31.06	10.28	31.62	11.11	31.89	12.07	22.40	13.81	33.61
S_5	0.34	17.16	32.84	2.45	21.47	1.09	22.29	2.06	23.29	3.70	24.86	4.71	26.12	5.49	26.88	5.86	27.41	7.24	29.33

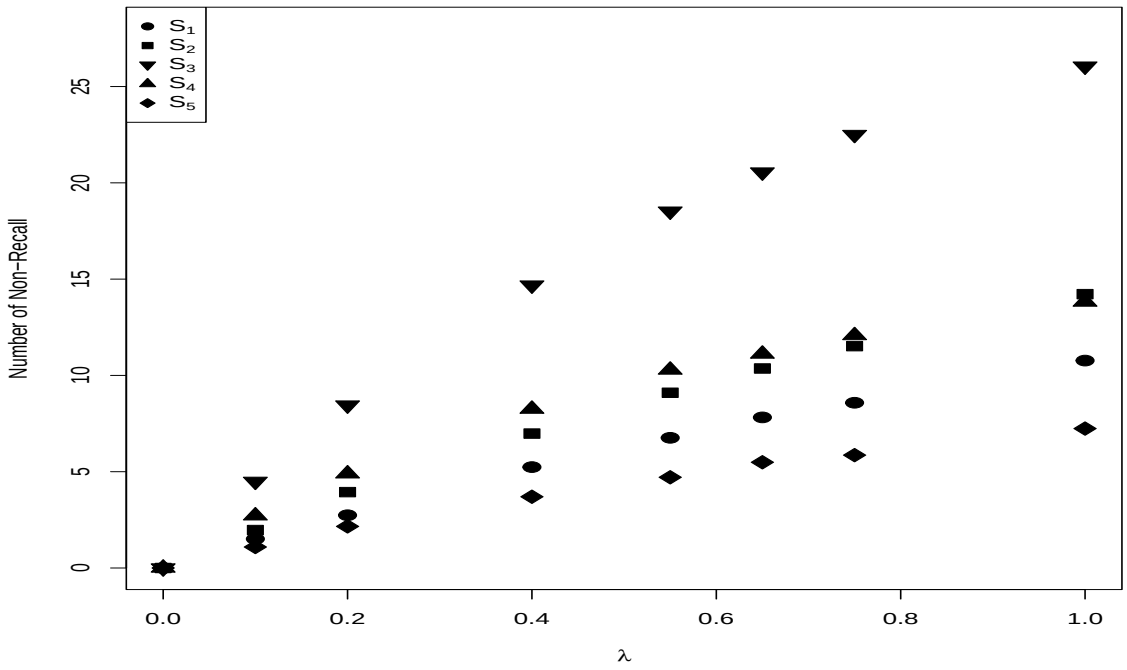

Figure 2.1: Number of non-recall observations under different monitoring patterns with varying λ.

From Table 2.1, we can see that the probability of experiencing the event and the corresponding number of non-censored observations under uniform monitoring is greater than that of exponential monitoring times. Also, as the value of the non-recall parameter λ increases, there is an increase in the number of non-recall observations as well as TTS. Figure 2.1 shows that the number of non-recall observations increases as the value of λ increases.

2.5 Simulation Study

In order to investigate the finite sample properties of estimators, we carry out the simulation study under both the classical and Bayesian frameworks. **R** software is used to carry out the simulation study. The time to event (T) of interest is assumed to follow the Weibull distribution with the shape parameter α and the scale β. Under the current status data assumption, each time to event value is monitored at a single time point (S). We consider uniform and exponential as monitoring time distributions in the simulation study. All estimates are obtained for both the monitoring patterns.

Let us set the initial values of unknown parameters $\alpha = 1.15$ and $\beta = 0.85$. For uniform

monitoring point, S is generated from uniform taking minimum and maximum of T as parameters. The non-recall probability parameter is chosen arbitrary as $\lambda = \{0.10, 0.15\}$. Corresponding to the first value of λ, we have approximately 15% observations in the non-recall category while 21% observations are right-censored. Corresponding to the second value of λ, the proportions of non-recall and right-censored are observed as 20% and 21% respectively. For the exponential monitoring pattern, S is generated with a rate of 0.2. Under exponential monitoring pattern, approximately 16% and 18% observations are under non-recall and right-censored categories for $\lambda = 0.05$. For $\lambda = 0.05$, we get the non-recall and right-censored observations at 27% and 18% levels. Based on the above setup, under both monitoring patterns, $n = (50, 150, 250)$ observations are generated and recall (n_r), non-recall (n_{nr}) and right-censored (n_c) observations are noted. For Bayesian inference under informative prior, hyper-parameters are chosen using moment matching criteria. For moment matching criteria, we match the mean and variances of MLEs (based on say N samples) with mean and variances of considered prior distribution and find out the values of hyper-parameters.

For varying sample sizes and prefixed set of parameter values, mean square error (MSE) and absolute bias (AB) in the classical and Bayesian frameworks are reported in Table 2.2 and Table 2.3. Average lengths (ALs), shape and coverage probabilities (CPs) for interval estimates obtained with the help of missing information principle, say asymptotic confidence intervals (ACIs) and highest posterior density (HPD) intervals based on MCMC samples and method proposed by Chen and Shao (1999) are reported in Table 2.4 and Table 2.5. The simulation process is replicated 1000 times to average out the sampling fluctuations. The shape is a measure of the symmetry of interval estimates, which is defined by the expression given below:

$$Shape = \frac{\hat{\eta}_{UL} - \hat{\eta}}{\hat{\eta} - \hat{\eta}_{LL}},$$

the symbol $\hat{\eta}$ represents the point estimate of parameter η; $\hat{\eta}_{LL}$ and $\hat{\eta}_{UL}$ are estimated the lower limit and estimated upper limit of intervals of the parameter η. When the value of shape is 1, it indicates the symmetry of intervals towards MLE as in the case of asymptotic confidence interval estimate. While the values of shape greater and less than 1 indicate

positive skewness and negative skewness respectively for the distribution of MLE.

Table 2.2: Mean square error and absolute bias for ML and Bayes methods under uniform monitoring points for varying sample sizes n.

			ML		$Bayes$	
λ	n		MSE	AB	MSE	AB
		$\hat{\alpha}$	0.0351	0.1432	0.0194	0.1050
	50	$\hat{\beta}$	0.0231	0.1153	0.0201	0.1087
		$\hat{\lambda}$	0.0018	0.0337	0.0012	0.0278
		$\hat{\alpha}$	0.0089	0.0747	0.0044	0.0515
0.10	150	$\hat{\beta}$	0.0085	0.0744	0.0078	0.0719
		$\hat{\lambda}$	0.0005	0.0177	0.0005	0.0175
		$\hat{\alpha}$	0.0054	0.0587	0.0025	0.0391
	250	$\hat{\beta}$	0.0057	0.0621	0.0053	0.0600
		$\hat{\lambda}$	0.0003	0.0132	0.0003	0.0120
		$\hat{\alpha}$	0.0351	0.1435	0.0205	0.1071
	50	$\hat{\beta}$	0.0197	0.1112	0.0173	0.1050
		$\hat{\lambda}$	0.0032	0.0436	0.0026	0.0391
		$\hat{\alpha}$	0.0119	0.0853	0.0056	0.0574
0.15	150	$\hat{\beta}$	0.0084	0.0749	0.0077	0.0720
		$\hat{\lambda}$	0.0007	0.0218	0.0006	0.0208
		$\hat{\alpha}$	0.0070	0.0661	0.0029	0.0418
	250	$\hat{\beta}$	0.0085	0.0765	0.0075	0.0720
		$\hat{\lambda}$	0.0005	0.0167	0.0004	0.0147

Simulated results under uniform monitoring time points are given in Table 2.2 and Table 2.4. Outputs for exponential monitoring points are reported in Table 2.3 and Table 2.5. From the simulated results given in Table 2.2 - Table 2.5, we can conclude that values of MSE and AB for parameters decrease as sample size increases, which validates the consistency of estimators. In Table 2.4 and Table 2.5, the average length of ACIs and HPD intervals decreases as sample size increases. The coverage probability indicates the proportion of time

Table 2.3: Mean square error and absolute bias for ML and Bayes methods under exponential monitoring points for varying sample sizes n.

			ML		$Bayes$	
λ	n		MSE	AB	MSE	AB
		$\hat{\alpha}$	0.0348	0.1418	0.0106	0.0785
	50	$\hat{\beta}$	0.0268	0.1280	0.0223	0.1170
		$\hat{\lambda}$	0.0003	0.0149	0.0002	0.0131
		$\hat{\alpha}$	0.0096	0.0767	0.0029	0.0422
0.05	150	$\hat{\beta}$	0.0080	0.0716	0.0072	0.0679
		$\hat{\lambda}$	0.0004	0.0079	0.0001	0.0069
		$\hat{\alpha}$	0.0052	0.0564	0.0016	0.0309
	250	$\hat{\beta}$	0.0051	0.0577	0.0047	0.0551
		$\hat{\lambda}$	0.0002	0.0060	0.0001	0.0059
		$\hat{\alpha}$	0.0439	0.1574	0.0101	0.0771
	50	$\hat{\beta}$	0.0290	0.1360	0.0228	0.1219
		$\hat{\lambda}$	0.0008	0.0221	0.0007	0.0219
		$\hat{\alpha}$	0.0100	0.0787	0.0025	0.0390
0.10	150	$\hat{\beta}$	0.0098	0.0804	0.0086	0.0761
		$\hat{\lambda}$	0.0003	0.0132	0.0003	0.0123
		$\hat{\alpha}$	0.0062	0.0631	0.0016	0.0315
	250	$\hat{\beta}$	0.0066	0.0661	0.0059	0.0626
		$\hat{\lambda}$	0.0002	0.0099	0.0002	0.0070

Table 2.4: Average length, shape and coverage probability for ML and Bayes methods under uniform monitoring points for varying sample sizes n.

			ACI		*HPD*		
λ	n		*AL*	*CP*	*AL*	*Shape*	*CP*
		$\hat{\alpha}$	0.8286	0.9844	0.5382	1.0229	0.9656
	50	$\hat{\beta}$	0.7614	0.9531	0.5696	1.0547	0.9578
		$\hat{\lambda}$	0.1680	0.9188	0.1453	1.1442	0.9656
		$\hat{\alpha}$	0.5675	0.9837	0.2932	1.0126	0.9817
0.10	150	$\hat{\beta}$	0.3661	0.9370	0.3197	1.0324	0.9126
		$\hat{\lambda}$	0.0838	0.9248	0.0834	1.0711	0.9431
		$\hat{\alpha}$	0.3931	0.9960	0.2216	1.0093	0.9820
	250	$\hat{\beta}$	0.2672	0.9319	0.2484	1.0283	0.9158
		$\hat{\lambda}$	0.0626	0.9238	0.0620	1.0572	0.9339
		$\hat{\alpha}$	0.8918	0.9830	0.5553	1.0252	0.9722
	50	$\hat{\beta}$	0.8047	0.9707	0.5668	1.0552	0.9599
		$\hat{\lambda}$	0.2126	0.9259	0.1948	1.1224	0.9475
		$\hat{\alpha}$	0.5562	0.9757	0.3009	1.0123	0.9594
0.15	150	$\hat{\beta}$	0.3583	0.9412	0.3252	1.0367	0.9270
		$\hat{\lambda}$	0.1059	0.9351	0.1046	1.0569	0.9533
		$\hat{\alpha}$	0.3800	0.9749	0.2260	1.0078	0.9724
	250	$\hat{\beta}$	0.2583	0.8333	0.2503	1.0273	0.8308
		$\hat{\lambda}$	0.0778	0.9311	0.0768	1.0440	0.9311

Table 2.5: Average length, shape and coverage probability for ML and Bayes methods under exponential monitoring points for varying sample sizes n.

			ACI		*HPD*		
λ	n		*AL*	*CP*	*AL*	*Shape*	*CP*
		$\hat{\alpha}$	0.7085	0.9406	0.4562	1.0169	0.9823
	50	$\hat{\beta}$	0.5614	0.9039	0.5564	1.0573	0.9292
		$\hat{\lambda}$	0.0673	0.9001	0.0620	1.1218	0.9330
		$\hat{\alpha}$	0.3911	0.9575	0.2593	1.0093	0.9900
0.05	150	$\hat{\beta}$	0.3161	0.9188	0.3028	1.0309	0.9463
		$\hat{\lambda}$	0.0390	0.9363	0.0389	1.0651	0.9500
		$\hat{\alpha}$	0.2986	0.9617	0.1996	1.0089	0.9883
	250	$\hat{\beta}$	0.2432	0.9017	0.2154	1.0221	0.9267
		$\hat{\lambda}$	0.0301	0.9450	0.0301	1.0533	0.9450
		$\hat{\alpha}$	0.6923	0.8976	0.4726	1.0143	0.9912
	50	$\hat{\beta}$	0.5441	0.8736	0.5391	1.0627	0.9254
		$\hat{\lambda}$	0.1058	0.9216	0.1027	1.0956	0.9406
		$\hat{\alpha}$	0.3799	0.9438	0.2672	1.0118	0.9950
0.10	150	$\hat{\beta}$	0.3082	0.8725	0.3044	1.0343	0.9188
		$\hat{\lambda}$	0.0602	0.9313	0.0576	1.0527	0.9375
		$\hat{\alpha}$	0.2935	0.9417	0.2069	1.0083	0.9983
	250	$\hat{\beta}$	0.2380	0.8533	0.2267	1.0284	0.8983
		$\hat{\lambda}$	0.0464	0.9200	0.0450	1.0446	0.9317

the interval contains the true value of the parameter and is found to be satisfactory. Based on the shape for HPD intervals, we can conclude the intervals become symmetric as the sample size increases. Also, as we increase the proportion of missing data in the model, results deteriorate which results in an increase in MSEs, ABs and ALs. All the results are found satisfactory under both monitoring patterns.

2.6 Real Data Study

2.6.1 Description of Data

In the current chapter, we are interested to estimate the duration of breastfeeding of a recent child. The data used in this section is taken from the fourth round National Family Health Survey (NFHS) conducted during the year $2015-16$ in India. Nath et al. (1994) studied the joint effect of Birth spacing, breastfeeding, and child mortality in a traditional Indian society using a hazards model analysis. Singh et al. (2007) studied the effect of breastfeeding and maternal health care program on infant mortality. Kumar et al. (2020) used current status, life-table and Kaplan-Meir approaches to estimate the duration of breastfeeding using NFHS-IV data. It is seen that there is a difference between the duration of breastfeeding of male and girl children. The NFHS-IV data contains information on several variables related to mothers and children. Here our interest is to estimate the mean and median age of duration of breastfeeding based on this data. From this data set, variables of interest like the date of birth of the child, the status of breastfeeding and duration of breastfeeding are used for analysis purposes. The data is based on the responses of mothers related to breastfeeding of current child and are recorded at different levels: (i) ever breastfed, but not currently doing (ii) still breastfeeding (iii) never breastfed (iv) breastfed until the child died (v) inconsistent and (vi) don't know. Thus, if the observation is exactly recalled we get the precise duration of the breastfeeding. The observation under the don't know the category is considered as a non-recall observation and for those observations, it is only known that event had been experienced before the interview/monitoring time. If the mother answers that she is still breastfeeding at the time of the interview, then it is categorized in the right-censored

category. Here, for study purposes, we consider the age of the child as monitoring time (S) and duration of breastfeeding (T) are recorded in months based on the mother's responses.

2.6.2 Fitting of Data

In this subsection, we discuss the suitability of considered time to event distribution with menarche data. To show that the considered time to event distribution is suitable for menarche data, we followed the approach based on modified chi-square test statistic proposed by Hope (1968) and recently used in the article Koley and Dewanji (2021). We define the modified chi-square statistic under recall-based data as

$$\chi^2_M = \sum_{i=1}^{n} \frac{(\delta_i \epsilon_i - p_i^*)^2}{p_i^* (1 - p_i^*)},$$

where p_i^* takes different values from the likelihood as per indicator variables defined below

$$p_i^* = \begin{cases} f(t_i;\theta)\exp\{-\lambda(s_i - t_i)\}; & \text{if} \quad \delta_i = 1 \ \& \ \epsilon_i = 1, \\ \int_0^{s_i} f(u;\theta)\Big(1 - \exp\{-\lambda(s_i - u)\}\Big)du; & \text{if} \quad \delta_i = 1 \ \& \ \epsilon_i = 0, \\ \bar{F}(s_i;\theta); & \text{if} \quad \delta_i = 0. \end{cases}$$

2.6.3 Estimates of Data

We consider 1,44,485 observations based on birth of child prior to 60 months of the survey. In the final sample, the number of exact recall (n_r) observations are $62,983$, non-recall observations (n_{nr}) are 386 and censored observations (n_c) are $81,116$ respectively. Using the modified chi-square test statistic, the p value obtained for breastfeeding data is 0.78, which indicates that Weibull distribution is suitable for considered data.

The present study focuses on modeling recall-based data. In the considered breastfeeding data, the proportion of non-recall observations is very small. So, if we remove these observations from the data, then the results are not affected too much. But for small sample size studies, these observations may give valuable information and we can't ignore them. Yamane (1967) gave an expression to calculate the representative sample size for a given

population with some margin of error. So, considering the whole sample as a population, we calculated the sample size with a 5% margin of error and got the effective sample of size 400.

To solve our purpose, we consider the three categories recall, non-recall and right-censored as strata and draw samples from each of the strata randomly to maintain different proportions of non-recall in the representative sample. Based on the discussed approach, we take a sample of size 500 in which the proportion of right-censored observations is fixed at 50%. For the remaining 50% of non-censored observations, we fixed the recall and non-recall proportions at three levels.

1. Level I: $n_r = 45\%$, $n_{nr} = 5\%$ and $n_c = 50\%$,
2. Level II: $n_r = 40\%$, $n_{nr} = 10\%$ and $n_c = 50\%$,
3. Level III: $n_r = 35\%$, $n_{nr} = 15\%$ and $n_c = 50\%$,

The point and interval estimates of parameters based on this data are calculated under frequentist and Bayesian approaches and reported in Table 2.6. While calculating the Bayesian estimates, we don't have prior knowledge of real data, so hyper-parameters are chosen such that the prior density becomes non-informative. The estimates of median duration of breast-feeding for the recent child using classical and Bayesian methods are obtained for three levels as

Table 2.6: The point and interval estimates for breastfeeding data under ML and Bayes approaches.

					ML		*Bayes*	
Level	n_r	n_{nr}	n_c		*Estimate*	*ACI*	*Estimate*	*HPD*
I	45%	5%	50%	$\hat{\alpha}$	1.8049	(1.6824, 1.9274)	1.8059	(1.7203, 1.8906)
				$\hat{\beta}$	0.6226	(0.5465, 0.7028)	0.6251	(0.5558,0.6895)
				$\hat{\lambda}$	0.1196	(0.0796, 0.1749)	0.1260	(0.0727, 0.1665)
II	40%	10%	50%	$\hat{\alpha}$	1.8003	(1.6786, 1.9221)	1.8000	(1.7168, 1.8874)
				$\hat{\beta}$	0.6401	(0.5590, 0.7227)	0.6432	(0.5721, 0.7080)
				$\hat{\lambda}$	0.2554	(0.1987,0.3436)	0.2664	(0.1839, 0.3249)
III	35%	15%	50%	$\hat{\alpha}$	1.7924	(1.6709, 1.9138)	1.7928	(1.7061, 1.8809)
				$\hat{\beta}$	0.6446	(0.5653, 0.7285)	0.6480	(0.5764, 0.7129)
				$\hat{\lambda}$	0.4040	(0.3330, 0.5256)	0.4249	(0.3125, 0.4954)

1. Level I: Under MLE, the mean and median duration of breastfeeding for the recent child comes out to be 27.75 & 25.47 months respectively. In case of the Bayesian approach, the mean and median duration of breastfeeding of the recent child comes out to be 27.74 & 25.46 months respectively.

2. Level II: Under MLE, the mean and median duration of breastfeeding for a recent child comes out to be 27.34 & 25.09 months while in case of the Bayesian, the mean and median duration of breastfeeding comes out to be 27.33 & 25.07 months respectively.

3. Level III: Under MLE, the mean and median duration of breastfeeding for the recent child comes out to be 27.27 & 24.99 months while in case of the Bayesian approach the mean and median duration comes out to be 27.25 & 24.97 months respectively.

2.7 Conclusions and Discussions

In this chapter, a latent variable approach for the analysis of recall-based data is developed under classical and Bayesian frameworks. Due to the flexibility of covering the increasing, constant and decreasing nature of hazard rates, Weibull distribution is considered to model the time to event of interest. Under the classical framework, the E-M algorithm is used for point estimation and the observed Fisher information matrix is calculated using the missing information principle. In the Bayesian approach, parameters are estimated by utilizing informative priors under SELF. Gibbs Sampling and M-H algorithms are used for generating samples from full conditional posteriors. HPD intervals are calculated using the obtained MCMC samples. To explore the impact of monitoring points, some statistics such as the probability of experiencing the event, expected time in study, total time in study, expected number of recall, and non-recall observations are calculated. For illustration purposes, an extensive simulation study is carried out considering two patterns of monitoring times as uniform and exponential. The estimates obtained under classical and Bayesian frameworks are satisfactory under both monitoring time patterns. Finally, for real data applications, the duration of breastfeeding is estimated using proposed methodologies.

Chapter 3

Parametric Estimation of Recall-Based Competing Risk Data

3.1 Introduction

In time to event investigations, sometimes it is not possible to follow the individuals under study in real-time. Therefore instead of getting actual information about the event researchers obtain incomplete information in terms of left, right, or interval-censored data. Many clinical or demographical surveys are conducted in such a way that the data are recorded in interval-censored form.

A particular form of interval-censored data is named Type-I Interval censored data or current status data. Such data arise mainly in the fields of reliability engineering, clinical trials, cross-sectional, demographic, sample survey and tumorigenicity studies. In the case of current status data, the individual is monitored only once in an experiment and the status of the event is noted at that monitoring time. For example, Krailo and Pike (1983) discussed the menopause data on US women observed between $1960 - 62$, where age at menopause with corresponding causes of menopause (natural/operative) is observed based on their responses at interview time. The estimation of the distribution of age at weaning under current status data setup is discussed in Diamond et al. (1986), where the age of a child at weaning is considered as time to event and the age of the child as monitoring time. In another example, current status data is utilized for the estimation of age-specific

immunization discussed in Keiding et al. (1996). Initially, non-parametric approaches were developed to deal with current status data. For the point process, current status data statistical methods were derived by Sun and Kalbfleisch (1993). A test statistic is developed for the testing of equality of the mean functions of point processes, and its asymptotic properties were derived. In a study Sun (1999), authors proposed a non-parametric test for the comparison of distinctive treatments based on current status data under the assumption that observation times for each individual may be different in different treatment groups. A broad review of current status data with applications in many fields can be seen in the article Jewell and van der Laan (2003). Construction of confidence intervals under current status data is proposed by Banerjee and Wellner (2005). Analysis of current status data under informative observation time points is discussed by Zhang et al. (2005).

The theory of competing risks is used in reliability theory and survival analysis when the occurrence of an event of interest experiences more than one cause. It is often of interest to estimate the cause-specific quantities in the presence of other causes. The presence of other causes affects the occurrence of events of interest. In survival analysis a patient is at risk of two or more than two mutually exclusive causes and the death of the patient may be due to any of these causes in presence of others. In a similar fashion, the failure of a system or machine may be due to the failure of any of its components. In the case of competing risks, the data can be observed in the form of bivariate distribution, i.e. time to the event and corresponding causes. The competing risk model is studied by several researchers in different areas under non-parametric and parametric approaches. The non-parametric models do not assume any specific functional form of time to event of interest. Most of the initial work on competing risk theory is done under a non-parametric setup. The concepts and estimation under competing risks in reliability engineering and survival analysis along with real-life examples can be seen in details in (Prentice et al. (1978), David and Moeschberger (1978), Sinha (1986), Crowder (2001), Lawless (2003) and Lee and Wang (2003).

In competing risks, it may be the case where the cause of failure is not observed fully due to time or cost constraints but belongs to any subset of causes. Such cases are known as masked data problems in literature. The competing risks analysis with partial information

on cause becomes more challenging to deal with. In a time of failure-based data on computer component systems, Reiser et al. (1995) reported a masked data where, failure occurs due to malfunctioning of one of the three components the motherboard, power supply, or disc drive. Maximum likelihood estimators (MLEs) for two-component series systems with exponential distribution as failure time model derived by Miyakawa (1984). Generalizing the results from Miyakawa (1984), Usher and Hodgson (1988) introduced the concept of masked data under exact and partial masking. Mukhopadhyay and Basu (1993) applied the E-M algorithm to solve the masked data problem for the K-component series system considering the component as exponentially distributed. The alternate closed-form solution to masked data problems discussed by Usher and Hodgson (1988) is obtained by Lin and Guess (1994). An alternative ad-hoc iterative approach is developed to find MLE by Usher (1996) with Weibull as lifetime distribution under masked data setup. Analytical solution to three-component masked data problem under Bayesian approach is discussed by Flehinger et al. (2002a) with the series system having exponentially distributed component's failure. Sarhan (2003) derived MLEs and Bayes estimates of reliability functions of individual components for the series system with Pareto distribution as lifetime model. MLE and Bayes estimates for a family of distribution are obtained by Singh and Tomer (2011) and Tomer et al. (2014). In a recent study, Rai et al. (2021), masked data is analyzed under classical and Bayesian approaches with Lindley distribution considering masking probabilities as cause and time-dependent.

The concept of competing risks and masking problems can be observed in the case of current status data also. Hudgens et al. (2001) derived the non-parametric MLE of cumulative incidence functions for competing risks survival data and applied the method to data from a cohort of injecting drug users in Thailand susceptible to infection from HIV-1 subtypes B and E. The thesis Maathuis (2006) discussed the data that arises in cross-sectional survival studies with more than one cause. HIV vaccine clinical trial data is analyzed under current status with competing risks data set up. The estimates of distribution functions were obtained with non-parametric MLE and naive estimator and some asymptotic properties are also discussed. Maathuis and Hudgens (2011) has analyzed competing risks current status

data with continuous, discrete and grouped observation times. In a recent study, current status competing risks data are analyzed under a non-parametric approach by Koley and Dewanji (2019a). Further, Koley and Dewanji (2019b) analyzed the competing risks current status data after re-parametrization and applied the E-M algorithm for non-parametric inferences. Current status data under competing risks set up assuming time-dependent masking probabilities are analyzed by Koley and Dewanji (2020a) in an unpublished technical report. In continuation, Koley and Dewanji (2020b) utilized some additional information available in current status data to avoid model identifiability and discussed parameter estimation under non-parametric and parametric approaches.

In many cross-sectional studies, if an individual has experienced the event then at the monitoring time the interviewer asks the respondent to recall the time of the event of interest. Such studies are known as recall-based studies in the literature and make the current status data more informative. Incorporation of this additional information in a model may result in the improvement of required inferences. In recall-based data, observations are more informative than usual current status data. A surveyor's interest is not only to know the current status of an individual but also to get the information at the maximum extent about the event. Turnbull and Weiss (1978) presented a study based on a survey of high school students who were possibly addicted to marijuana, a psychoactive drug. In this study, 191 California students were asked whether they were using marijuana. Among those who answered 'Yes' some students were able to remember the exact age when they started taking marijuana. But many of them were not able to recall the exact age at the time of the interview and thus their replies generated left censored data. Whereas few of them never used marijuana by the date of the interview and their responses became right-censored observations. In literature, many recall-based studies like age at menopause in women, age at first sexual intercourse, age at marriage, or age at a first child are conducted frequently. The Biological Anthropology Unit of the Indian Statistical Institute, Calcutta also conducted an anthropometric survey from 2005 to 2011. The data were reported in Salehabadi et al. (2015), in which 2195 randomly selected girls aged between 7 years to 21 years, were surveyed. The data set contains an individual's age, menarcheal status, age at menarche and

some other related information. In the article Salehabadi et al. (2015), authors addressed the recall-based problem for menarcheal age and drew the inferences using a parametric model. The exponential and piecewise functional forms for the non-recall probabilities are considered in this article for modeling purposes. The same data is later explored under non-parametric approaches in presence of covariates by Salehabadi and Sengupta (2015), Salehabadi and Sengupta (2016) and Mirzaei Salehabadi et al. (2020). Sukumaran and Dewan (2019) studied recall-based competing risks data using the cause-specific hazard approach. They considered a piecewise functional form of non-recall probability and discussed the estimation of parameters along with some large sample properties. They also established the identifiability of the considered model. They had utilized the data on age at menarche and considered the order of births of girls as competing causes.

The present chapter proposes a new approach to model the recall-based competing risks data using a latent failure time-based model. In Section 3.2, we present the data setup and construction of the likelihood function for the observed data. In Section 3.3, we obtain ML estimates of parameters using the E-M algorithm and evaluate asymptotic confidence intervals by utilizing the missing information principle. Section 3.5 is dedicated to the Bayesian analysis under which point and interval estimators of parameters are obtained. In Section 3.6, a simulation study is performed for derived methodology with different sample sizes for uniform and exponential monitoring patterns. Section 3.7 describes the analysis of menarche data (Salehabadi et al. (2015)) considering the birth order of girls as competing causes. The conclusion and important remarks about the study are given in Section 3.8.

3.2 Data Structure and Construction of Likelihood

For inference purposes, if the data structure is well defined then the construction of the likelihood function is straightforward. However, if the data provide partial information that can not be recalled at the time of the study, one needs to modify the existing methodology. So, in this section, we established a set-up for the parametric study of the non-recall current status data. Suppose there are n individuals in an experiment and each individual is exposed

to k causes. The event of interest can occur due to any one of these K causes. If the latent event times for an individual for the K causes are $T_1, T_2, \cdots, T_K$ then the time to occurrence of an event (event time) T for an individual can be defined as $T = \min(T_1, T_2, \cdots, T_K)$. The event time due to a cause is a random variable and follows a distribution such as $T_k \sim F_k(t_k),\ k = 1, 2, \cdots, K$. It is assumed that each individual will be inspected only once during the experiment. The monitoring time, S, of an individual, is also a random variable and follows a distribution such as $S \sim G(s)$. In the formation of likelihood functions, the density of S can be treated as a nuisance variable. Also, the event time T and monitoring time S are assumed to be independent.

At a monitoring time point, one observes the status of whether an event has occurred or not. If the event has not happened till that time then it is considered as right-censored. If it has already occurred then there arise two possibilities; either an individual is able to recall the exact event time or it will be a non-recall observation. In this study, observation is considered as recalled if one specifies the time as well as the cause of the event exactly at the monitoring time. In the case of non-recall, we have the information that the event has occurred before the monitoring time but the exact event time, as well as the cause of its occurrence, are unknown.

For i^{th} individual let us define two indicators δ_i and ϵ_i to structure the observed current status data set in presence of competing risks. The indicator function $\delta_i = I(T_i \leq S_i)$ will return 1 if the event of interest has occurred before monitoring time and it will give 0, otherwise. Further for an event with $\delta_i = 1$, that is, an event has already occurred at the monitoring time point, the indicator function ϵ_i takes value 1 if an individual is able to recall the exact event time and cause whereas it is considered as 0 if the event time and cause are unknown. Thus, if $\delta_i = 1$ and $\epsilon_i = 1$ that means the i^{th} individual exactly recalls the event time, $T_i = t_i$ as well as its actual cause, $C_i = k;\ k = 1, 2, \cdots, K$. Alternatively, if $\delta_i = 1$ and $\epsilon_i = 0$ then it indicates that $T_i < S_i$ with a case of non-recall information and we get a left-censored observation for that individual. Also for an individual, if the event does not occur till the monitoring it will be considered as the right-censored observation and then indicator δ_i will take value 0. Under recall-based competing risk setup the observed data

vector can be written as $\underline{D} = (S, \delta, \epsilon, T\delta\epsilon, C\delta\epsilon)$. For the i^{th} individual the observed data can be denoted $d_i = (s_i, \delta_i, \epsilon_i, t_i\delta_i\epsilon_i, c_i\delta_i\epsilon_i);\ i = 1, 2, \ldots, n$.

For the available information in data, d, and parameter vector η, the likelihood function can be defined as

$$L(\eta|d) = \prod_{i=1}^{n} \left[\prod_{k} \left\{ f(t_i, \eta_k) \prod_{\substack{l=1 \\ l\neq k}}^{K} \bar{F}(t_i, \eta_l)(1 - \psi_k(s_i, t_i)) \right\} \right]^{\delta_i\epsilon_i} \left[\sum_{k} \left\{ \int_0^{s_i} f(u, \eta_k) \prod_{\substack{l=1 \\ l\neq k}}^{K} \bar{F}(u, \eta_l)\psi_k(s_i, u)du \right\} \right]^{\delta_i(1-\epsilon_i)} \left[\bar{F}(s_i, \eta)\right]^{(1-\delta_i)}, \tag{3.1}$$

where $\psi_k(s, t)$ represents the non-recall probability when actual event occurred due to k^{th} cause at time t. Here, it is assumed that the non-recall behavior of an individual depends on time as well as cause of occurrence of event of interest. So, if the i^{th} individual experience the event but unable to recall the time and cause then the non-recall probability can be written as

$$\psi_k(s_i, t_i) = 1 - \exp\{-\lambda_k(s_i - t_i)\},\ \lambda_k > 0,\ k = 1, 2, \cdots, K. \tag{3.2}$$

Alternatively, the recall probability can be written as $1 - \psi_k(s, t) = \exp\{-\lambda_k(s - t)\}$. This probability function is based on the difference between the event time and monitoring time of an individual i.e. $(s - t)$. The probability of recall comes out to be 1 in the case when the monitoring time coincides with the event time. However, as time passes the recall probability reduces exponentially with the rate λ_k.

3.3 Expectation-Maximization Implementation

In this section, we consider non-recall observations as missing observations and use E-M algorithm for the estimation of the model parameters. Here it can be observed that the likelihood expression in (3.1) is not in a convenient form for further analysis. So it becomes necessary to deal with the terms in the expression that involve non-recall observations.

For a given sample of size n, let n_r, n_{nr} and n_c respectively denote the number of recall, non-recall and censored observations. Also for calculation purpose we define $n_r = \sum \delta_i \epsilon_i$, $n_{nr} = \sum \delta_i(1-\epsilon_i)$; $i = 1, 2, \cdots, n$ and $n_c = n - n_r - n_{nr}$. Using the approach of Rai et al. (2021), we defined a latent variable Z, following multinomial distribution for non-recall observations. For i^{th} individual under non-recall category $Z_i \sim \mathcal{MD}(1, p_{i1}, p_{i2}, \cdots, p_{iK})$. For this multinomial variate, the probabilities of success can be defined as

$$p_{ik} = \frac{I_k(S_i)}{\sum\limits_k I_k(S_i)}; \quad i = 1, 2, \cdots, n_{nr};\ k = 1, 2, \cdots, K, \tag{3.3}$$

where, $I_k(.)$ is given by

$$I_k(s) = \int_0^s f(u, \theta_k) \prod_{\substack{l=1 \\ l \neq k}}^{K} \bar{F}(u, \theta_l) \psi_k(s, u) du.$$

By using the posterior probabilities defined in (3.3) and receptively the value of latent variable, we will be able to assign a cause, $k(k = 1, 2, \cdots, K)$, to each of the non-recall observation. Now by using the multinomial variable Z, the likelihood expression (3.1) explicitly can be re-written as

$$\begin{aligned} L(\eta|d) = & \prod_{k=1}^{K} \prod_{i=1}^{n_k} \left[f(t_i, \eta_k) \prod_{\substack{l=1 \\ l \neq k}}^{K} \bar{F}(t_i, \eta_l) \Big(1 - \psi_k(s_i, t_i)\Big) \right] \\ & \prod_{i=1}^{n_{nr}} \left[\prod_{k=1}^{K} \left\{ \int_0^{s_i} f(u, \eta_k) \prod_{\substack{l=1 \\ l \neq k}}^{K} \bar{F}(u, \eta_l) \psi_k(s_i, u) du \right\}^{z_i} \right] \prod_{i=1}^{n_c} \left[\bar{F}(s_i, \eta) \right], \end{aligned} \tag{3.4}$$

where, n_k is the total number of individuals who recalled that the occurrence of event is due to k^{th} cause and hence $\sum\limits_{k=1}^{K} n_k = n_r$.

As of now the likelihood function (3.4) is more convenient than (3.1) but still, we have to deal with expression due to non-recall observations. As in the case of non-recall observation, the only available information regarding the event time is that it is observed during $(0, S)$. So in the next step, the issue of integral term has been resolved by using the equivalent

quantity estimation (EQE) method (see Tan (2007), Wang (2016), Fan et al. (2019)). Let us introduce an equivalent quantity T^* for the non-recall observations lying in interval $(0, S)$. For i^{th} individual with monitoring time interval $(0, s_i)$ and given associated cause, the conditional density of t^*_{ik} can be written as

$$f\Big(t^*_{ik}|t^*_{ik} \in (0, s_i), z_i, \eta_k\Big) = \frac{f(t^*_{ik}, \eta_k)}{1 - \bar{F}(s_i, \eta_k)}; \quad i = 1, 2, \ldots, n_{nr}, k = 1, 2, \ldots, K. \tag{3.5}$$

Now the likelihood function (3.4) can be modified using the density (3.5) for non-recall set as follows

$$\begin{aligned} L(\eta|d) = &\prod_{k=1}^{K}\prod_{i=1}^{n_k}\left[f(t_i, \eta_k)\prod_{\substack{l=1\\l\neq k}}^{K}\bar{F}(t_i, \eta_l)\Big(1 - \psi_k(s_i, t_i)\Big)\right] \\ &\prod_{i=1}^{n_{nr}}\left[\prod_{k=1}^{K}\left\{f(t^*_{ik}, \eta_k)\prod_{\substack{l=1\\l\neq k}}^{K}\bar{F}(t^*_{ik}, \eta_l)\psi_k(s_i, t^*_{ik})\right\}^{z_i}\right]\prod_{i=1}^{n_c}\left[\bar{F}(s_i, \eta)\right]. \end{aligned} \tag{3.6}$$

For the likelihood function given in (3.6), we have utilized two latent variables. Now on the basis of these two variables, Z and T^*, we can introduce a new latent variable, say W^*, for non-recall probability terms to make the computation easy. Let the random variable W^* follows an exponential distribution with mean $1/\lambda_k$ and truncated at $(S - T^*)$. Hence, the missing data vector can be written as $\underline{D}^* = (Z, T^*, W^*)$ be the missing data vector. Then $d^*_i = (z_i, t^*_{ik}, w^*_{ik});\ i = 1, 2, \ldots, n_{nr}, k = 1, 2, \ldots, K$ denotes the observed value for i^{th} individual falling in the non-recall category.

Now we proceed with considering the events time to be exponential distribution for all causes. The exponential distribution has constant hazard rate and used to model the remission time of chronic diseases (heart disease, cancer, and diabetes) and the lifetime of youth period. Let assume that event time under k^{th} cause follows an exponential lifetime with mean $1/\theta_k$, i.e. $T \sim \mathcal{E}(t; \theta_k);\ k = 1, 2, \cdots, K$. Also, the set of parameters can be represented by $\Theta = \{\theta_1, \theta_2, \cdots, \theta_K\}$ and $\Lambda = \{\lambda_1, \lambda_2, \cdots, \lambda_K\}$. Now under the assumption of exponential lifetime and for complete data set (d, d^*), the likelihood function (3.6) can be

expressed as a complete likelihood function as

$$L_c(\Theta,\Lambda|d,d^*) = \prod_{k=1}^{K}\prod_{i=1}^{n_k}\left[\theta_k \exp\{-\theta_k t_i\}\prod_{\substack{l=1\\l\neq k}}^{K}\exp\{-\theta_l t_i\}\exp\{-\lambda_k(s_i-t_i)\}\right] \prod_{i=1}^{n_{nr}}\prod_{k=1}^{K}\left[\theta_k\lambda_k \exp\{-\theta_k t^*_{ik}\}\prod_{\substack{l=1\\l\neq k}}^{K}\exp\{-\theta_l t^*_{ik}\}\exp\{-\lambda_k w^*_{ik}\}\right]^{z_i} \prod_{i=1}^{n_c}\left[\exp\left\{-s_i\sum_{k=1}^{K}\theta_k\right\}\right]. \tag{3.7}$$

The log-likelihood function of complete data is obtained by taking logarithm of (3.7) as below

$$l_c(\Theta,\Lambda|d,d^*) = \sum_{k=1}^{K}\sum_{i=1}^{n_k}\left[\ln(\theta_k)-\theta_k t_i-\sum_{\substack{l=1\\l\neq k}}^{K}\theta_l t_i-\lambda_k(s_i-t_i)\right]-\sum_{k=1}^{K}\theta_k\sum_{i=1}^{n_c}s_i + \sum_{i=1}^{n_{nr}}\sum_{k=1}^{K}z_i\left[\ln(\theta_k)+\ln(\lambda_k)-\theta_k t^*_{ik}-\sum_{\substack{l=1\\l\neq k}}^{K}\theta_l t^*_{ik}-\lambda_k w^*_{ik}\right]. \tag{3.8}$$

As we introduced latent variables to make this likelihood function complete, the E-M technique is one of the suitable methods to obtain the ML estimators of the parameters. In order to apply the E step of E-M algorithm let us define the quantity $Q\left(\Theta,\Lambda|\hat{\Theta}^{(m)},\hat{\Lambda}^{(m)}\right) = E\left[l_c(\Theta,\Lambda|d,d^*)|d,\hat{\Theta}^{(m)},\hat{\Lambda}^{(m)}\right]$, where $\hat{\Theta}^{(m)}$ and $\hat{\Lambda}^{(m)}$ be the possible estimates of parameters at m^{th} iteration. A nested expectation is taken with respect to the three latent variables, Z, T^* and W^* which are introduced earlier in this section. By using the nested expectation, the expression of $Q\left(\Theta,\Lambda|\hat{\Theta}^{(m)},\hat{\Lambda}^{(m)}\right)$ can be derived as follows

$$
\begin{aligned}
Q\Big(\Theta,\Lambda|\hat{\Theta}^{(m)},\hat{\Lambda}^{(m)}\Big) &= E_{W^*}\left[E_{T^*}\left[E_Z\left[l_c(\Theta,\Lambda|d,d^*)|d,\hat{\Theta}^{(m)},\hat{\Lambda}^{(m)}\right]\right]\right] \\
&= \sum_{k=1}^{K}\sum_{i=1}^{n_k}\left[\ln\theta_k - \theta_k t_i - \lambda_k(s_i-t_i) - \sum_{\substack{l=1\\l\neq k}}^{K}\theta_l t_i\right] \\
&+ \sum_{i=1}^{n_{nr}}\sum_{k=1}^{K}\xi_1\Big(s_i;\hat{\theta}_k^{(m)},\hat{\lambda}_k^{(m)}\Big)\left[\ln\theta_k + \ln\lambda_k - \theta_k\xi_2\Big(t_{ik}^*;\hat{\theta}_k^{(m)}\Big)\right] \\
&+ \sum_{i=1}^{n_{nr}}\sum_{k=1}^{K}\xi_1\Big(s_i;\hat{\theta}_k^{(m)},\hat{\lambda}_k^{(m)}\Big)\left[-\lambda_k\xi_3\Big(w_{ik}^*;\hat{\lambda}_k^{(m)}\Big)\right] \\
&- \sum_{i=1}^{n_{nr}}\sum_{k=1}^{K}\xi_1\Big(s_i;\hat{\theta}_k^{(m)},\hat{\lambda}_k^{(m)}\Big)\left[\sum_{\substack{l=1\\l\neq k}}^{K}\theta_l\xi_2\Big(t_{ik}^*;\hat{\theta}_k^{(m)}\Big)\right] - \left(\sum_{k=1}^{K}\theta_k\sum_{i=1}^{n_c}s_i\right).
\end{aligned}
\tag{3.9}
$$

The latent variable Z is following a multinomial distribution whereas the variables T^* and W^* having truncated exponential distribution with parameter $1/\theta_k$ and $1/\lambda_k$, respectively. Hence, the expectation terms used at $(m+1)^{th}$ iteration of the E-M algorithm is calculated based on given observed data and values of parameters at m^{th} iteration and given as

$$
\begin{aligned}
&\xi_1(s_i;\theta_k,\lambda_k) = E\big[z_i\big] = p_{ik}, \\
&\xi_2(t_{ik}^*;\theta_k) = E\Big[t_{ik}^*\big|t_{ik}^* < s_i, z_i, \theta_k\Big] = \frac{\int_0^{s_i} u\theta_k\exp\{-\theta_k u\}du}{1-\exp\{-\theta_k s_i\}} \\
&\qquad\qquad = \frac{1}{\theta_k}\left[1 - \frac{\theta_k s_i\exp\{-\theta_k s_i\}}{1-\exp\{-\theta_k s_i\}}\right], \\
&\xi_3(w_{ik}^*;\lambda_k) = E\Big[w_{ik}^*\big|w_{ik}^* < (s_i - t_{ik}^*), z_i, \lambda_k\Big] = \frac{\int_0^{s_i-t_{ik}^*} u\lambda_k\exp\{-\lambda_k u\}du}{1-\exp\{-\lambda_k(s_i-t_{ik}^*)\}} \\
&\qquad\qquad = \frac{1}{\lambda_k}\left[\frac{1-\Big(1+\lambda_k(s_i-t_{ik}^*)\Big)\exp\{-\lambda_k(s_i-t_{ik}^*)\}}{1-\exp\{-\lambda_k(s_i-t_{ik}^*)\}}\right].
\end{aligned}
$$

In M step of E-M algorithm, we differentiate the function obtained in (3.9) partially with respect to θ_k and λ_k; $k = 1, 2, \cdots, K$ and then equate to zero. The estimated value of parameters at $(m+1)^{th}$ iteration can be computed on the basis of m^{th} step by using the

following expressions

$$\hat{\theta}_k^{(m+1)} = \frac{n_k + \sum_{i=1}^{n_{nr}} \xi_1\left(s_i; \theta_k^{(m)}, \lambda_k^{(m)}\right)}{\sum_{i=1}^{n_k} t_i + \sum_{i=1}^{n_{nr}} \xi_1\left(s_i; \theta_k^{(m)}, \lambda_k^{(m)}\right)\xi_2\left(t_{ik}^*; \theta_k^{(m)}\right) + \sum_{i=1}^{n_c} s_i} \tag{3.10}$$

$$\hat{\lambda}_k^{(m+1)} = \frac{\sum_{i=1}^{n_{nr}} \xi_1\left(s_i; \theta_k^{(m)}, \lambda_k^{(m)}\right)}{\sum_{i=1}^{n_k}(s_i - t_i) + \sum_{i=1}^{n_{nr}} \xi_1\left(s_i; \theta_k^{(m)}, \lambda_k^{(m)}\right)\xi_3\left(w_{ik}^*; \lambda_k^{(m)}\right)}. \tag{3.11}$$

Now, the steps of the E-M algorithm are repeated until the estimates converge, simultaneously. The iterations can be terminate when $|\hat{\theta}_k^{(m+1)} - \hat{\theta}_k^{(m)}| < \epsilon$ and $|\hat{\lambda}_k^{(m+1)} - \hat{\lambda}_k^{(m)}| < \epsilon$ where ϵ is a sufficiently small positive real number.

3.4 Observed Information Matrix

In a current status experiment, one does not observe the complete information about the event so a direct calculation of the Fisher information matrix is not sufficient. In this situation, a computation technique proposed by Louis (1982) is a cogent alternative. In this method when the E-M algorithm is used to find the ML estimates then the information matrix can be obtained by using the second derivatives of the complete data log-likelihood given the observed data. For the set of parameters (θ, λ) the observed information matrix can be obtained by using the equation $I(\hat{\theta}, \hat{\lambda}) = I_1 - I_2 {I_2}^T$, where the matrix I_1 is the negative of second order derivatives of (3.9) and I_2 is the gradient vector obtained by differentiating (3.9) with respect to parameters. The structure of observed information matrix is given by

$$I(\hat{\Theta}, \hat{\Lambda}) = \left[\begin{array}{ccc:ccc}
-\frac{\partial^2 Q}{\partial \theta_1^2} - \left(\frac{\partial Q}{\partial \theta_1}\right)^2 & \cdots & -\frac{\partial^2 Q}{\partial \theta_1 \partial \theta_K} - \frac{\partial Q}{\partial \theta_1}\frac{\partial Q}{\partial \theta_K} & -\frac{\partial^2 Q}{\partial \theta_1 \partial \lambda_1} - \frac{\partial Q}{\partial \theta_1}\frac{\partial Q}{\partial \lambda_1} & \cdots & -\frac{\partial^2 Q}{\partial \theta_1 \partial \lambda_K} - \frac{\partial Q}{\partial \theta_1}\frac{\partial Q}{\partial \lambda_K} \\
-\frac{\partial^2 Q}{\partial \theta_2 \partial \theta_1} - \frac{\partial Q}{\partial \theta_2}\frac{\partial Q}{\partial \theta_1} & \cdots & -\frac{\partial^2 Q}{\partial \theta_2 \partial \theta_K} - \frac{\partial Q}{\partial \theta_2}\frac{\partial Q}{\partial \theta_K} & -\frac{\partial^2 Q}{\partial \theta_2 \partial \lambda_1} - \frac{\partial Q}{\partial \theta_2}\frac{\partial Q}{\partial \lambda_1} & \cdots & -\frac{\partial^2 Q}{\partial \theta_2 \partial \lambda_K} - \frac{\partial Q}{\partial \theta_2}\frac{\partial Q}{\partial \lambda_K} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
-\frac{\partial^2 Q}{\partial \theta_K \partial \theta_1} - \frac{\partial Q}{\partial \theta_K}\frac{\partial Q}{\partial \theta_1} & \cdots & -\frac{\partial^2 Q}{\partial \theta_K^2} - \left(\frac{\partial Q}{\partial \theta_K}\right)^2 & -\frac{\partial^2 Q}{\partial \theta_K \partial \lambda_1} - \frac{\partial Q}{\partial \theta_K}\frac{\partial Q}{\partial \lambda_1} & \cdots & -\frac{\partial^2 Q}{\partial \theta_K \partial \lambda_K} - \frac{\partial Q}{\partial \theta_K}\frac{\partial Q}{\partial \lambda_K} \\
\hdashline
-\frac{\partial^2 Q}{\partial \lambda_1 \partial \theta_1} - \frac{\partial Q}{\partial \lambda_1}\frac{\partial Q}{\partial \theta_1} & \cdots & -\frac{\partial^2 Q}{\partial \lambda_1 \partial \theta_K} - \frac{\partial Q}{\partial \lambda_1}\frac{\partial Q}{\partial \theta_K} & -\frac{\partial^2 Q}{\partial \lambda_1^2} - \left(\frac{\partial Q}{\partial \lambda_1}\right)^2 & \cdots & -\frac{\partial^2 Q}{\partial \lambda_1 \partial \lambda_K} - \frac{\partial Q}{\partial \lambda_1}\frac{\partial Q}{\partial \lambda_K} \\
-\frac{\partial^2 Q}{\partial \lambda_2 \partial \theta_1} - \frac{\partial Q}{\partial \lambda_2}\frac{\partial Q}{\partial \theta_1} & \cdots & -\frac{\partial^2 Q}{\partial \lambda_2 \partial \theta_K} - \frac{\partial Q}{\partial \lambda_2}\frac{\partial Q}{\partial \theta_K} & -\frac{\partial^2 Q}{\partial \lambda_2 \partial \lambda_1} - \frac{\partial Q}{\partial \lambda_2}\frac{\partial Q}{\partial \lambda_1} & \cdots & -\frac{\partial^2 Q}{\partial \lambda_2 \partial \lambda_K} - \frac{\partial Q}{\partial \lambda_2}\frac{\partial Q}{\partial \lambda_K} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
-\frac{\partial^2 Q}{\partial \lambda_K \partial \theta_1} - \frac{\partial Q}{\partial \lambda_K}\frac{\partial Q}{\partial \theta_1} & \cdots & -\frac{\partial^2 Q}{\partial \lambda_K \partial \theta_K} - \frac{\partial Q}{\partial \lambda_K}\frac{\partial Q}{\partial \theta_K} & -\frac{\partial^2 Q}{\partial \lambda_K \partial \lambda_1} - \frac{\partial Q}{\partial \lambda_K}\frac{\partial Q}{\partial \lambda_1} & \cdots & -\frac{\partial^2 Q}{\partial \lambda_K^2} - \left(\frac{\partial Q}{\partial \lambda_K}\right)^2
\end{array}\right]\Bigg|_{(\Theta,\Lambda)=(\hat{\Theta},\hat{\Lambda})}$$

Further using the partition of the matrix, the observed Fisher information matrix can be re-written as

$$I(\hat{\Theta}, \hat{\Lambda}) = \left[\begin{array}{c:c} I(\Theta,\Theta) & I(\Theta,\Lambda) \\ \hdashline I(\Lambda,\Theta) & I(\Lambda,\Lambda) \end{array}\right]\Bigg|_{(\Theta,\Lambda)=(\hat{\Theta},\hat{\Lambda})}$$

The variance-covariance matrix can be found by taking inverse of observed Fisher information matrix. The expression used to construct of observed Fisher information matrix for

$k = 1, 2, \cdots, K$ and $l = 1, 2, \cdots, K$ are as follows

$$\frac{\partial Q}{\partial \theta_k} = \frac{n_r}{\theta_k} - \sum_{i=1}^{n_r} t_i - \sum_{i=1}^{n_c} s_i + \sum_{i=1}^{n_{nr}} \frac{1}{\theta_k}\xi_1(s_i;\theta_k,\lambda_k) - \sum_{i=1}^{n_{nr}} \xi_1(s_i;\theta_k,\lambda_k)\xi_2(t^*_{ik};\theta_k),$$
$$\frac{\partial^2 Q}{\partial \theta_k \partial \theta_l} = \frac{\partial^2 Q}{\partial \theta_k \partial \lambda_k} = 0,$$
$$\frac{\partial^2 Q}{\partial \theta_k^2} = -\frac{n_r}{\theta_k^2} - \sum_{i=1}^{n_{nr}} \frac{1}{\theta_k^2}\xi_1(s_i;\theta_k,\lambda_k),$$
$$\frac{\partial Q}{\partial \lambda_k} = -\sum_{i=1}^{n_r}(s_i - t_i) + \sum_{i=1}^{n_{nr}} \frac{1}{\lambda_k}\xi_1(s_i;\theta_k,\lambda_k) - \sum_{i=1}^{n_{nr}} \xi_1(s_i;\theta_k,\lambda_k)\xi_3(w^*_{ik};\lambda_k),$$
$$\frac{\partial Q}{\partial \lambda_k \partial \theta_k} = \frac{\partial Q}{\partial \lambda_k \partial \lambda_l} = 0,$$
$$\frac{\partial^2 Q}{\partial \lambda_k^2} = -\sum_{i=1}^{n_{nr}} \frac{1}{\lambda_k^2}\xi_1(s_i;\theta_k,\lambda_k).$$

Under regularity conditions and from the asymptotic property of ML estimator, Θ and Λ follows Multivariate Normal distribution, i.e.,

$$\sqrt{n}(\hat{\Theta} - \Theta) \sim \mathcal{MVN}\Big(0, nI^{-1}(\hat{\Theta},\hat{\Theta})\Big); \quad \text{as} \quad n \to \infty,$$

and

$$\sqrt{n}(\hat{\Lambda} - \Lambda) \sim \mathcal{MVN}\Big(0, nI^{-1}(\hat{\Lambda},\hat{\Lambda})\Big); \quad \text{as} \quad n \to \infty,$$

where $I^{-1}(\hat{\Theta},\hat{\Theta})$ and $I^{-1}(\hat{\Lambda},\hat{\Lambda})$ are the upper diagonal and bottom diagonal elements of variance-covariance matrix $I(\hat{\Theta},\hat{\Lambda})$. We can calculation the standard error by taking square root of these diagonal elements of variance-covariance matrix. We can construct the $100(1-\alpha)\%$ ACIs for Θ and Λ by $\hat{\Theta} \pm z_{\alpha/2}\sqrt{I^{-1}(\hat{\Theta},\hat{\Theta})}$ and $\hat{\Lambda} \pm z_{\alpha/2}\sqrt{I^{-1}(\hat{\Lambda},\hat{\Lambda})}$. Here, $z_{\alpha/2}$ denotes the $100(\alpha/2)\%$ upper percentile of standard normal distribution.

3.5 Bayesian Estimation

In Bayesian estimation, parameters are considered as random variables which are assumed as fixed but unknown in case of classical framework. Here, we are assuming gamma distribution which is a conjugate prior for θ_k as well as for λ_k i.e. $\theta_k \sim \mathcal{G}(a_k, b_k)$ and $\lambda_k \sim \mathcal{G}(c_k, d_k)$; $k =$

$1, 2, \ldots, K$. Let $\pi(\Theta, \Lambda) = \pi(\theta_1, \theta_1, \cdots, \theta_K, \lambda_1, \lambda_2, \cdots, \lambda_K)$ be the joint prior distribution of (Θ, Λ). Under independence assumption the joint prior density $\pi(\Theta, \Lambda)$ is given by

$$\pi(\Theta, \Lambda) \propto \prod_{k=1}^{K} \theta_k^{a_k - 1} \lambda_k^{c_k - 1} \exp\{-\theta_k b_k - \lambda_k d_k\}; \quad a_k, b_k, c_k, d_k > 0, k = 1, 2, \cdots, K. \tag{3.12}$$

Using the conditional likelihood given in (3.7), the joint prior given in (3.12) and the Bayes theorem, the conditional joint posterior distribution, is given by

$$\begin{aligned}\Pi(\Theta, \Lambda | d, d^*) \propto & \prod_{k=1}^{K} \prod_{i=1}^{n_k} \left[\theta_k \exp\{-\theta_k t_i\} \prod_{\substack{l=1 \\ l \neq k}}^{K} \exp\{-\theta_l t_i\} \exp\{-\lambda_k (s_i - t_i)\} \right] \\ & \prod_{i=1}^{n_{nr}} \prod_{k=1}^{K} \left[\theta_k \lambda_k \exp\{-\theta_k t_{ik}^*\} \prod_{\substack{l=1 \\ l \neq k}}^{K} \exp\{-\theta_l t_{ik}^*\} \exp\{-\lambda_k w_{ik}^*\} \right]^{z_i} \\ & \prod_{i=1}^{n_c} \left[\exp\left\{ -s_i \sum_{k=1}^{K} \theta_k \right\} \right] \prod_{k=1}^{K} \theta_k^{a_k - 1} \lambda_k^{c_k - 1} \exp\{-\theta_k b_k - \lambda_k d_k\}\end{aligned} \tag{3.13}$$

As in earlier section the samples from t^* and w^* can be generated independently from exponential random variables with means $1/\theta_k$ and $1/\lambda_k$ truncated at time points s and $(s - t^*)$, respectively. The full conditionals for θ_k and λ_k; $k = 1, 2, \cdots, K$ are obtained from posterior distribution (3.13), which are given below

$$\theta_k | \{\Theta, \Lambda\}_{(-\theta_k)} \sim \mathcal{G}\left(n_k + Z^* + a_k, \ \sum_{i=1}^{n_k} t_i + \sum_{i=1}^{n_{nr}} z_i t_{ik}^* + \sum_{i=1}^{n_c} s_i + b_k \right), \tag{3.14}$$

$$\lambda_k | \{\Theta, \Lambda\}_{(-\lambda_k)} \sim \mathcal{G}\left(Z^* + c_k, \ \sum_{i=1}^{n_k} (s_i - t_i) + \sum_{i=1}^{n_{nr}} z_i w_{ik}^* + d_k \right). \tag{3.15}$$

where, $Z^* = \sum_{i=1}^{n_{nr}} z_i$. Here from (3.14) and (3.15), we can see that the parameters θ_k and λ_k are conditionally independent from each other for given z, t^* and w^*. The value of (Θ, Λ) can be updated simultaneously using the Gibbs Sampler algorithm. The algorithm used for generating samples from full conditionals using the Gibbs sampling is given below:

1. Step 1: Set initial values of parameters, say $\left(\Theta^{(0)}, \Lambda^{(0)}\right)$ and generate

$$z_i \sim \mathcal{MD}(1, p_{i1}, p_{i2}, \cdots, p_{iK}); \ i = 1, 2, \cdots, n_{nr}.$$

2. Step 2: For given values of z_i, θ_k and λ_k, generate t^*_{ik} using

$$t^*_{ik} = -\frac{1}{\theta_k} \ln\Big[1 - u_i\Big(1 - \exp\{-\theta_k s_i\}\Big)\Big];\ i = 1, 2, \cdots, n_{nr};\ k = 1, 2, \cdots, K, \quad (3.16)$$

where $u \sim \mathcal{U}(0, 1)$.

3. Step 3: Based on values of z_i, t^*_{ik} and λ_k, observations on w^*_{ik} can be generated by using expression

$$w^*_{ik} = -\frac{1}{\lambda_k} \ln\Big[1 - u_i\Big(1 - \exp\{-\lambda_k(s_i - t^*_{ik})\}\Big)\Big];\ i = 1, 2, \cdots, n_{nr};\ k = 1, 2, \cdots, K. \quad (3.17)$$

4. Step 4: Finally for given values of z_i, t^*_{ik} and w^*_{ik}, generate observations on $\theta^{(1)}_k$ and $\lambda^{(1)}_k$ using (3.14) and (3.15) respectively.

The current state is $\big(\Theta^{(1)}, \Lambda^{(1)}\big)$. Now steps 1 to 4 are replicated M times to obtain a sequence of Markov Chain Monte Carlo (MCMC) chain of random variables

$$\left\{\big(\Theta^{(1)}, \Lambda^{(1)}\big), \big(\Theta^{(2)}, \Lambda^{(2)}\big), \cdots, \big(\Theta^{(M)}, \Lambda^{(M)}\big)\right\}.$$

After discarding the burn-in sample from generated chains the stationarity and convergence can be checked through the cumsum plot and Gelman and Rubin's test statistics (Gelman and Rubin (1992)). With a suitable choice of the thinning interval, we remove the auto-correlation from the generated MCMC chain. Finally, we get a chain of length M' and these samples from conditional posteriors can be used to find the point and interval estimates in the Bayesian framework. The highest posterior density intervals are obtained by utilizing the method of Chen and Shao (1999).

3.6 Simulation Study

In this section, we have presented a simulation study to check the performance of our proposed methodology. The computation is executed using the **R** software. The methodologies

in the classical and Bayesian sections are discussed generally for K causes, but for simulation study purposes, we consider $K = 2$, i.e. there are only two causes responsible for the happening of the event. Particular expressions for two causes are given in Appendix 3.9. For simulation study, first we have generated two samples, each of size n, such as $T_1 \sim \mathcal{E}(\theta_1)$ and $T_2 \sim \mathcal{E}(\theta_2)$ for the assumed values of parameters $\theta_1 = 1.10$ and $\theta_2 = 1.12$. These T_1 and T_2 are considered to be the latent event times for individuals due to causes 1 and 2, respectively. According to assumption the event time is $T = \min(T_1, T_2)$. Also for an individual if $T = T_1$ then the cause of the event is cause 1 or if $T = T_2$ then it is cause 2. To explore the impact of monitoring time on the inference of various estimators we consider two different patterns of monitoring points. In the first case, it is assumed that monitoring time, S, is uniformly distributed over the individual's event time whereas in the second case exponential distribution is considered to generate monitoring times with a rate of 0.50.

As in the current status set up, we may have three types of individuals in the study: those who have experienced the event and are able to recall the exact event time and cause as well; those who have experienced the event but are not able to recall the exact event time and cause and finally some are in right-censored category for which event of interest has not been experienced at the monitoring time point. So, to introduce these three categories artificially in generated sample we first fix a set of parameter $(\lambda_1, \lambda_2) = \{(0.10, 0.11), (0.20, 0.25)\}$ for non-recall category. These sets of parameters are chosen in such a way that we can maintain a proportion of non-recall and right-censored observations at around 30% and 40% as per the sample of size n. An observation is randomly assigned as non-recall or censored in the sample.

By using the E-M algorithm discussed in Section 3.3, we calculated the estimates of parameters and their mean square errors and absolute biases are presented in various tables. Also, the interval estimates of parameters are evaluated using the asymptotic properties of the ML estimator and corresponding ALs and CPs are given in tables. Under Bayesian analysis, with the gamma distribution as a conjugate prior, samples are generated using the Gibbs sampling algorithm. The hyper-parameters of prior distributions for Bayesian analysis are chosen using moment matching criteria. Under this approach, to find the values

of hyper-parameters, we match the mean and variance of MLEs (based on M samples) with the mean and variance of chosen prior density. Under the Bayesian approach, along with the MSE and AB of point estimates, the HPD intervals with their average lengths and coverage probabilities are reported. The whole simulation process is replicated 1000 times to average out results under classical and Bayes approaches.

Table 3.1: The mean square error and absolute bias for parameters set $(\theta_1, \theta_2) = (1.10, 1.12)$ and $(\lambda_1, \lambda_2) = \{(0.10, 0.11), (0.20, 0.25)\}$ under uniform monitoring points for varying sample size n.

			ML			*Bayes*		
(λ_1, λ_2)		$n \rightarrow$	80	150	250	80	150	250
(0.10, 0.11)	$\hat{\theta}_1$	*MSE*	0.0433	0.0249	0.0172	0.0383	0.0233	0.0165
		AB	0.1720	0.1285	0.1088	0.1615	0.1241	0.1064
	$\hat{\theta}_2$	*MSE*	0.0460	0.0259	0.0216	0.0408	0.0242	0.0207
		AB	0.1773	0.1315	0.1222	0.1666	0.1269	0.1197
	$\hat{\lambda}_1$	*MSE*	0.0044	0.0012	0.0006	0.0015	0.0007	0.0005
		AB	0.0527	0.0263	0.0193	0.0319	0.0198	0.0168
	$\hat{\lambda}_2$	*MSE*	0.0036	0.0011	0.0006	0.0013	0.0006	0.0004
		AB	0.0464	0.0259	0.0200	0.0289	0.0199	0.0173
(0.20, 0.25)	$\hat{\theta}_1$	MSE	0.0437	0.0323	0.0312	0.0384	0.0300	0.0298
		AB	0.1732	0.1475	0.1504	0.1692	0.1419	0.1407
	$\hat{\theta}_2$	*MSE*	0.0597	0.0484	0.0481	0.0517	0.0453	0.0451
		AB	0.2047	0.1879	0.1797	0.1903	0.1813	0.1723
	$\hat{\lambda}_1$	*MSE*	0.0079	0.0042	0.0024	0.0041	0.0035	0.0020
		AB	0.0657	0.0483	0.0398	0.0489	0.0443	0.0354
	$\hat{\lambda}_2$	*MSE*	0.0055	0.0032	0.0015	0.0031	0.0026	0.0013
		AB	0.0556	0.0454	0.0302	0.0424	0.0400	0.0285

As in this study, we consider two monitoring time patterns such as uniform and exponen-

Table 3.2: The mean square error and absolute bias for parameters set $(\theta_1, \theta_2) = (1.10, 1.12)$ and $(\lambda_1, \lambda_2) = \{(0.10, 0.11), (0.20, 0.25)\}$ under exponential monitoring points for varying sample size n.

			ML			*Bayes*		
(λ_1, λ_2)		n	80	150	250	80	150	250
$(0.10, 0.11)$	$\hat{\theta}_1$	*MSE*	0.0567	0.0422	0.0342	0.0492	0.0390	0.0325
		AB	0.2023	0.1770	0.1632	0.1876	0.1696	0.1587
	$\hat{\theta}_2$	*MSE*	0.0642	0.0522	0.0384	0.0561	0.0483	0.0365
		AB	0.2128	0.2002	0.1772	0.1982	0.1922	0.1727
	$\hat{\lambda}_1$	*MSE*	0.0015	0.0008	0.0005	0.0008	0.0006	0.0004
		AB	0.0277	0.0220	0.0173	0.0203	0.0192	0.0164
	$\hat{\lambda}_2$	*MSE*	0.0012	0.0007	0.0004	0.0007	0.0006	0.0004
		AB	0.0259	0.0215	0.0160	0.0195	0.0187	0.0149
$(0.20, 0.25)$	$\hat{\theta}_1$	*MSE*	0.0979	0.0858	0.0825	0.0806	0.0785	0.0781
		AB	0.2794	0.2720	0.2715	0.2515	0.2493	0.2366
	$\hat{\theta}_2$	*MSE*	0.1208	0.1195	0.1178	0.1200	0.1156	0.1124
		AB	0.3196	0.3028	0.2933	0.2887	0.2630	0.2573
	$\hat{\lambda}_1$	*MSE*	0.0059	0.0035	0.0021	0.0041	0.0035	0.0020
		AB	0.0572	0.0458	0.0359	0.0488	0.0447	0.0323
	$\hat{\lambda}_2$	*MSE*	0.0045	0.0025	0.0014	0.0032	0.0026	0.0016
		AB	0.0537	0.0394	0.0298	0.0456	0.0374	0.0252

Table 3.3: The average lengths and coverage probabilities for ACIs and HPD for parameters set $(\theta_1, \theta_2) = (1.10, 1.12)$ and $(\lambda_1, \lambda_2) = \{(0.10, 0.11), (0.20, 0.25)\}$ under uniform monitoring points for varying sample size n.

			ACI			*HPD*		
(λ_1, λ_2)		$n \rightarrow$	80	150	250	80	150	250
$(0.10, 0.11)$	$\hat{\theta}_1$	*AL*	0.6817	0.5191	0.3952	0.6708	0.5036	0.3741
		CP	0.8588	0.8800	0.8300	0.9000	0.9060	0.8720
	$\hat{\theta}_2$	*AL*	0.6826	0.5203	0.3958	0.6710	0.5137	0.3841
		CP	0.8450	0.8820	0.8756	0.8863	0.9040	0.8020
	$\hat{\lambda}_1$	AL	0.2694	0.1701	0.1229	0.2146	0.1628	0.1131
		CP	1.0000	0.9840	0.9840	1.0000	0.9940	0.9980
	$\hat{\lambda}_2$	*AL*	0.2720	0.1712	0.1233	0.2234	0.1676	0.1033
		CP	1.0000	0.9680	0.9620	1.0000	0.9900	0.9960
$(0.20, 0.25)$	$\hat{\theta}_1$	*AL*	0.6886	0.5249	0.4037	0.6725	0.5130	0.3841
		CP	0.8469	0.8000	0.8640	0.8996	0.8600	0.8732
	$\hat{\theta}_2$	*AL*	0.6978	0.5483	0.4036	0.6871	0.5197	0.3969
		CP	0.7804	0.8658	0.8450	0.8494	0.8722	0.8532
	$\hat{\lambda}_1$	*AL*	0.3905	0.2607	0.1899	0.3593	0.2577	0.1722
		CP	0.9987	0.9780	0.9760	1.0000	0.9920	0.9940
	$\hat{\lambda}_2$	*AL*	0.4096	0.2703	0.1993	0.3954	0.2496	0.1723
		CP	0.9862	0.9760	0.9740	1.0000	0.9940	0.9940

Table 3.4: The average lengths and coverage probabilities for ACIs and HPD for parameters set $(\theta_1, \theta_2) = (1.10, 1.12)$ and $(\lambda_1, \lambda_2) = \{(0.10, 0.11), (0.20, 0.25)\}$ under exponential monitoring points for varying sample size n.

			ACI			*HPD*		
(λ_1, λ_2)		$n \rightarrow$	80	150	250	80	150	250
$(0.10, 0.11)$	$\hat{\theta}_1$	*AL*	0.6391	0.4711	0.3660	0.6271	0.4529	0.3413
		CP	0.7966	0.8666	0.8578	0.8640	0.8785	0.8570
	$\hat{\theta}_2$	*AL*	0.6429	0.4668	0.3671	0.6272	0.4524	0.3415
		CP	0.8573	0.8572	0.8518	0.8160	0.8698	0.8638
	$\hat{\lambda}_1$	*AL*	0.1887	0.1311	0.1005	0.1727	0.1307	0.1001
		CP	0.9960	0.9920	0.9900	1.0000	0.9980	0.9960
	$\hat{\lambda}_2$	*AL*	0.1875	0.1338	0.1009	0.1767	0.1241	0.1001
		CP	1.0000	0.9860	0.9760	1.0000	0.9960	0.9820
$(0.20, 0.25)$	$\hat{\theta}_1$	*AL*	0.6564	0.4810	0.3717	0.6478	0.4750	0.3587
		CP	0.8505	0.9296	0.9131	0.9687	0.9468	0.9326
	$\hat{\theta}_2$	*AL*	0.6555	0.4701	0.3809	0.6464	0.4686	0.3764
		CP	0.8539	0.9163	0.9043	0.9360	0.9288	0.9309
	$\hat{\lambda}_1$	*AL*	0.2935	0.2103	0.1588	0.2900	0.2034	0.1484
		CP	0.9938	0.9788	0.9600	1.0000	0.9863	0.9625
	$\hat{\lambda}_2$	*AL*	0.3138	0.2274	0.1738	0.3032	0.2059	0.1620
		CP	0.9710	0.9713	0.9700	0.9979	0.9925	0.9925

tial distribution. So, in Table 3.1 and Table 3.3 results are shown under uniform monitoring time points. For exponential monitoring points outputs are reported in Table 3.2 and Table 3.4. From the simulated results given in Table 3.1 - Table 3.4, we can give some straightforward statements about the methodology. Overall the values of MSE and AB are small enough for all sample sizes as well as for both the approaches, classical and Bayesian. So we can state that both methods are performing quite well. Also, the values of MSE and AB decrease as the sample size increases which indicates the consistency of point estimators. It can be noted that as the proportion of non-recall and censoring in data increases the MSE and AB also increase. This is obvious because we get less information. But even at the 40% of such observation in the sample the derived methodology is performing well. Similar characteristics can be observed from interval estimates. As the sample size increases the AL of interval estimates decreases. Also, with the increase in the proportion of non-recall and censored observations in the sample the average lengths increase. The coverage probabilities are also indicating that the interval estimates are suitable good for parameters. For both, uniform and exponential monitoring time distribution, we are getting satisfactory results. We are confident enough that the derived methodology will also be successful for other monitoring patterns.

3.7 Real Data Analysis

This entire section is dedicated to the analysis of the menarche data set. The considered data is related to age at menarche which is a part of a project titled "Physical Growth, Body Composition and Nutritional Status of Bengali School aged Children, Adolescents and Young adults of Calcutta, India: Effects of Socioeconomic Factors on Secular Trends". The data was collected by the Biological Anthropology Unit of the Indian Statistical Institute in and around the city of Kolkata, India from 2005 to 2011. Menarche data was reported in the article by Salehabadi et al. (2015). One objective of this survey was to study the distribution of menarcheal age with associated factors. In this study, a total of 2195 girls were interviewed. At the time of the interview, i.e. at the monitoring time point, 1420 girls

have already experienced the events of interest, i.e. menarche. When it is asked to recall the event time, they were able to recall it at the day level, month level, or year level. After categorizing the data in such a way it is found that 443 girls recalled the age of menarche at the day level, 251 girls recalled age at menarche at month level and 190 girls recalled the age at menarche at the year level. Also, 536 such girls are available in the data who had not recalled the event time and hence their responses are categorized as non-recall observations. Besides all this, we have 775 girls for which the event did not happen till the monitoring time point. With the event time the order of birth, first or later birth, is also assigned to the individuals. By considering the event of interest as the onset of menarche, the data collection procedure is satisfying all the assumptions of the current status scenario. As assumed by Sukumaran and Dewan (2019), for the competing risk set up the order of birth can be considered as the cause for the menarche data. So here the first-born child corresponds to cause I and the later-born child as cause II.

For analysis purposes, we take recall phenomena at two levels. First, if an individual is able to recall the event time at day level then it is considered as recalled observation otherwise it will be a non-recalled one. Second, if an individual is able to recall exactly up to month level then it is considered as recall. In the month level case, the day level observations are also trimmed at the month level and included as recall observations. In Figure 3.1 and Figure 3.2, the bar plot respectively shows the distribution of exact event time whereas the box plots present the distribution of age at menarche at day and month level. In a paired box plot, the first belongs to the first-born child and the second to the later-born child. It can be noticed from these paired box plots that the median age at menarche of later-born children is larger than the first-born child.

To check the fitting of menarche data with exponential distribution, we have used the goodness of fit test based on chi-square statistic defined in article Hope (1968) and recently used by Koley and Dewanji (2021). Let us define the modified chi-square test statistic as follows

$$\chi^2_M = \sum_{i=1}^{n}\sum_{j=1}^{J} \frac{\left(\delta_i\epsilon_i - p^*_{ij}\right)^2}{p^*_{ij}\left(1-p^*_{ij}\right)},$$

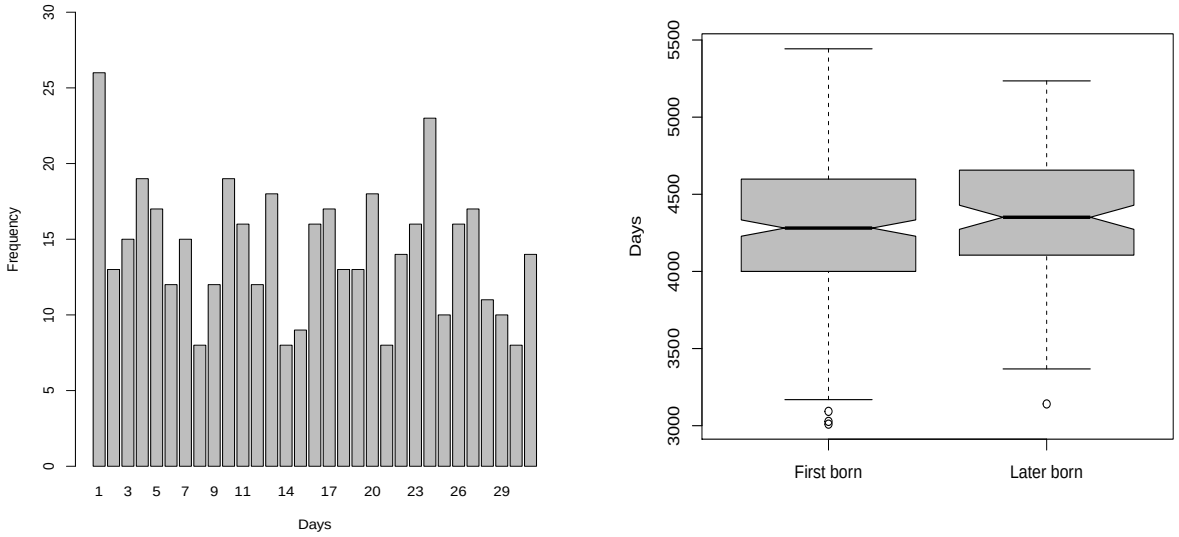

Figure 3.1: A set of bar-plot and box-plot of age at menarche data recalled at day level.

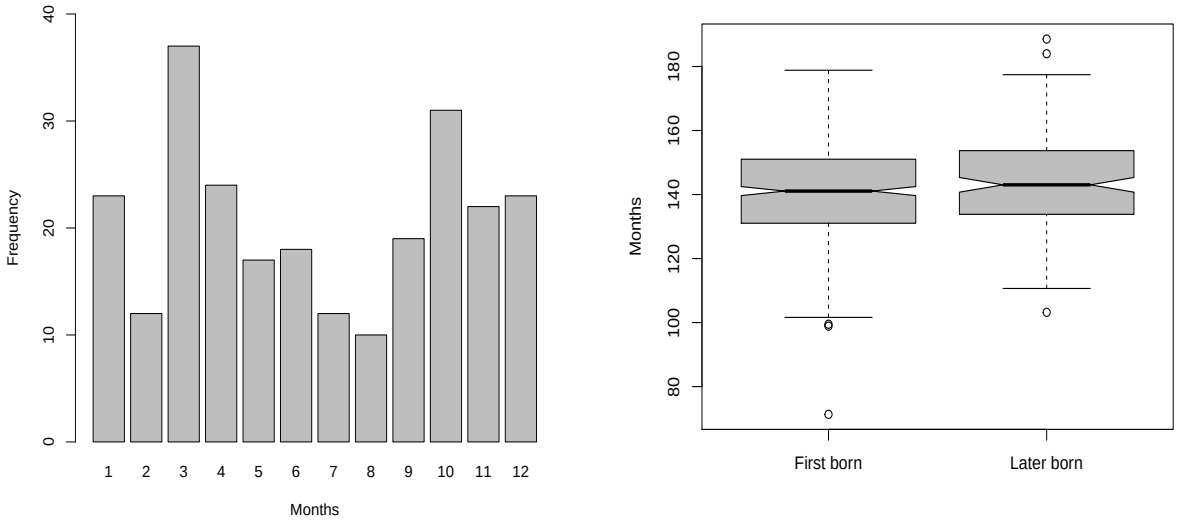

Figure 3.2: A set of bar-plot and box-plot of age at menarche data recalled at month level.

where p_{ij}^* takes different values from the likelihood as per indicator variables defined below

$$p_{ij}^* = \begin{cases} f(t_i,\eta_k)\prod\limits_{\substack{l=1\\l\neq k}}^{K}\bar{F}(t_i,\eta_l)\Big(1-\psi_k(s_i,t_i)\Big); & \text{if} \quad \delta_i=1 \;\&\; \epsilon_i=1, \\ \sum\limits_{k=1}^{K}\Big[\int\limits_0^{s_i} f(u,\eta_k)\prod\limits_{\substack{l=1\\l\neq k}}^{K}\bar{F}(u,\eta_l)\psi_k(s_i,u)du\Big]; & \text{if} \quad \delta_i=1 \;\&\; \epsilon_i=0, \\ \bar{F}(s_i,\eta); & \text{if} \quad \delta_i=0 \;\&\; \epsilon_i=0. \end{cases}$$

The calculation of the p value for the defined test statistic can be done as described ahead. The parameters of the model are first calculated using the proposed method and then the modified χ^2 statistic is calculated using observed data. Then a large number of datasets (say, 1000) is generated of the same sample size n with estimated model parameters and then we calculated modified chi-square statistics for each simulated data. The p value is calculated as the proportion of times the modified χ^2 statistic exceeds the observed value. The calculated p values for day level and month level datasets are found to be 0.89 and 0.90 which indicates that the considered exponential distribution is suitable for this data.

The point and interval estimates under classical and Bayesian frameworks for menarche data are reported in Table 3.5 to 3.6.

Table 3.5: The point and interval estimates for menarche data under ML and Bayes estimation when recalled at day level.

	ML	*Bayes*					
		Q_1	*Mean*	*Median*	Q_3	*SD*	*HPD*
$\hat{\theta}_1$	0.0606 (0.0576, 0.0636)	0.0583	0.0607	0.0613	0.0638	0.0045	(0.0515, 0.0687)
$\hat{\theta}_2$	0.0578 (0.0540, 0.0592)	0.0554	0.0580	0.0583	0.0598	0.0048	(0.0482, 0.0623)
$\hat{\lambda}_1$	0.1387 (0.1265, 0.1509)	0.1021	0.1906	0.1620	0.2586	0.1059	(0.0617, 0.4061)
$\hat{\lambda}_2$	0.1751 (0.1604, 0.1898)	0.1217	0.3096	0.2141	0.4055	0.2567	(0.0674, 0.6921)

A summary statistics of first-born and later-born girls based on exactly recall observations along with the point and interval estimates of median age using ML and Bayesian estimation methods for the age at menarche (in years) are given in Table 3.7. The confidence intervals for the median age at menarche is calculated by the method given in article Abu-Shawiesh (2010). As raw data is showing that the average menarcheal age of first-born

Table 3.6: The point and interval estimates for menarche data under ML and Bayes estimation when recalled at month level.

	ML	*Bayes*					
		Q_1	*Mean*	*Median*	Q_3	*SD*	*HPD*
$\hat{\theta}_1$	0.0663 (0.0635, 0.0691)	0.0642	0.0665	0.0669	0.0692	0.0039	(0.0582, 0.0736)
$\hat{\theta}_2$	0.0584 (0.0545, 0.0620)	0.0556	0.0588	0.0593	0.0625	0.0042	(0.0525, 0.0698)
$\hat{\lambda}_1$	0.0610 (0.0547, 0.0673)	0.0531	0.0649	0.0647	0.0768	0.0146	(0.0400, 0.0898)
$\hat{\lambda}_2$	0.1117 (0.1005, 0.1229)	0.0888	0.1353	0.1261	0.1774	0.0545	(0.0558, 0.2366)

Table 3.7: The menarche data summary statistics and estimates for median age at menarche with 95% confidence intervals when recalled at (i) day level and (ii) month level.

			Summary Statistics					*Median Age (in years)*	
Level	(n_1, n_2, n_{nr}, n_c)	Born	*Min*	*Max*	*SD*	*Mean*	Q_2	*ML*	*Bayes*
Day	(339, 129, 952, 775)	First	8.24	14.90	1.23	11.74	11.72	11.44 (10.97,11.93)	11.42 (10.96,11.91)
		Later	8.60	14.33	1.08	11.93	11.91	11.99 (11.51,12.51)	11.95 (11.47,12.47)
Month	(536, 202, 682, 775)	First	5.94	14.90	1.23	11.74	11.75	10.45 (10.03,10.91)	10.42 (10.00,10.87)
		Later	8.60	15.72	1.19	11.96	11.90	11.87 (11.39,12.38)	11.79 (11.31,12.30)

is lower than later-born, we get similar results based on proposed estimators. Also, it can be noticed that at day level study the average age of menarche is slightly different than the month level study. It is because at day level study we are more precise about the exact information but at the same time, non-recall observations are comparatively higher than the month level study. For this data, Salehabadi et al. (2015) reported the average age at menarche to be 11.78 years whereas Zeglen et al. (2020) found it 11.80 years. But in both the studies data are not analyzed under competing risks setup. Also, there are several studies available in demography that report the average age at menarche in the range of 9 to 16 years. We can conclude that results obtained from our model are not far away from results reported in earlier studies and at the same time it is able to give a new direction for further studies.

3.8 Conclusions and Discussions

In this chapter, the latent failure approach is developed for modeling the recall-based data under a competing risks framework. The non-recall probability is modeled using the assumption of difficulty in recalling with passes of time. ML estimates are obtained with the help of the nested E-M algorithm and for interval estimates, the observed Fisher information matrix is obtained using the missing information principle. The study is further extended to the Bayesian paradigm under suitable conjugate priors. Under this approach, point, and interval estimates are obtained using posterior samples which are generated with the help of the Gibbs sampling method. A simulation study is done for varying proportions of non-recall and censored observations under the consideration of uniform and exponential monitoring patterns. The results are found to be satisfactory under both classical and Bayesian approaches. Finally, menarche data is analyzed using proposed methodologies.

The concern of this study is to deal with the non-recall observations in current status data. Hence the choice of the functional form of non-recall probability is an important issue and should be dealt with carefully. Since the memory fades with the passes of time, the non-recall probability is taken as a function of the difference between monitoring time and event time. Some well-known functional forms of non-recall probability which are suggested in the literature are exponential form, piecewise constant function and multinomial logit model. The fuzzy approach may also be a good choice for modeling the non-recall probability. The proposed analysis can further be extended with any other distribution such as Weibull, log-normal, and inverse Gaussian. Further, the inclusion of covariates with existing event time distribution may be considered.

3.9 Appendix

Two Causes Set Up

We have developed the whole methodology for K causes in earlier sections. But in Section 3.7, two causes set up has been utilized for real data analysis. So to keep things simple and straightforward for researchers, we are presenting the procedure briefly by considering $K = 2$ in this Section.

Data Structure and Construction of Likelihood

Let n_1, n_2 and n_{nr} denotes the number of individual expose due to cause 1, cause 2 and non-recall observations, respectively. Now for the parameter $\Theta = (\theta_1, \theta_2)$ and $\Lambda = (\lambda_1, \lambda_2)$ the likelihood function can be written as

$$\begin{aligned} L(\Theta, \Lambda|d) &= \theta_1^{n_1}\theta_2^{n_2} \prod_{i=1}^{n_1} \exp\{-(\theta_1+\theta_2)t_i\} \exp\{-\lambda_1(s_i - t_i)\} \\ &\quad \prod_{i=1}^{n_2} \exp\{-(\theta_1+\theta_2)t_i\} \exp\{-\lambda_2(s_i - t_i)\} \prod_{i=1}^{n_c} \exp\{-(\theta_1+\theta_2)s_i\} \\ &\quad \prod_{i=1}^{n_{nr}} \left[\int_0^{s_i} \theta_1 \exp\{-(\theta_1+\theta_2)u\} \Big(1 - \exp\{-\lambda_1(s_i - u)\}\Big) du \right. \\ &\quad \left. + \int_0^{s_i} \theta_2 \exp\{-(\theta_1+\theta_2)u\} \Big(1 - \exp\{-\lambda_2(s_i - u)\}\Big) du \right] \end{aligned} \tag{3.18}$$

Expectation-Maximization Implementation

For all non-recall observation, let introduce a latent variable $z_i|(i = 1, 2, \cdots, n_{nr})$, from Bernoulli distribution. The probability of success can be defined as

$$P_i = \frac{I_1(s_i)}{I_1(s_i) + I_2(s_i)}, \tag{3.19}$$

where, $I_1(.)$ and $I_2(.)$ can be given by:

$$I_1(s_i) = \int_0^{s_i} \theta_1 \exp\{-(\theta_1+\theta_2)u\} \Big(1 - \exp\{-\lambda_1(s_i - u)\}\Big) du,$$

$$I_2(s_i) = \int_0^{s_i} \theta_2 \exp\{-(\theta_1+\theta_2)u\}\Big(1-\exp\{-\lambda_2(s_i-u)\}\Big)du.$$

By using the above idea of introducing z_i the sum terms reduced into product form. Also for integral part a left truncated time point say, t_i^*, is introduced following equivalent failure approach. For non-recall probabilities two latent variable u_i^* and u_i^* are introduced following exponential distribution with means $\frac{1}{\lambda_1}$ and $\frac{1}{\lambda_2}$ truncated at $(s_i - t_i^*)$ given $z_i = 1$ and $z = 0$ respectively. Here, after introducing the latent variable z_i, u_i^* and v_i^*, missing data vector is denoted by $d_i^* = (z_i, t_i^*, u_i^*, v_i^*)$. In light of observed and missing data, the complete likelihood is written as

$$\begin{aligned} L_c(\Theta,\Lambda|d,d^*) = {} & \theta_1^{(n_1+Z^*)}\theta_2^{(n_2+n_{nr}-Z^*)}\lambda_1^{Z^*}\lambda_2^{(n_{nr}-Z^*)}\prod_{i=1}^{n_1}\exp\{-(\theta_1+\theta_2)t_i\}\exp\{-\lambda_1(s_i-t_i)\} \\ & \prod_{i=1}^{n_2}\exp\{-(\theta_1+\theta_2)t_i\}\exp\{-\lambda_2(s_i-t_i)\} \\ & \prod_{i=1}^{n_{nr}}\Big[\exp\{-(\theta_1+\theta_2)t_i^*\}\exp\{-\lambda_1 u_i^*\}\Big]^{z_i}\prod_{i=1}^{n_c}\exp\{-(\theta_1+\theta_2)s_i\} \\ & \prod_{i=1}^{n_{nr}}\Big[\exp\{-(\theta_1+\theta_2)t_i^*\}\exp\{-\lambda_2 v_i^*\}\Big]^{1-z_i}. \end{aligned} \tag{3.20}$$

Taking logarithm of (3.20), we get the log-likelihood of complete data as follows

$$\begin{aligned} l_c(\Theta,\Lambda|d,d^*) = {} & (n_1+Z^*)\ln(\theta_1) + (n_2+n_{nr}-Z^*)\ln(\theta_2) + Z^*\ln(\lambda_1) + (n_{nr}-Z^*)\ln(\lambda_2) \\ & - \theta_1\sum_{i=1}^{n_1}t_i - \theta_2\sum_{i=1}^{n_1}t_i - \lambda_1\sum_{i=1}^{n_1}(s_i-t_i) - \theta_1\sum_{i=1}^{n_2}t_i - \theta_2\sum_{i=1}^{n_2}t_i - \lambda_2\sum_{i=1}^{n_2}(s_i-t_i) \\ & - \theta_1\sum_{i=1}^{n_{nr}}z_i t_i^* - \theta_2\sum_{i=1}^{n_{nr}}z_i t_i^* - \lambda_1\sum_{i=1}^{n_{nr}}z_i u_i^* - \theta_1\sum_{i=1}^{n_{nr}}(1-z_i)t_i^* - \theta_2\sum_{i=1}^{n_{nr}}(1-z_i)t_i^* \\ & - \lambda_2\sum_{i=1}^{n_{nr}}(1-z_i)v_i^* - \theta_1\sum_{i=1}^{n_c}s_i - \theta_2\sum_{i=1}^{n_c}s_i. \end{aligned} \tag{3.21}$$

For the estimation purpose, we define a new quantity $Q(\Theta,\Lambda|\Theta^{(m)},\Lambda^{(m)})$ for given values of parameters and taking expectation of complete log-likelihood with respect to the latent variables such as

$$
\begin{aligned}
Q\Big(\Theta,\Lambda|\Theta^{(m)},\Lambda^{(m)}\Big) &= \Big(n_1+P^{(m)}\Big)\ln(\theta_1)+\Big(n_2+n_{nr}-P^{(m)}\Big)\ln(\theta_2)+P^{(m)}\ln(\lambda_1)\\
&+\Big(n_{nr}-P^{(m)}\Big)\ln(\lambda_2)-(\theta_1+\theta_2)\sum_{i=1}^{n_1}t_i-\lambda_1\sum_{i=1}^{n_1}(s_i-t_i)\\
&-(\theta_1+\theta_2)\sum_{i=1}^{n_2}t_i-\lambda_2\sum_{i=1}^{n_2}(s_i-t_i)-(\theta_1+\theta_2)\sum_{i=1}^{n_{nr}}P_i\xi_4\Big(t_i^*,\theta^{(m)}\Big)\\
&-\lambda_1\sum_{i=1}^{n_{nr}}P_i\xi_6\Big(u_i^*,\lambda^{(m)}\Big)-(\theta_1+\theta_2)\sum_{i=1}^{n_c}s_i-\theta_1\sum_{i=1}^{n_{nr}}(1-P_i)\xi_5\Big(t_i^*,\theta^{(m)}\Big)\\
&-\theta_2\sum_{i=1}^{n_{nr}}(1-P_i)\xi_5\Big(t_i^*,\theta^{(m)}\Big)-\lambda_2\sum_{i=1}^{n_{nr}}(1-P_i)\xi_7\Big(v_i^*,\lambda^{(m)}\Big). \qquad (3.22)
\end{aligned}
$$

Since, z_i is a Bernoulli variate, therefore, $E(z_i)=P_i$ and $E(Z^*)=E\Big(\sum_{i=1}^{n_{nr}}z_i\Big)=P(say)$. The expected values used in E step at an iteration are given by

$$
\begin{aligned}
\xi_4(t_i^*,\theta_1) &= E\Big(t_i^*\big|t_i^*<s_i,k=1,\theta_1\Big)=\frac{1}{\theta_1}\left[1-\frac{\theta_1 s_i\exp\{-\theta_1 s_i\}}{1-\exp\{-\theta_1 s_i\}}\right],\\
\xi_5(t_i^*,\theta_2) &= E\Big(t_i^*\big|t_i^*<s_i,k=2,\theta_2\Big)=\frac{1}{\theta_2}\left[1-\frac{\theta_2 s_i\exp\{-\theta_2 s_i\}}{1-\exp\{-\theta_2 s_i\}}\right],\\
\xi_6(u_i^*,\lambda_1) &= E\Big(u_i^*\big|u_i^*<(s_i-t_i^*),k=1,\lambda_1\Big)=\frac{1}{\lambda_1}\left[\frac{1-\Big(1+\lambda_1(s_i-t_i^*)\Big)\exp\{-\lambda_1(s_i-t_i^*)\}}{1-\exp\{-\lambda_1(s_i-t_i^*)\}}\right],\\
\xi_7(v_i^*,\lambda_2) &= E\Big(v_i^*\big|v_i^*<(s_i-t_i^*),k=2,\lambda_2\Big)=\frac{1}{\lambda_2}\left[\frac{1-\Big(1+\lambda_2(s_i-t_i^*)\Big)\exp\{-\lambda_2(s_i-t_i^*)\}}{1-\exp\{-\lambda_2(s_i-t_i^*)\}}\right].
\end{aligned}
$$

Now in order to maximize log-likelihood (the M-step), we differentiated (3.22) partially with respect to θ_1, θ_2, λ_1 and λ_2 and after some calculations, get the following expressions:

$$
\hat{\theta}_1^{(m+1)}=\frac{\Big(n_1+P^{(m)}\Big)}{\sum_{i=1}^{n_1}t_i+\sum_{i=1}^{n_2}t_i+\sum_{i=1}^{n_{nr}}P_i^{(m)}\xi_4\Big(t_i^*,\theta_1^{(m)}\Big)+\sum_{i=1}^{n_{nr}}\Big(1-P_i^{(m)}\Big)\xi_5\Big(t_i^*,\theta_2^{(m)}\Big)+\sum_{i=1}^{n_c}s_i}, \qquad (3.23)
$$

$$
\hat{\theta}_2^{(m+1)}=\frac{\Big(n_2+n_{nr}-P^{(m)}\Big)}{\sum_{i=1}^{n_1}t_i+\sum_{i=1}^{n_2}t_i+\sum_{i=1}^{n_{nr}}P_i^{(m)}\xi_4\Big(t_i^*,\theta_1^{(m)}\Big)+\sum_{i=1}^{n_{nr}}\Big(1-P_i^{(m)}\Big)\xi_5\Big(t_i^*,\theta_2^{(m)}\Big)+\sum_{i=1}^{n_c}s_i}, \qquad (3.24)
$$

$$
\hat{\lambda}_1^{(m+1)}=\frac{P^{(m)}}{\sum_{i=1}^{n_1}(s_i-t_i)+\sum_{i=1}^{n_{nr}}P_i^{(m)}\xi_6\Big(u_i^*,\lambda_1^{(m)}\Big)}, \qquad (3.25)
$$

$$
\hat{\lambda}_2^{(m+1)}=\frac{\Big(n_{nr}-P^{(m)}\Big)}{\sum_{i=1}^{n_2}(s_i-t_i)+\sum_{i=1}^{n_{nr}}\Big(1-P_i^{(m)}\Big)\xi_7\Big(v_i^*,\lambda_2^{(m)}\Big)}. \qquad (3.26)
$$

To observe the interval estimators, the expressions to get the observed Fisher information

matrix for two component setup can be given as

$$\frac{\partial Q}{\partial \theta_1} = \frac{n_1 + P}{\theta_1} - \sum_{i=1}^{n_1} t_i - \sum_{i=1}^{n_2} t_i - \sum_{i=1}^{n_{nr}} P_i \xi_4(t_i^*, \theta_1) - \sum_{i=1}^{n_{nr}} (1 - P_i)\xi_5(t_i^*, \theta_2) - \sum_{i=1}^{n_c} s_i,$$

$$\frac{\partial Q}{\partial \theta_2} = \frac{n_2 + n_{nr} - P}{\theta_2} - \sum_{i=1}^{n_1} t_i - \sum_{i=1}^{n_2} t_i - \sum_{i=1}^{n_{nr}} P_i \xi_4(t_i^*, \theta_1) - \sum_{i=1}^{n_{nr}} (1 - P_i)\xi_5(t_i^*, \theta_2) - \sum_{i=1}^{n_c} s_i,$$

$$\frac{\partial Q}{\partial \lambda_1} = \frac{P}{\lambda_1} - \sum_{i=1}^{n_1} (s_i - t_i) - \sum_{i=1}^{n_{nr}} P_i \xi_6(u_i^*, \lambda_1)$$

$$\frac{\partial Q}{\partial \lambda_2} = \frac{n_{nr} - P}{\lambda_2} - \sum_{i=1}^{n_2} (s_i - t_i) - \sum_{i=1}^{n_{nr}} (1 - P_i)\xi_7(v_i^*, \lambda_2)$$

$$\frac{\partial^2 Q}{\partial \theta_1^2} = -\frac{n_1 + P}{\theta_1^2}, \quad \frac{\partial^2 Q}{\partial \theta_2^2} = -\frac{n_2 + n_{nr} - P}{\theta_2^2}, \quad \frac{\partial^2 Q}{\partial \lambda_1^2} = -\frac{P}{\lambda_1^2}, \quad \frac{\partial^2 Q}{\partial \lambda_2^2} = -\frac{n_{nr} - P}{\lambda_2^2}.$$

The rest of the terms comes out to be zero. If $(\hat{\Theta}, \hat{\Lambda})$ be the ML estimates of (Θ, Λ) then the corresponding interval estimates can be written as

$$\hat{\Theta} \pm z_{\alpha/2} SE(\hat{\Theta}) \text{ and } \hat{\Lambda} \pm z_{\alpha/2} SE(\hat{\Lambda}).$$

Bayesian Estimation

For the Bayesian estimation in case of two causes with Gamma distribution as priors the joint prior density is written as

$$\pi(\Theta, \Lambda) \propto \theta_1^{a_1 - 1} \theta_2^{a_2 - 1} \lambda_1^{a_3 - 1} \lambda_2^{a_4 - 1} \exp\{-\theta_1 b_1 - \theta_2 b_2 - \lambda_1 b_3 - \lambda_2 b_4\}; \; a_l, b_l > 0, l = 1, 2, 3, 4 \tag{3.27}$$

Using the conditional likelihood given in (3.20), joint prior given in (3.27) and Bayes theorem, joint posterior distribution is given by

$$\begin{aligned}\Pi(\Theta,\Lambda|d,d^*) \propto & \theta_1^{(n_1+Z^*+a_1-1)}\theta_2^{(n_2+n_{nr}-Z^*+a_2-1)}\lambda_1^{(Z^*+a_3-1)}\lambda_2^{(n_{nr}-Z^*+a_4-1)} \\ & \prod_{i=1}^{n_1}\exp\{-(\theta_1+\theta_2)t_i-\lambda_1(s_i-t_i)\}\prod_{i=1}^{n_2}\exp\{-(\theta_1+\theta_2)t_i-\lambda_2(s_i-t_i)\} \\ & \prod_{i=1}^{n_{nr}}\Big[\exp\{-(\theta_1+\theta_2)t_i^*-\lambda_1u_i^*\}\Big]^{z_i}\Big[\exp\{-(\theta_1+\theta_2)t_i^*-\lambda_2v_i^*\}\Big]^{1-z_i} \\ & \prod_{i=1}^{n_c}\exp\{-(\theta_1+\theta_2)s_i\}\exp\{-\theta_1b_1-\theta_2b_2-\lambda_1b_3-\lambda_2b_4\} \qquad (3.28)\end{aligned}$$

The full conditionals for θ_1, θ_2, λ_1 and λ_2 are obtained form posterior distribution (3.28), which are given below

$$\theta_1|\{\Theta,\Lambda\}_{(-\theta_1)} \sim \mathcal{G}\left(n_1+Z^*+a_1,\ \sum_{i=1}^{n_1}t_i+\sum_{i=1}^{n_2}t_i+\sum_{i=1}^{n_{nr}}z_it_i^*+\sum_{i=1}^{n_{nr}}(1-z_i)t_i^*+\sum_{i=1}^{n_c}s_i+b_1\right) \qquad (3.29)$$

$$\theta_2|\{\Theta,\Lambda\}_{(-\theta_2)} \sim \mathcal{G}\left(n_2+n_{nr}-Z^*+a_2,\ \sum_{i=1}^{n_1}t_i+\sum_{i=1}^{n_2}t_i+\sum_{i=1}^{n_{nr}}z_it_i^*+\sum_{i=1}^{n_{nr}}(1-z_i)t_i^*+\sum_{i=1}^{n_c}s_i+b_2\right) \qquad (3.30)$$

$$\lambda_1|\{\Theta,\Lambda\}_{(-\lambda_1)} \sim \mathcal{G}\left(Z^*+a_3,\ \sum_{i=1}^{n_1}(s_i-t_i)+\sum_{i=1}^{n_{nr}}z_iu_i^*+b_3\right) \qquad (3.31)$$

$$\lambda_2|\{\Theta,\Lambda\}_{(-\lambda_2)} \sim \mathcal{G}\left(n_{nr}-Z^*+a_4,\ \sum_{i=1}^{n_2}(s_i-t_i)+\sum_{i=1}^{n_{nr}}(1-z_i)v_i^*+b_4\right) \qquad (3.32)$$

Here from (3.29) to (3.32), we can see that the parameters θ_1, θ_2, λ_1 and λ_2 are conditionally independent from each other for given z_i, t_i^*, u_i^* and v_i^*. To obtain the estimators the sample can be generated simultaneously by using the Gibbs sampling method. For $w \sim \mathcal{U}(0,1)$ the expressions which are used in generation of data from t_i^*, u_i^* and v_i^* are given by

$$\begin{aligned}t_i^* &= -\frac{1}{\theta_k}\ln\Big[1-w_i\Big(1-\exp\{-\theta_ks_i\}\Big)\Big], \\ u_i^* &= -\frac{1}{\lambda_1}\ln\Big[1-w_i\Big(1-\exp\{-\lambda_1(s_i-t_i^*)\}\Big)\Big], \\ v_i^* &= -\frac{1}{\lambda_2}\ln\Big[1-w_i\Big(1-\exp\{-\lambda_2(s_i-t_i^*)\}\Big)\Big].\end{aligned}$$

Chapter 4

Analysis of Recall-Based Competing Risk Data using Weibull Distribution

4.1 Introduction

Time to event studies is of great interest and found their applicability in the fields of clinical trials, survival analysis, reliability engineering and cross-sectional studies. Data based on time to events can be obtained in a study by observing the individuals over time. For example, in clinical trials, a patient is followed up for a time until the event of interest occurs or the individual may withdraw himself/herself from the study due to some personal reasons. In survival analysis, the living beings are followed up in an experiment to get the time to event. In life testing experiments, individuals under test generate time to event data. Moreover, time to event data may be collected in cross-sectional studies and sample surveys by following individuals over time. If the individuals are followed throughout their lifetimes, the complete data may be observed. But in real-life, to follow an individual over the life span is not possible all the time. Henceforth, the collected datasets may have censored, truncated, or missing observations.

There are two well-known censoring schemes Type-I and Type-II, that generate in real-life situations, due to restrictions on time of experiment and number of failures. Besides these, there may be situations when individuals are checked at pre-scheduled periodic monitoring times to see the status of the disease or any other event of interest. If the status of an event

of interest is found changed in the current visit, then it is considered that it is changed somewhere between the previous and current visit, but the exact time to event is not known. Such types of studies generate interval-censored data. Right and Left censored data can be treated as special cases of interval-censored data. Right censored data occurs in the case when the upper bound of the interval-censored window is infinity, while in the case of left-censored data the lower bound of the window becomes zero. Left censoring occurs when an event has happened prior to entering the study. For example, a permanent tooth has already emerged before a dental study that aims to estimate its appearance distribution. In the same study, interval censoring can be present if the permanent tooth is present in the mouth in the current monitoring which was not at the last monitoring point. A similar case is true for the diagnosis of AIDS, which is diagnosed at the current monitoring point which was absent at the last visit. For more on interval censoring one can refer to articles Finkelstein (1986) and Sun (2007).

The theory of competing risk arises in the case when the event of interest may be exposed to more than one risk/cause at any particular time point. Several fundamental studies on competing risk data are done using the exponential distribution due to its applicability in reliability theory and survival analysis. The exponential distribution has a constant hazard rate and can be used to lifetime data possessing such kind of nature. The exponential distribution can be used to model the lifetime of electronic devices, a lifetime of youth period with respect to death from any disease and duration between arrival times. Weibull distribution is able to capture the decreasing, constant and increasing hazard rates when the shape parameter is less than 1, equal to 1 and greater than 1 respectively. Thus, it provides more flexibility to model the lifetime data. In literature, under competing risks setup, several fundamental types of research considering Weibull as a lifetime distribution are done. Flehinger et al. (2002b) developed a parametric model for the life testing situation in which the systems are individual to failure from the independent competing risks. The study is explained assuming Weibull distribution for the failure times of competing risks. Pascual (2007) developed the model for ALT when there are two competing risks that are dependent on one accelerating factor. The failure times are assumed to follow the Weibull distribution

with known and common shape parameters. Expressions for the test plan criteria based on the Fisher information matrix are presented. Finally, the proposed methodology is applied to the ALT of Class-H insulation for motorettes data, where the temperature is considered as the accelerating factor. Pascual (2008) studied the planning of accelerated life testing (ALT) models under competing risks assuming Weibull distribution for time to failure due to a specific risk. The method is applied to Class-B insulation of motorettes. Pareek et al. (2009) considered Weibull distribution with progressive censoring under competing risk setup assuming same shape and different scales for different causes. The point and interval estimates are obtained. Further, the optimal censoring scheme is developed for the considered setup. Hashemi and Amiri (2011) analyzed the progressive Type-II censoring with the Weibull model for competing risks data assuming Binomial removals and obtained the MLE and confidence intervals for the unknown parameters. The simulation studies for different sample sizes are performed to access the performance of estimators. Progressively censored competing risks data analysis assuming the lifetime distributions as Weibull distribution is done by Kundu and Pradhan (2011). The authors assumed the latent cause of failures to be independent Weibull distributed with the common shape parameter, but different scale parameters. Under suitable choice of priors, the Bayesian analysis of the model is done and an optimal censoring scheme is developed under the Bayesian paradigm. Balakrishnan et al. (2015) studied the Weibull distribution for testing one-shot devices under competing risks. The point and interval estimates are obtained by using the E-M algorithm and the missing information principle respectively. For illustration purposes, modified Class-B insulation data is used. The Weibull distribution is assumed for the latent failure times with the same shape under adaptive type-I progressive hybrid censoring scheme by Ashour and Nassar (2017). The classical and Bayesian inferences are drawn for the considered setup. Liu and Shi (2017) considered the step-stress model with competing risks for failure from Weibull distribution under progressive Type-II censoring. Under the proportional hazard model, the authors obtained the MLE and confidence intervals. The Bayesian analysis is done with different choices of prior distributions. The performance of estimators is shown with help of simulation studies. Chacko and Mohan (2019) studied the Weibull distribution for the anal-

ysis of competing risk data under progressive type-II censoring by assuming the removals as binomial distribution and drew classical and Bayesian inferences. Baghestani and Hosseini-Baharanchi (2019) proposed an improper extension of Weibull distribution which consists of two distributions Weibull before a pre-specified time point and exponential for the tail of the time axis. The models are studied using the Bayesian approach. The simulation study is done to show the identifiability and appropriate convergence of the developed model. The proposed model is applied to the acute lymphoblastic leukemia dataset. Haghighi and Castagliola (2019) studied a control chart to monitor the Weibull shape parameter where the observations are censored due to competing risks. It is assumed that the failure occurs due to competing risks that are assumed to follow Weibull distributions with different shapes and scales. The control chart is developed to monitor the shape parameters. The constructed control limits depend on the sample size, the number of failures due to each risk and the desired stable average run length (ARL). The study is further extended to making a data setup where the parameters are estimated using the E-M algorithm. Under both cases without and with masking, the performance of the proposed chart is analyzed through simulation studies. Finally, real data is analyzed for showing the applicability of the proposed methodology.

This chapter aims to develop the mathematical models to deal with recall-based competing risk data assuming latent time to events due to different causes from Weibull distribution. In Section 4.2 basic data structure of recall-based current status data under competing risks setup and formation of likelihood is discussed. Section 4.3 deals with finding point and interval estimates through the E-M algorithm utilizing the missing information principle. Under the Bayesian paradigm, estimates are obtained assuming gamma prior which are discussed in Section 4.4. Here, a three-stage nested algorithm based on Gibbs Sampling is utilized for drawing samples from conditional posteriors. An extensive simulation study is performed in Section 4.5 to check the performance of proposed methodologies for different choices of parameters and sample sizes. In Section 4.6, for real data applications, the age at menarche is estimated for first born and later born girls. Finally, Section 4.7 is dedicated to the discussion and future scope of the study.

4.2 Data Structure and Likelihood

Here, we assume the same setup of recall-based data under competing risks scenario as discussed in Chapter 3. Under the same assumption and notations described in 3.2, we can define the likelihood for parameter vector θ in the light of observed data d as

$$L(\theta|d) = \prod_{i=1}^{n} \left[\prod_{k} \left\{ f(t_i, \theta_k) \prod_{\substack{l=1 \\ l\neq k}}^{K} \bar{F}(t_i, \theta_l)\{1 - \psi_k(s_i, t_i)\} \right\} \right]^{\delta_i \epsilon_i} \left[\sum_{k} \left\{ \int_0^{s_i} f(u, \theta_k) \prod_{\substack{l=1 \\ l\neq k}}^{K} \bar{F}(u, \theta_l)\psi_k(s_i, u)du \right\} \right]^{\delta_i(1-\epsilon_i)} \Big[\bar{F}(s_i, \theta)\Big]^{(1-\delta_i)}, \tag{4.1}$$

where, parameter vector $\theta = \{\theta_1, \theta_2, \ldots, \theta_K\}$ and $\psi_k(s,t)$ denotes the non-recall probability associated to k^{th} cause. Here, it can be noted that probability associated to different causes is not same and depends on the function of difference between monitoring time and time to event. Alternatively, we can define the probability of recall as $1 - \psi_k(s,t)$.

Using the same concept discussed in 3.2, the exponential functional form defined in equation 3.2 is considered for non-recall probability. The probability of recall decreases with rate λ_k as the difference between monitoring time and time to event becomes larger. Under this assumption, the likelihood (4.1) can be re-written as

$$L(\theta, \lambda|d) = \prod_{k=1}^{K} \prod_{i=1}^{n_k} \left[f(t_i, \theta_k) \prod_{\substack{l=1 \\ l\neq k}}^{K} \bar{F}(t_i, \theta_l) \exp\{-\lambda_k(s_i - t_i)\} \right] \prod_{i=1}^{n_{nr}} \left[\sum_{k=1}^{K} \left\{ \int_0^{s_i} f(u, \theta_k) \prod_{\substack{l=1 \\ l\neq k}}^{K} \bar{F}(u, \theta_l) \Big(1 - \exp\{-\lambda_k(s_i - u)\}\Big) du \right\} \right] \prod_{i=1}^{n_c} \Big[\bar{F}(s_i, \theta)\Big], \tag{4.2}$$

The likelihood function (4.2) is not easy for further calculations due to unknown quantities time and cause of occurrence of an event for non-recall observations. So, some alternative methods should be used for the estimation of unknown parameters. The role of the E-M algorithm comes in situations where data is missing due to many reasons such as censoring,

truncation, grouping or masking. The E-M algorithm is proposed by Dempster et al. (1977). In the next section, we proposed an E-M-based algorithm to solve the purpose.

4.3 Classical Inference

4.3.1 Expectation Maximization Algorithm

From (4.2), it can be observed the non-recall observations have partial information on the time to event and the cause of occurrence of the event is unknown. So, we treat them as a missing data problem and use the E-M algorithm for further analysis. As Kuo and Yang (2000) and Rai et al. (2021) used latent variable in case of masked data, in similar fashion for i^{th} observation under non-recall category, a latent variable $Z_i \sim \mathcal{MD}(1, p_{i1}, p_{i2}, \cdots, p_{iK})$ with probability of success is defined by

$$p_{ik} = \frac{I_k(s_i)}{\sum\limits_k I_k(s_i)}; \quad i = 1, 2, \ldots, n_{nr};\ k = 1, 2, \cdots, K \tag{4.3}$$

where, $I_k(.)$ is given by

$$I_k(s) = \int_0^s f(u, \eta_k) \prod_{\substack{l=1 \\ l \neq k}}^{K} \bar{F}(u, \eta_l) \psi_k(s, u) du.$$

With the help of latent variable Z, we assign the cause of occurrence to each individual lying under non-recall category. Further based on multinomial variable Z, likelihood (4.1) can be re-written as

$$L(\theta, \lambda | d) = \prod_{k=1}^{K} \prod_{i=1}^{n_k} \left[f(t_i, \theta_k) \prod_{\substack{l=1 \\ l \neq k}}^{K} \bar{F}(t_i, \theta_l) \exp\{-\lambda_k(s_i - t_i)\} \right] \prod_{i=1}^{n_c} \left[\bar{F}(s_i, \theta) \right]$$
$$\prod_{i=1}^{n_{nr}} \left[\prod_{k=1}^{K} \left\{ \int_0^{s_i} f(u, \theta_k) \prod_{\substack{l=1 \\ l \neq k}}^{K} \bar{F}(u, \theta_l) \Big(1 - \exp\{-\lambda_k(s_i - u)\} \Big) du \right\}^{z_i} \right], \tag{4.4}$$

Once the causes of occurrence for non-recall observations is detected next quantity remains time to event which lies in interval $(0, S)$. The principle of equivalent failure quantity as discussed in Chapter 3 is used to deal with it. Thus, in light of given causes of occurrence of event, a latent variable T^* is introduced, which lies in the interval $(0, S)$. For i^{th} individual having monitoring interval $(0, s_i)$ and given cause of occurrence of event, the conditional density of t^*_{ik} can be given by

$$f\Big(t^*_{ik}\big|t^*_{ik} \in (0, s_i), z_i, \theta_k\Big) = \frac{f(t^*_{ik}; \theta_k)}{1 - \bar{F}(s_i; \theta_k)}; \quad i = 1, 2, \ldots, n_{nr},\ k = 1, 2, \cdots, K. \tag{4.5}$$

Now with the help of conditional density (4.5), the likelihood function (4.4) can be re-written as

$$\begin{aligned} L(\theta, \lambda|d) = &\prod_{k=1}^{K}\prod_{i=1}^{n_k}\left[f(t_i, \theta_k)\prod_{\substack{l=1\\ l\neq k}}^{K}\bar{F}(t_i, \theta_l)\exp\{-\lambda_k(s_i - t_i)\}\right]\prod_{i=1}^{n_c}\Big[\bar{F}(s_i, \theta)\Big] \\ &\prod_{i=1}^{n_{nr}}\left[\prod_{k=1}^{K}\left\{f(t^*_{ik}, \theta_k)\prod_{\substack{l=1\\ l\neq k}}^{K}\bar{F}(t^*_{ik}, \theta_l)\Big(1 - \exp\{-\lambda_k(s_i - t^*_{ik})\}\Big)\right\}^{z_i}\right]. \end{aligned} \tag{4.6}$$

Based on two variables, Z and T^*, we introduce another latent variable, say W^*, to make the functional form of non-recall probability into more convenient form. The latent variable W^* follows exponential distribution with mean $1/\lambda_k$ and truncated at $(S - T^*)$. Hence, the missing data vector can be written as $\underline{D}^* = (Z, T^*, W^*)$ be the missing data vector. Then $d^*_i = (z_i, t^*_{ik}, w^*_{ik});\ i = 1, 2, \ldots, n_{nr}, k = 1, 2, \ldots, K$ denotes the observed value for i^{th} individual falling in the non-recall category.

Let us assume that the k^{th} cause of occurrence of event follows Weibull distribution with shape α_k and scale β_k, i.e., $T \sim \mathcal{W}(\alpha_k, \beta_k),\ k = 1, 2, \cdots, K$. The set of parameters can be represented as $\alpha = \{\alpha_1, \alpha_2, \cdots, \alpha_K\}$, and $\beta = \{\beta_1, \beta_2, \cdots, \beta_K\}$. Now, let us denote the set of parameter vector by $\Theta = (\alpha, \beta, \lambda)$. Now under the Weibull distribution and complete

data set (d, d^*), the complete likelihood function can be written as

$$L_c(\Theta|d,d^*) = \prod_{k=1}^{K}\prod_{i=1}^{n_k}\left[\alpha_k\beta_k t_i^{\alpha_k-1}\exp\{-\beta_k t_i^{\alpha_k}\}\prod_{\substack{l=1\\l\neq k}}^{K}\exp\{-\beta_l t_i^{\alpha_l}\}\exp\{-\lambda_k(s_i-t_i)\}\right]$$
$$\prod_{i=1}^{n_{nr}}\prod_{k=1}^{K}\left[\alpha_k\beta_k\lambda_k {t_{ik}^*}^{\alpha_k-1}\exp\{-\beta_k {t_{ik}^*}^{\alpha_k}\}\prod_{\substack{l=1\\l\neq k}}^{K}\exp\{-\beta_l {t_{ik}^*}^{\alpha_l}\}\exp\{-\lambda_k w_{ik}^*\}\right]^{z_i}$$
$$\prod_{i=1}^{n_c}\left[\exp\left\{-\sum_{k=1}^{K}\beta_k s_i^{\alpha_k}\right\}\right]. \tag{4.7}$$

The log-likelihood function of complete data is obtained by taking logarithm of (4.7) as below

$$l_c(\Theta|d,d^*) = \sum_{k=1}^{K}\sum_{i=1}^{n_k}\left[\ln(\alpha_k)+\ln(\beta_k)+(\alpha_k-1)\ln(t_i)-\beta_k t_i^{\alpha_k}-\sum_{\substack{l=1\\l\neq k}}^{K}\beta_l t_i^{\alpha_l}-\lambda_k(s_i-t_i)\right]$$
$$+\sum_{i=1}^{n_{nr}}\sum_{k=1}^{K}z_i\left[\ln(\alpha_k)+\ln(\beta_k)+\ln(\lambda_k)+(\alpha_k-1)\ln(t_{ik}^*)-\beta_k {t_{ik}^*}^{\alpha_k}-\sum_{\substack{l=1\\l\neq k}}^{K}\beta_l {t_{ik}^*}^{\alpha_l}\right]$$
$$+\sum_{i=1}^{n_{nr}}\sum_{k=1}^{K}z_i\left[-\lambda_k w_{ik}^*\right]-\sum_{k=1}^{K}\sum_{i=1}^{n_c}\beta_k s_i^{\alpha_k}. \tag{4.8}$$

In order to apply the E step of E-M algorithm, let us define the quantity $Q\Big(\Theta|\hat{\Theta}^{(m)}\Big) = E\Big[l_c(\Theta|d,d^*)|d,\hat{\Theta}^{(m)}\Big]$, where $\hat{\Theta}^{(m)}$ be the estimate of parameters at m^{th} iteration. A nested expectation is taken with respect to the introduced latent variables, Z, T^* and W^*. By using the nested expectation, the expression of $Q\Big(\Theta|\hat{\Theta}^{(m)}\Big)$ can be derived as follows

$$Q\Big(\Theta|\hat{\Theta}^{(m)}\Big) = E_{W^*}\Big[E_{T^*}\Big[E_Z\big[l_c(\Theta|d,d^*)|d,\hat{\Theta}^{(m)}\big]\Big]\Big]$$
$$=\sum_{k=1}^{K}\sum_{i=1}^{n_k}\left[\ln(\alpha_k)+\ln(\beta_k)+(\alpha_k-1)\ln(t_i)-\beta_k t_i^{\alpha_k}-\sum_{\substack{l=1\\l\neq k}}^{K}\beta_l t_i^{\alpha_l}-\lambda_k(s_i-t_i)\right]$$
$$+\sum_{i=1}^{n_{nr}}\sum_{k=1}^{K}E\big[z_i\big]\left[\ln(\alpha_k)+\ln(\beta_k)+\ln(\lambda_k)+(\alpha_k-1)E\Big[\ln(t_{ik}^*)\mid t_{ik}^*<s_i\Big]\right]$$
$$+\sum_{i=1}^{n_{nr}}\sum_{k=1}^{K}E\big[z_i\big]\left[-\beta_k E\Big[{t_{ik}^*}^{\alpha_k}\mid t_{ik}^*<s_i\Big]-\sum_{\substack{l=1\\l\neq k}}^{K}\beta_l E\Big[{t_{ik}^*}^{\alpha_l}\mid t_{ik}^*<s_i\Big]\right]$$

$$+\sum_{i=1}^{n_{nr}}\sum_{k=1}^{K} E\big[z_i\big]\Bigg[-\lambda_k E\Big[w_{ik}^* \mid w_{ik}^* < (s_i - t_{ik}^*)\Big]\Bigg] - \sum_{k=1}^{K}\sum_{i=1}^{n_c}\beta_k s_i^{\alpha_k}. \tag{4.9}$$

Now for finding the estimators of unknown parameters, (4.9) is differentiated with respect to α_k, β_k and λ_k and thus we get

$$\frac{\partial Q}{\partial \alpha_k} = \frac{n_k}{\alpha_k} + \sum_{i=1}^{n_k}\ln(t_i) - \beta_k\sum_{i=1}^{n_k} t_i^{\alpha_k}\ln(t_i) + \frac{1}{\alpha_k}\sum_{i=1}^{n_{nr}} E\big[z_i\big] + \sum_{i=1}^{n_{nr}} E\big[z_i\big]E\Big[\ln(t_{ik}^*) \mid t_{ik}^* < s_i\Big]$$
$$- \beta_k\sum_{i=1}^{n_{nr}} E\big[z_i\big]E\Big[{t_{ik}^*}^{\alpha_k}\ln(t_{ik}^*) \mid t_{ik}^* < s_i\Big] - \beta_k\sum_{i=1}^{n_c} s_i^{\alpha_k}\ln(s_i) \tag{4.10}$$

$$\frac{\partial Q}{\partial \beta_k} = \frac{n_k}{\beta_k} - \sum_{i=1}^{n_k} t_i^{\alpha_k} + \frac{1}{\beta_k}\sum_{i=1}^{n_{nr}} E\big[z_i\big] - \sum_{i=1}^{n_{nr}} E\big[z_i\big]E\Big[{t_{ik}^*}^{\alpha_k} \mid t_{ik}^* < s_i\Big] - \sum_{i=1}^{n_c} s_i^{\alpha_k}\ln(s_i) \tag{4.11}$$

$$\frac{\partial Q}{\partial \lambda_k} = -\sum_{i=1}^{n_k}(s_i - t_i) - \sum_{i=1}^{n_{nr}} E\big[z_i\big]E\Big[w_{ik}^* \mid w_{ik}^* < (s_i - t_{ik}^*)\Big] + \frac{1}{\lambda_k}\sum_{i=1}^{n_{nr}} E\big[z_i\big] \tag{4.12}$$

To calculate the required conditional expectations at E-step of the E-M algorithm, we use the multinomial distribution and the conditional density defined in 4.5. Thus, the required conditional expectation terms are

$$\xi_1(s_i;\alpha_k,\beta_k,\lambda_k) = E\Big[z_i\Big] = p_{ik},$$

$$\xi_2(t_{ik}^*;\alpha_k,\beta_k) = E\Big[\ln(t_{ik}^*)\big|t_{ik}^* < s_i, z_i, \alpha_k, \beta_k\Big] = \frac{\int_0^{s_i}\ln(u)\alpha_k\beta_k u^{\alpha_k-1}\exp\{-\beta_k u^{\alpha_k}\}du}{1-\exp\{-\beta_k s_i^{\alpha_k}\}}$$
$$= \frac{K_1(s_i) - \ln(\beta_k)\Big(1-\exp\{-\beta_k s_i^{\alpha_k}\}\Big)}{\alpha_k\Big(1-\exp\{-\beta_k s_i^{\alpha_k}\}\Big)},$$

$$\xi_3(t_{ik}^*;\alpha_k,\beta_k) = E\Big[{t_{ik}^*}^{\alpha_k}\ln(t_{ik}^*)\big|t_{ik}^* < s_i, z_i, \alpha_k, \beta_k\Big]$$
$$= \frac{\int_0^{s_i} u^{\alpha_k}\ln(u)\alpha_k\beta_k u^{\alpha_k-1}\exp\{-\beta_k u^{\alpha_k}\}du}{1-\exp\{-\beta_k s_i^{\alpha_k}\}}$$
$$= \frac{K_2(s_i) - \ln(\beta_k)\Big(1-(1+\beta_k s_i^{\alpha_k})\exp\{-\beta_k s_i^{\alpha_k}\}\Big)}{\alpha_k\beta_k\Big(1-\exp\{-\beta_k s_i^{\alpha_k}\}\Big)},$$

$$\xi_4(t_{ik}^*;\alpha_k,\beta_k) = E\Big[{t_{ik}^*}^{\alpha_k}\big|t_{ik}^* < s_i, z_i, \alpha_k, \beta_k\Big] = \frac{\int_0^{s_i} u^{\alpha_k}\alpha_k\beta_k u^{\alpha_k-1}\exp\{-\beta_k u^{\alpha_k}\}du}{1-\exp\{-\beta_k s_i^{\alpha_k}\}}$$
$$= \frac{1-(1+\beta_k s_i^{\alpha_k})\exp\{-\beta_k s_i^{\alpha_k}\}}{\beta_k\Big(1-\exp\{-\beta_k s_i^{\alpha_k}\}\Big)},$$

$$\xi_5(w_{ik}^*;\lambda_k) = E\Big[w_{ik}^*\big|w_{ik}^* < (s_i - t_{ik}^*), z_i, \lambda_k\Big] = \frac{\int_0^{s_i - t_{ik}^*} u\lambda_k\exp\{-\lambda_k u\}}{1-\exp\{-\lambda_k(s_i - t_{ik}^*)\}}$$

$$= \frac{1}{\lambda_k}\left[\frac{1-\{1+\lambda_k(s_i-t^*_{ik})\}\exp\{-\lambda_k(s_i-t^*_{ik})\}}{1-\exp\{-\lambda_k(s_i-t^*_{ik})\}}\right],$$

$$\xi_6(t^*_{ik};\alpha_k,\beta_k) = E\Big[{t^*_{ik}}^{\alpha_k}\Big(\ln(t^*_{ik})\Big)^2 \Big| t^*_{ik} < s_i, z_i, \alpha_k, \beta_k\Big]$$

$$= \frac{\int_0^{s_i} u^{\alpha_k}\Big(\ln(u)\Big)^2 \alpha_k\beta_k u^{\alpha_k-1}\exp\{-\beta_k u^{\alpha_k}\}du}{1-\exp\{-\beta_k s_i^{\alpha_k}\}}$$

$$= \frac{K_3(s_i) + \Big(\ln(\beta_k)\Big)^2\Big(1-(1+\beta_k s_i^{\alpha_k})\exp\{-\beta_k s_i^{\alpha_k}\}\Big) - 2\ln(\beta_k)K_2(s_i)}{\alpha_k^2\beta_k\Big(1-\exp\{-\beta_k s_i^{\alpha_k}\}\Big)},$$

where, $K_1(y) = \int_0^{\beta_k y^{\alpha_k}} \ln(y)\exp\{-y\}dy$, $\quad K_2(y) = \int_0^{\beta_k y^{\alpha_k}} y\ln(y)\exp\{-y\}dy$, and $K_3(y) = \int_0^{\beta_k y^{\alpha_k}} y\Big(\ln(y)\Big)^2\exp\{-y\}dy$.

Thus, in the M-step of the $(m+1)^{th}$ iteration of the E-M algorithm, the values of $\hat{\alpha}_k^{(m+1)}$, $\hat{\beta}_k^{(m+1)}$ and $\hat{\lambda}_k^{(m+1)}$ can be obtained by solving following expressions

$$\hat{\alpha}_k^{(m+1)} = \frac{n_k + \sum_{i=1}^{n_{nr}}\xi_1\Big(s_i;\alpha_k^{(m)},\beta_k^{(m)},\lambda_k^{(m)}\Big)}{\left[\begin{array}{l} -\sum_{i=1}^{n_k}\ln(t_i) - \sum_{i=1}^{n_{nr}}\xi_1\Big(s_i;\alpha_k^{(m)},\beta_k^{(m)},\lambda_k^{(m)}\Big)\xi_2\Big(t^*_{ik};\alpha_k^{(m)},\beta_k^{(m)}\Big) \\ \quad + \hat{\beta}_k^{(m)}\Big\{\sum_{i=1}^{n_k} t_i^{\hat{\alpha}_k^{(m)}}\ln(t_i) + \sum_{i=1}^{n_{nr}}\xi_1\Big(s_i;\alpha_k^{(m)},\beta_k^{(m)},\lambda_k^{(m)}\Big)\xi_3\Big(t^*_{ik};\alpha_k^{(m)},\beta_k^{(m)}\Big) + \sum_{i=1}^{n_c} s_i^{\hat{\alpha}_k^{(m)}}\ln(s_i)\Big\} \end{array}\right]}, \tag{4.13}$$

$$\hat{\beta}_k^{(m+1)} = \frac{n_k + \sum_{i=1}^{n_{nr}}\xi_1\Big(s_i;\alpha_k^{(m)},\beta_k^{(m)},\lambda_k^{(m)}\Big)}{\sum_{1=1}^{n_k} t_i^{\hat{\alpha}_k^{(m+1)}} + \sum_{i=1}^{n_{nr}}\xi_1\Big(s_i;\alpha_k^{(m)},\beta_k^{(m)},\lambda_k^{(m)}\Big)\xi_4\Big(t^*_{ik};\alpha_k^{(m)},\beta_k^{(m)}\Big) + \sum_{i=1}^{n_c} s_i^{\hat{\alpha}_k^{(m+1)}}\ln(s_i)}, \tag{4.14}$$

$$\hat{\lambda}_k^{(m+1)} = \frac{\sum_{i=1}^{n_{nr}}\xi_1\Big(s_i;\alpha_k^{(m)},\beta_k^{(m)},\lambda_k^{(m)}\Big)}{\sum_{i=1}^{n_k}(s_i-t_i) + \sum_{i=1}^{n_{nr}}\xi_1\Big(s_i;\alpha_k^{(m)},\beta_k^{(m)},\lambda_k^{(m)}\Big)\xi_5\Big(w^*_{ik};\lambda_k^{(m)}\Big)} \tag{4.15}$$

4.3.2 Information Matrix

To obtain the observed Fisher information matrix, we use the method proposed by Louis (1982). The method is based on the missing information principle. For set of parameter vector Θ, the observed Fisher information matrix can be calculated using the identity $I(\hat{\Theta}) = I_1 - I_2{I_2}^T$, where matrix I_1 can be obtained by taking negative of second derivatives 4.9 while matrix I_2 is obtained by taking gradients (4.10) to (4.11). The structure of the observed

Fisher information matrix is given by

$$I(\hat{\Theta}) = \left[\begin{array}{ccc|ccc|ccc|ccc} -\frac{\partial^2 Q}{\partial \alpha_1^2} - \left(\frac{\partial Q}{\partial \alpha_1}\right)^2 & \cdots & -\frac{\partial^2 Q}{\partial \alpha_1 \partial \alpha_K} - \frac{\partial Q}{\partial \alpha_1}\frac{\partial Q}{\partial \alpha_K} & -\frac{\partial^2 Q}{\partial \alpha_1 \partial \beta_1} - \frac{\partial Q}{\partial \alpha_1}\frac{\partial Q}{\partial \beta_1} & \cdots & -\frac{\partial^2 Q}{\partial \alpha_1 \partial \beta_K} - \frac{\partial Q}{\partial \alpha_1}\frac{\partial Q}{\partial \beta_K} & -\frac{\partial^2 Q}{\partial \alpha_1 \partial \lambda_1} - \frac{\partial Q}{\partial \alpha_1}\frac{\partial Q}{\partial \lambda_1} & \cdots & -\frac{\partial^2 Q}{\partial \alpha_1 \partial \lambda_K} - \frac{\partial Q}{\partial \alpha_1}\frac{\partial Q}{\partial \lambda_K} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -\frac{\partial^2 Q}{\partial \alpha_K \partial \alpha_1} - \frac{\partial Q}{\partial \alpha_K}\frac{\partial Q}{\partial \alpha_1} & \cdots & -\frac{\partial^2 Q}{\partial \alpha_K^2} - \left(\frac{\partial Q}{\partial \alpha_K}\right)^2 & -\frac{\partial^2 Q}{\partial \alpha_K \partial \beta_1} - \frac{\partial Q}{\partial \alpha_K}\frac{\partial Q}{\partial \beta_1} & \cdots & -\frac{\partial^2 Q}{\partial \alpha_K \partial \beta_K} - \frac{\partial Q}{\partial \alpha_K}\frac{\partial Q}{\partial \beta_K} & -\frac{\partial^2 Q}{\partial \alpha_K \partial \lambda_1} - \frac{\partial Q}{\partial \alpha_K}\frac{\partial Q}{\partial \lambda_1} & \cdots & -\frac{\partial^2 Q}{\partial \alpha_K \partial \lambda_K} - \frac{\partial Q}{\partial \alpha_K}\frac{\partial Q}{\partial \lambda_K} \\ \hline -\frac{\partial^2 Q}{\partial \beta_1 \partial \alpha_1} - \frac{\partial Q}{\partial \beta_1}\frac{\partial Q}{\partial \alpha_1} & \cdots & -\frac{\partial^2 Q}{\partial \beta_1 \partial \alpha_K} - \frac{\partial Q}{\partial \beta_1}\frac{\partial Q}{\partial \alpha_K} & -\frac{\partial^2 Q}{\partial \beta_1^2} - \left(\frac{\partial Q}{\partial \beta_1}\right)^2 & \cdots & -\frac{\partial^2 Q}{\partial \beta_1 \partial \beta_K} - \frac{\partial Q}{\partial \beta_1}\frac{\partial Q}{\partial \beta_K} & -\frac{\partial^2 Q}{\partial \beta_1 \partial \lambda_1} - \frac{\partial Q}{\partial \beta_1}\frac{\partial Q}{\partial \lambda_1} & \cdots & -\frac{\partial^2 Q}{\partial \beta_1 \partial \lambda_K} - \frac{\partial Q}{\partial \beta_1}\frac{\partial Q}{\partial \lambda_K} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -\frac{\partial^2 Q}{\partial \beta_K \partial \alpha_1} - \frac{\partial Q}{\partial \beta_K}\frac{\partial Q}{\partial \alpha_1} & \cdots & -\frac{\partial^2 Q}{\partial \beta_K \partial \alpha_K} - \frac{\partial Q}{\partial \beta_K}\frac{\partial Q}{\partial \alpha_K} & -\frac{\partial^2 Q}{\partial \beta_K \partial \beta_1} - \frac{\partial Q}{\partial \beta_K}\frac{\partial Q}{\partial \beta_1} & \cdots & -\frac{\partial^2 Q}{\partial \beta_K^2} - \left(\frac{\partial Q}{\partial \beta_K}\right)^2 & -\frac{\partial^2 Q}{\partial \beta_K \partial \lambda_1} - \frac{\partial Q}{\partial \beta_K}\frac{\partial Q}{\partial \lambda_1} & \cdots & -\frac{\partial^2 Q}{\partial \beta_K \partial \lambda_K} - \frac{\partial Q}{\partial \beta_K}\frac{\partial Q}{\partial \lambda_K} \\ \hline -\frac{\partial^2 Q}{\partial \lambda_1 \partial \alpha_1} - \frac{\partial Q}{\partial \lambda_1}\frac{\partial Q}{\partial \alpha_1} & \cdots & -\frac{\partial^2 Q}{\partial \lambda_1 \partial \alpha_K} - \frac{\partial Q}{\partial \lambda_1}\frac{\partial Q}{\partial \alpha_K} & -\frac{\partial^2 Q}{\partial \lambda_1 \partial \beta_1} - \frac{\partial Q}{\partial \lambda_1}\frac{\partial Q}{\partial \beta_1} & \cdots & -\frac{\partial^2 Q}{\partial \lambda_1 \partial \beta_K} - \frac{\partial Q}{\partial \lambda_1}\frac{\partial Q}{\partial \beta_K} & -\frac{\partial^2 Q}{\partial \lambda_1^2} - \left(\frac{\partial Q}{\partial \lambda_1}\right)^2 & \cdots & -\frac{\partial^2 Q}{\partial \lambda_1 \partial \lambda_K} - \frac{\partial Q}{\partial \lambda_1}\frac{\partial Q}{\partial \lambda_K} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -\frac{\partial^2 Q}{\partial \lambda_K \partial \alpha_1} - \frac{\partial Q}{\partial \lambda_K}\frac{\partial Q}{\partial \alpha_1} & \cdots & -\frac{\partial^2 Q}{\partial \lambda_K \partial \alpha_K} - \frac{\partial Q}{\partial \lambda_K}\frac{\partial Q}{\partial \beta_1} & -\frac{\partial^2 Q}{\partial \lambda_K \partial \beta_1} - \frac{\partial Q}{\partial \lambda_K}\frac{\partial Q}{\partial \beta_1} & \cdots & -\frac{\partial^2 Q}{\partial \lambda_K \partial \beta_K} - \frac{\partial Q}{\partial \lambda_K}\frac{\partial Q}{\partial \beta_K} & -\frac{\partial^2 Q}{\partial \lambda_K \partial \lambda_1} - \frac{\partial Q}{\partial \lambda_K}\frac{\partial Q}{\partial \lambda_1} & \cdots & -\frac{\partial^2 Q}{\partial \lambda_K^2} - \left(\frac{\partial Q}{\partial \lambda_K}\right)^2 \end{array}\right]\Bigg|_{\Theta=\hat{\Theta}}$$

The terms used for the construction of observed Fisher information matrix are given below

$$\begin{aligned}
\frac{\partial^2 Q}{\partial \alpha_k^2} &= -\frac{n_k + \sum_{i=1}^{n_{nr}} \xi_1(s_i;\alpha_k,\beta_k,\lambda_k)}{\alpha_k^2} - \beta_k \sum_{i=1}^{n_{nr}} \xi_1(s_i;\alpha_k,\beta_k,\lambda_k)\xi_6(t_{ik}^*;\alpha_k,\beta_k) \\
&\quad - \beta_k \sum_{i=1}^{n_c} s_i^{\alpha_k}\Big(\ln(s_i)\Big)^2, \\
\frac{\partial^2 Q}{\partial \beta_k^2} &= -\frac{n_k + \sum_{i=1}^{n_{nr}} \xi_1(s_i;\alpha_k,\beta_k,\lambda_k)}{\beta_k^2}, \\
\frac{\partial^2 Q}{\partial \lambda_k^2} &= -\frac{\sum_{i=1}^{n_{nr}} \xi_1(s_i;\alpha_k,\beta_k,\lambda_k)}{\lambda_k^2}, \\
\frac{\partial^2 Q}{\partial \beta_k \partial \alpha_k} &= -\sum_{i=1}^{n_k} t_i^{\alpha_k}\ln(t_i) - \sum_{i=1}^{n_{nr}} \xi_1(s_i;\alpha_k,\beta_k,\lambda_k)\xi_3(t_{ik}^*;\alpha_k,\beta_k) - \sum_{i=1}^{n_c} s_i^{\alpha_k}\ln(s_i).
\end{aligned}$$

The variance-covariance matrix can be found by inverting the matrix $I(\hat{\Theta})$. The square root of diagonal elements of the variance-covariance matrix gives the standard error (SE) of $\hat{\Theta}$ which can be used in the construction of confidence intervals for unknown parameters. Thus, a $100(1-\gamma)\%$ confidence interval for $\hat{\Theta}$ is given by $\hat{\Theta} \pm z_{\gamma/2}\sqrt{I^{-1}(\hat{\Theta})}$.

4.4 Bayesian Inference

In the classical section, the parameters are considered to be a fixed but unknown quantity. However, in the Bayesian approach, parameters are considered as random variables with some prior density. This prior information available on parameters is utilized along with the

likelihood to make the posterior distribution. All the inferences in the Bayesian paradigm are drawn from the posterior distribution. Here, it can be observed that the posterior analysis with likelihood function (4.4) is not easy. So, for posterior analysis, the complete likelihood function (4.7) which is constructed for the E-M implementation is used. The data augmentation approach based on Gibbs sampling is utilized for sample generation from the conditional posteriors.

4.4.1 Prior Distribution

In literature, different priors are discussed for the shape and scale for Weibull distribution under Bayesian analysis. We have given a discussion on the choice of prior for the shape and scale in 2.3.1. In this study, we consider α_k, β_k, λ_k; $k = 1, 2, \cdots, K$ to be Gamma distributed, i.e. $\alpha_k \sim \mathcal{G}(a_k, b_k)$, $\beta_k \sim \mathcal{G}(c_k, d_k)$ and $\lambda_k \sim \mathcal{G}(e_k, f_k)$. Assuming the independence of priors the joint prior density is written up to proportionality constants as below

$$\pi(\Theta) \propto \prod_{k=1}^{K} {\alpha_k}^{a_k-1} {\beta_k}^{c_k-1} {\lambda_k}^{e_k-1} \exp\{-(\alpha_k b_k + \beta_k d_k + \lambda_k f_k)\}\,;\; a_k, b_k, c_k, d_k, e_k, f_k > 0, k = 1, 2, \cdots, K. \tag{4.16}$$

The values of hyperparameters can be obtained by using the moment matching approach. For this, we equate the mean and variance of MLEs based on samples (say, M) with the mean and variance of considered gamma priors.

4.4.2 Gibbs Sampling

For applying the Gibbs Sampling the posterior distribution is required which can be obtained by merging complete likelihood (4.7) and the joint prior (4.16). The posterior distribution up-to proportionality constant is written below

$$\Pi(\Theta|d,d^*) \propto L_c(\Theta|d,d^*)\pi(\Theta)$$

$$\propto \prod_{k=1}^{K}\prod_{i=1}^{n_k}\left[\alpha_k\beta_k t_i^{\alpha_k-1}\exp\{-\beta_k t_i^{\alpha_k}\}\prod_{\substack{l=1\\l\neq k}}^{K}\exp\{-\beta_l t_i^{\alpha_l}\}\exp\{-\lambda_k(s_i-t_i)\}\right]$$

$$\prod_{i=1}^{n_{nr}}\prod_{k=1}^{K}\left[\alpha_k\beta_k\lambda_k {t_{ik}^*}^{\alpha_k-1}\exp\{-\beta_k {t_{ik}^*}^{\alpha_k}\}\prod_{\substack{l=1\\l\neq k}}^{K}\exp\{-\beta_l {t_{ik}^*}^{\alpha_l}\}\exp\{-\lambda_k w_{ik}^*\}\right]^{z_i}$$

$$\prod_{i=1}^{n_c}\left[\exp\left\{-\sum_{k=1}^{K}\beta_k s_i^{\alpha_k}\right\}\right]\prod_{k=1}^{K}{\alpha_k}^{a_k-1}{\beta_k}^{c_k-1}{\lambda_k}^{e_k-1}\exp\{-(\alpha_k b_k+\beta_k d_k+\lambda_k f_k)\} \tag{4.17}$$

Proceeding further in applying Gibbs Sampling, we need to write full conditionals for parameters based on the posterior distribution (4.17). The full conditionals for α_k's, β_k's and λ_k's are given below

$$\alpha_k|\Theta_{(-\alpha_k)} \propto \alpha_k^{(n_k+Z^*+a_k-1)}\left(\prod_{i=1}^{n_k}t_i^{\alpha_k-1}\right)\left(\prod_{i=1}^{n_{nr}}{t_{ik}^*}^{z_i(\alpha_k-1)}\right)\exp\left\{-\beta_k\left(\sum_{i=1}^{n_k}t_i^{\alpha_k}+\sum_{i=1}^{n_{nr}}z_i{t_{ik}^*}^{\alpha_k}+\sum_{i=1}^{n_c}s_i^{\alpha_k}+b_k\right)\right\}, \tag{4.18}$$

$$\beta_k|\Theta_{(-\beta_k)} \sim \mathcal{G}\left(n_k+Z^*+c_k,\ \sum_{i=1}^{n_k}t_i^{\alpha_k}+\sum_{i=1}^{n_{nr}}z_i{t_{ik}^*}^{\alpha_k}+\sum_{i=1}^{n_c}s_i^{\alpha_k}+d_k\right), \tag{4.19}$$

$$\lambda_k|\Theta_{(-\lambda_k)} \sim \mathcal{G}\left(Z^*+e_k,\ \sum_{i=1}^{n_k}(s_i-t_i)+\sum_{i=1}^{n_{nr}}z_i w_{ik}^*+f_k\right), \tag{4.20}$$

From (4.18) to (4.20), we can see that the full conditionals of α_k are not in any standard distribution form. So, samples from these full conditionals are generated by using Metropolis-Hasting (M-H) (Metropolis and Ulam (1949) and Hastings (1970)) algorithm taking normal distribution as proposal density. Further, β_k and λ_k are following gamma distribution with varying shapes and scales and observations from these can be drawn directly for given values of latent variables Z, T^*, and W^*.

4.4.3 Data Augmentation Algorithm

The algorithm used for generating samples from full conditionals using three stage Gibbs sampling is given below:

1. Step 1: Set initial values of parameters, say $\Theta^{(0)} = \left(\alpha_k^{(0)}, \beta_k^{(0)}, \lambda_k^{(0)}\right)$ and generate $z_i \sim \mathcal{MD}(1, p_{i1}, p_{i2}, \ldots, p_{iK});\ i = 1, 2, \cdots, n_{nr},\ k = 1, 2, \cdots, K.$

2. Step 2: For given values of z_i and Θ, generate observations on t_{ik}^* using expression

$$t_{ik}^* = \left[-\frac{1}{\beta_k} \ln\left\{1 - u_i\left(1 - \exp\{-\beta_k s_i^{\alpha_k}\}\right)\right\}\right]^{1/\alpha_k};\ i = 1, 2, \ldots, n_{nr},\ k = 1, 2, \cdots, K \tag{4.21}$$

   where $u \sim \mathcal{U}(0, 1)$.

3. Step 3: Based on generated values of z_i, t_{ik}^* and given Θ from previous steps, observations on w_{ik}^* can be generated by using expression

$$w_{ik}^* = -\frac{1}{\lambda_k} \ln\left[1 - u_i\Big(1 - \exp\{-\lambda_k(s_i - t_{ik}^*)\}\Big)\right];\ i = 1, 2, \ldots, n_{nr},\ k = 1, 2, \cdots, K. \tag{4.22}$$

4. Step 4: For given values of z_i, t_{ik}^* and w_{ik}^* generate observations on $\alpha_k^{(1)}$ using M-H algorithm taking normal as proposal density.

5. Step 5: In the next step observations on $\beta_k^{(1)}$ and $\lambda_k^{(1)}$ can be generated for given $\alpha_k^{(1)}$ using (4.19) to (4.20) respectively.

Now, the current state is $\Theta^{(1)} = \left(\alpha_k^{(1)}, \beta_k^{(1)}, \lambda_k^{(1)}\right)$. Step 1 to step 5 are replicated M times to obtain a sequence of random variables $\left(\Theta^{(1)}, \Theta^{(2)}, \cdots, \Theta^{(M)}\right)$. After discarding the burn-in from generated chains, the stationarity of the chains is checked by trace plot and Gelman and Rubin's test statistics (see Gelman and Rubin (1992)), we left out with a reduced chain of length M'. All the Bayesian inferences are drawn based on these samples. The HPD

intervals based on these samples are obtained using a method proposed by Chen and Shao (1999).

4.5 Simulation Study

In this section, a simulation study is carried out to access the performance of the proposed methodology. The computation is carried out using the **R** software. The proposed methodology is developed for the K causes setup, but in the simulation study, we consider only two causes, which are responsible for the occurrence of the event. First, we generated two samples of size n such as $T_1 \sim \mathcal{W}(\alpha_1, \beta_1)$ and $T_2 \sim \mathcal{W}(\alpha_2, \beta_2)$ for the prefixed value of parameters $(\alpha_1, \beta_1) = (0.95, 1.10)$ and $(\alpha_1, \beta_1) = (0.97, 1.12)$. These T_1 and T_2 are considered to be the latent event times for individuals due to causes 1 and 2, respectively.Then the event time becomes $T = \min(T_1, T_2)$. Also for an individual if $T = T_1$ then it is considered that the event has occurred due to cause 1 or if $T = T_2$ then it is due to 2. To explore the impact of monitoring time on the inference of various estimators we consider uniform as well as exponential monitoring patterns. For the uniform monitoring pattern, S is generated in the range of T while in the case of exponential monitoring patterns, the observations are generated by taking the rate parameter 0.4. As in this setup, we have three types of individuals in the study: those who have experienced the event and are able to recall the exact event time and cause as well; those who have experienced the event but are not able to recall the exact event time and cause and finally some are in right-censored category for which event of interest has not been experienced at the monitoring time point. So to introduce these three categories artificially in the generated sample under both monitoring patterns, we first fix a set of parameter $(\lambda_1, \lambda_2) = \{(0.10, 0.12), (0.17, 0.18)\}$ for non-recalled category. This set of parameters produces almost 30% and 40% proportion of non-recall observations from non-censored observations. An observation is randomly assigned as non-recall or censored in the sample.

By using the E-M algorithm discussed in section 4.3, for point estimates, mean square error and absolute bias are reported in Table 4.1 and Table 4.2. The interval estimates of

parameters are obtained using the asymptotic properties of MLE and corresponding average lengths, shape, and coverage probability are reported in Table 4.3 and Table 4.4. We denote the asymptotic confidence intervals by symbol ACI. Under Bayesian analysis, taking the gamma as priors for both scale and shape parameters, samples are generated using the Gibbs sampling algorithm. The hyper-parameters of prior distributions are chosen by using the moment matching criteria approach. The whole simulation process is replicated 500 times to average out the results.

Table 4.1: Mean square error and absolute bias for point estimates under uniform monitoring points for varying sample sizes.

			ML			*Bayes*		
(λ_1,λ_2)		$n \rightarrow$	80	150	250	80	150	250
(0.10, 0.12)	$\hat{\alpha}_1$	*MSE*	0.0177	0.0121	0.0073	0.0160	0.0111	0.0052
		AB	0.1048	0.0881	0.0690	0.0916	0.0707	0.0503
	$\hat{\beta}_1$	*MSE*	0.0717	0.0441	0.0390	0.0414	0.0229	0.0214
		AB	0.2258	0.1651	0.1616	0.1745	0.1250	0.1238
	$\hat{\lambda}_1$	*MSE*	0.0204	0.0117	0.0084	0.0191	0.0101	0.0075
		AB	0.1128	0.0928	0.0760	0.0955	0.0761	0.0558
	$\hat{\alpha}_2$	*MSE*	0.0174	0.0083	0.0073	0.0109	0.0061	0.0054
		AB	0.1008	0.0710	0.0648	0.0917	0.0607	0.0404
	$\hat{\beta}_2$	*MSE*	0.0909	0.0488	0.0465	0.0756	0.0261	0.0242
		AB	0.2513	0.1830	0.1789	0.1893	0.1324	0.1305
	$\hat{\lambda}_2$	*MSE*	0.0235	0.0112	0.0070	0.0204	0.0101	0.0050
		AB	0.1289	0.0913	0.0716	0.0915	0.0882	0.0573
(0.17, 0.18)	$\hat{\alpha}_1$	*MSE*	0.0204	0.0126	0.0076	0.0187	0.0121	0.0053
		AB	0.1145	0.0889	0.0696	0.0919	0.0809	0.0604
	$\hat{\beta}_1$	*MSE*	0.0770	0.0502	0.0451	0.0535	0.0250	0.0242
		AB	0.2319	0.1856	0.1740	0.1953	0.1372	0.1303
	$\hat{\lambda}_1$	*MSE*	0.0322	0.0179	0.0123	0.0202	0.0125	0.0102
		AB	0.1510	0.1095	0.0918	0.1090	0.0870	0.0672
	$\hat{\alpha}_2$	*MSE*	0.0205	0.0103	0.0086	0.0190	0.0101	0.0068
		AB	0.1058	0.0824	0.0654	0.0921	0.0608	0.0505
	$\hat{\beta}_2$	*MSE*	0.0965	0.0590	0.0484	0.0574	0.0340	0.0260
		AB	0.2610	0.1962	0.1869	0.1967	0.1479	0.1352
	$\hat{\lambda}_2$	*MSE*	0.0289	0.0183	0.0111	0.0153	0.0120	0.0101
		AB	0.1427	0.1186	0.0874	0.0939	0.1034	0.0684

Table 4.2: Mean square error and absolute bias for point estimates under exponential monitoring points for varying sample sizes.

			ML			*Bayes*		
(λ_1,λ_2)		$n \rightarrow$	80	150	250	80	150	250
(0.10, 0.12)	$\hat{\alpha}_1$	*MSE*	0.0216	0.0118	0.0094	0.0208	0.0102	0.0054
		AB	0.1191	0.0923	0.0816	0.1011	0.0805	0.0702
	$\hat{\beta}_1$	*MSE*	0.1292	0.0800	0.0710	0.0743	0.0570	0.0573
		AB	0.3249	0.2540	0.2422	0.2341	0.1857	0.1685
	$\hat{\lambda}_1$	*MSE*	0.0047	0.0032	0.0017	0.0041	0.0030	0.0010
		AB	0.0577	0.0454	0.0322	0.0466	0.0264	0.0222
	$\hat{\alpha}_2$	*MSE*	0.0212	0.0106	0.0096	0.0182	0.0087	0.0078
		AB	0.1058	0.0845	0.0839	0.0911	0.0705	0.0702
	$\hat{\beta}_2$	*MSE*	0.1001	0.0673	0.0671	0.0793	0.0369	0.0346
		AB	0.2797	0.2288	0.2039	0.2136	0.1668	0.1590
	$\hat{\lambda}_2$	*MSE*	0.0063	0.0039	0.0027	0.0030	0.0028	0.0019
		AB	0.0643	0.0472	0.0395	0.0348	0.0313	0.0300
(0.17, 0.18)	$\hat{\alpha}_1$	*MSE*	0.0218	0.0131	0.0099	0.0215	0.0120	0.0073
		AB	0.1207	0.0944	0.0827	0.0102	0.0813	0.0704
	$\hat{\beta}_1$	*MSE*	0.1298	0.1288	0.0963	0.0751	0.0678	0.0591
		AB	0.3320	0.3033	0.2879	0.2351	0.2311	0.1932
	$\hat{\lambda}_1$	*MSE*	0.0094	0.0052	0.0029	0.0059	0.0037	0.0028
		AB	0.0725	0.0567	0.0428	0.0471	0.0460	0.0409
	$\hat{\alpha}_2$	*MSE*	0.0216	0.0126	0.0107	0.0210	0.0091	0.0086
		AB	0.1066	0.0954	0.0891	0.0914	0.0906	0.0803
	$\hat{\beta}_2$	*MSE*	0.1385	0.1251	0.1237	0.0905	0.0636	0.0525
		AB	0.3352	0.3287	0.3140	0.2300	0.2197	0.2034
	$\hat{\lambda}_2$	*MSE*	0.0111	0.0066	0.0030	0.0072	0.0048	0.0026
		AB	0.0815	0.0635	0.0451	0.0542	0.0526	0.0428

Table 4.3: Average length, shape and coverage probability for interval estimates under uniform monitoring points for varying sample sizes.

			ACI			*HPD*		
$(\hat{\lambda}_1,\hat{\lambda}_2)$		$n \rightarrow$	80	150	250	80	150	250
(0.10, 0.12)		*AL*	0.5069	0.3749	0.2818	0.4863	0.3350	0.2603
	$\hat{\alpha}_1$	*Shape*	1.0000	1.0000	1.0000	1.0092	1.0176	1.0068
		CP	0.9533	0.9067	0.9200	1.0000	1.0000	1.0000
		AL	0.6731	0.5357	0.3920	0.6589	0.5204	0.3871
	$\hat{\beta}_1$	*Shape*	1.0000	1.0000	1.0000	1.0741	1.0531	1.0242
		CP	0.8733	0.8780	0.8640	0.9600	0.9733	0.9500
		AL	0.2066	0.1424	0.1106	0.2021	0.1298	0.1051
	$\hat{\lambda}_1$	*Shape*	1.0000	1.0000	1.0000	1.1080	1.0692	1.0623
		CP	0.8460	0.8360	0.8390	1.0000	1.0000	1.0000
		AL	0.5029	0.3805	0.2881	0.4850	0.3769	0.2612
	$\hat{\alpha}_2$	*Shape*	1.0000	1.0000	1.0000	1.0899	1.0549	1.0132
		CP	0.9333	0.9467	0.9400	1.0000	1.0000	1.0000
		AL	0.6902	0.5318	0.3931	0.6732	0.5150	0.3852
	$\hat{\beta}_2$	*Shape*	1.0000	1.0000	1.0000	1.0926	1.0479	1.0433
		CP	0.8707	0.8747	0.8560	0.9000	0.9533	0.9900
		AL	0.2507	0.1434	0.1083	0.2448	0.1225	0.1055
	$\hat{\lambda}_2$	*Shape*	1.0000	1.0000	1.0000	1.0462	1.0401	1.0168
		CP	0.8500	0.8467	0.8420	0.9933	1.0000	0.9900
(0.17, 0.18)		*AL*	0.5328	0.3771	0.2944	0.5029	0.3564	0.2720
	$\hat{\alpha}_1$	*Shape*	1.0000	1.0000	1.0000	1.0996	1.0699	1.0363
		CP	0.9500	0.9000	0.9300	1.0000	1.0000	1.0000
		AL	0.7362	0.5427	0.3959	0.7173	0.5361	0.3915
	$\hat{\beta}_1$	*Shape*	1.0000	1.0000	1.0000	1.0908	1.0804	1.0574
		CP	0.8000	0.8710	0.8610	0.9600	0.9500	0.9500
		AL	0.2806	0.1888	0.1483	0.2749	0.1793	0.1280
	$\hat{\lambda}_1$	*Shape*	1.0000	1.0000	1.0000	1.0902	1.0581	1.0376
		CP	0.8505	0.8520	0.8440	1.0000	1.0000	1.0000
		AL	0.5395	0.3849	0.2947	0.5147	0.3832	0.2822
	$\hat{\alpha}_2$	*Shape*	1.0000	1.0000	1.0000	1.0153	1.0009	1.0001
		CP	0.9750	0.9650	0.9400	1.0000	1.0000	1.0000
		AL	0.7508	0.5320	0.4007	0.7499	0.5189	0.3934
	$\hat{\beta}_2$	*Shape*	1.0000	1.0000	1.0000	1.0876	1.0771	1.0574
		CP	0.8150	0.8690	0.8630	0.9500	0.9350	1.0000
		AL	0.2904	0.1983	0.1479	0.2719	0.1865	0.1289
	$\hat{\lambda}_2$	*Shape*	1.0000	1.0000	1.0000	1.0849	1.0497	1.0301
		CP	0.8560	0.8485	0.8440	1.0000	0.9950	1.0000

Table 4.4: Average length, shape and coverage probability for interval estimates under exponential monitoring points for varying sample sizes.

			ACI			*HPD*		
$(\hat{\lambda}_1,\hat{\lambda}_2)$		$n \rightarrow$	80	150	250	80	150	250
(0.10, 0.12)		*AL*	0.4719	0.3428	0.2550	0.4515	0.3301	0.2466
	$\hat{\alpha}_1$	*Shape*	1.0000	1.0000	1.0000	1.0100	1.0073	1.0014
		CP	0.8600	0.8900	0.8760	1.0000	1.0000	1.0000
		AL	0.5699	0.4394	0.3411	0.5441	0.4250	0.3319
	$\hat{\beta}_1$	*Shape*	1.0000	1.0000	1.0000	1.0941	1.0370	1.0235
		CP	0.8900	0.8810	0.8320	0.9200	0.9300	0.9800
		AL	0.1558	0.1026	0.0842	0.1321	0.1012	0.0798
	$\hat{\lambda}_1$	*Shape*	1.0000	1.0000	1.0000	1.1336	1.0895	1.0869
		CP	0.9100	0.8769	0.8680	1.0000	0.9700	0.9800
		AL	0.4666	0.3400	0.2590	0.4563	0.3296	0.2517
	$\hat{\alpha}_2$	*Shape*	1.0000	1.0000	1.0000	1.0198	1.0025	1.0023
		CP	0.9000	0.8900	0.8200	1.0000	1.0000	1.0000
		AL	0.6050	0.4514	0.3422	0.5862	0.4328	0.3350
	$\hat{\beta}_2$	*Shape*	1.0000	1.0000	1.0000	1.0770	1.0426	1.0090
		CP	0.8600	0.8947	0.8526	0.9800	0.9800	1.0000
		AL	0.1873	0.1250	0.0925	0.1748	0.1096	0.0887
	$\hat{\lambda}_2$	*Shape*	1.0000	1.0000	1.0000	1.1124	1.0923	1.0645
		CP	0.8200	0.8700	0.8700	1.0000	0.9800	0.9400
(0.17, 0.18)		*AL*	0.4861	0.3592	0.2649	0.4784	0.3412	0.2579
	$\hat{\alpha}_1$	*Shape*	1.0000	1.0000	1.0000	1.0036	1.0024	1.0010
		CP	0.8800	0.8500	0.8760	1.0000	1.0000	1.0000
		AL	0.5847	0.4939	0.3608	0.5780	0.4799	0.3544
	$\hat{\beta}_1$	*Shape*	1.0000	1.0000	1.0000	1.1434	1.1118	1.0916
		CP	0.8373	0.8200	0.8780	0.9800	0.9500	0.9800
		AL	0.2219	0.1638	0.1238	0.2089	0.1438	0.1187
	$\hat{\lambda}_1$	*Shape*	1.0000	1.0000	1.0000	1.1455	1.1322	1.1019
		CP	0.8067	0.8760	0.9150	0.9867	0.9900	1.0000
		AL	0.4761	0.3728	0.2599	0.4575	0.3532	0.2689
	$\hat{\alpha}_2$	*Shape*	1.0000	1.0000	1.0000	1.0498	1.0298	1.0019
		CP	0.9200	0.8000	0.8740	1.0000	1.0000	1.0000
		AL	0.6539	0.5139	0.3601	0.6492	0.5085	0.3499
	$\hat{\beta}_2$	*Shape*	1.0000	1.0000	1.0000	1.1408	1.1145	1.0940
		CP	0.8367	0.8205	0.9070	0.9667	0.9400	0.9500
		AL	0.2252	0.1674	0.1251	0.2117	0.1559	0.1137
	$\hat{\lambda}_2$	*Shape*	1.0000	1.0000	1.0000	1.1403	1.1273	1.1014
		CP	0.8740	0.8685	0.8720	0.9867	0.9900	0.9700

From the simulation study, we can see that under both monitoring patterns, the value of MSE and AB decreases as the sample size increases which validates the consistency of estimators. The average length for ACIs and HPD intervals decreases as we increase the sample size. Further, as the proportion of incomplete data increases, the average values of MSE, AB, and AL increases. The coverage probabilities in the case of ACIs and HPD intervals are at the nominal levels. The values of shape for HPD approach symmetry with the increase in sample size. Thus, we can conclude that all the simulated results under both monitoring patterns are found to be satisfactory.

4.6 Real Data Analysis

For illustration purposes, we consider the menarche data described in Chapter 3. In the real data, variables related to age at menarche are given. Here, the age at onset of menarche is taken as the event of interest. The girls aged between $7-21$ years are interviewed at any age and asked to recall the age at which they experienced the event. There may be two possible cases whether they have experienced the event or not and accordingly they are categorized into non-censored and censored categories. Further, under the non-censored category, the responses of girls are available at the day level, month level and year level. For calculation purposes, all the observations are converted in years. Here, for the analysis purposes, we utilize data available at day and month levels. Since age at menarche differs for girls as per their birth order. So, we consider the first and later born as competing causes and the age at menarche is estimated.

When we consider data recalled at day level, we treat month and year level data as non-recall and we pool them into non-recall category. In second case, when we consider the data recalled at month level, then we set day and month level information into recall and year level recall data is pooled into non-recall category. Under these two scenarios, we get different categories of the data at three levels as below

1. Level I: Considering recall at the day level, we get different categories of data as $n_{r_1} = 339$, $n_{r_2} = 129$, $n_{nr} = 952$ and $n_c = 775$.

2. Level II: Considering recall at month levels, we have different categories of data as $n_{r_1} = 536$, $n_{r_2} = 202$, $n_{nr} = 682$ and $n_c = 775$.

For checking the suitability of Weibull distribution to assumed menarche data, we use the simulation based modified chi-square method as discussed earlier in real data Section 3.7 in Chapter 3. We calculate the p values based on above testing procedure for available three level datasets. We get p values 0.50 and 0.72 for data available at day levels and month levels respectively. Thus, Weibull distribution is a suitable choice for the time to event distribution. For calculation purposes, we have scaled data by dividing observations from 15. The point and interval estimates of parameters based on considered data are calculated under frequentist and Bayesian approaches and reported in Table 4.5. While calculating the Bayesian estimates, we don't have prior knowledge of real data, so hyper-parameters are chosen such that the prior density becomes non-informative. We report the mean and

Table 4.5: The point and interval estimates for menarche data under ML and Bayes approaches.

		ML		*Bayes*	
Level		*Estimate*	*ACI*	*Estimate*	*HPD*
I	$\hat{\alpha}_1$	5.7779	(5.5444, 6.0114)	5.7770	(5.7513, 5.8067)
	$\hat{\beta}_1$	1.3119	(1.2334, 1.3903)	1.3638	(1.2732, 1.4063)
	$\hat{\lambda}_1$	3.9393	(3.6611, 4.2174)	4.7969	(4.6473, 5.1063)
	$\hat{\alpha}_2$	7.4277	(6.9101, 7.9453)	7.4314	(7.3018, 7.5744)
	$\hat{\beta}_2$	0.3791	(0.3369, 0.4213)	0.3739	(0.3566, 0.4317)
	$\hat{\lambda}_2$	3.6560	(3.1240, 4.1879)	3.9484	(3.2441, 4.1085)
II	$\hat{\alpha}_1$	6.5865	(6.3186, 6.8544)	6.5872	(6.5539, 6.6228)
	$\hat{\beta}_1$	1.5396	(1.4431, 1.6362)	1.5876	(1.3657, 1.6283)
	$\hat{\lambda}_1$	2.3849	(2.1810, 2.5889)	2.6664	(2.2891, 2.5478)
	$\hat{\alpha}_2$	7.9565	(7.4411, 8.4719)	7.9539	(7.8189, 8.0916)
	$\hat{\beta}_2$	0.5553	(0.4971, 0.6134)	0.5553	(0.4951, 0.6048)
	$\hat{\lambda}_2$	2.2516	(1.8989, 2.6043)	2.3089	(1.9098, 2.4052)

median ages of age at menarche for first and later born girls obtained at all three levels as

1. Level I: Under ML, the mean and median age at menarche for first born girls comes out to be 13.25 and 13.43 years while under Bayesian paradigm, the mean and median age at menarche for first born girls comes out to be 13.17 and 13.35 years respectively. For the later born girls, under ML, the mean and median age at menarche for first born girls comes out to be 16.04 and 16.27 years while under Bayesian approach, the

mean and median age at menarche for first born girls comes out to be 16.08 and 16.31 years respectively.

2. Level II: Under ML, the mean and median age at menarche for first born girls comes out to be 13.09 and 13.28 years while under Bayesian approach, the mean and median age at menarche for first born girls comes out to be 13.05 and 13.23 years respectively. For the later born girls, under ML, the mean and median age at menarche for first born girls comes out to be 15.21 and 15.42 years while under Bayesian approach, the mean and median age at menarche for first born girls comes out to be 15.21 and 15.43 years respectively.

4.7 Concluding Remarks

The present chapter focuses on developing the latent failure-based approach for recall-based data under competing risks scenario. To capture the flexibility of data , we model the time to event due to different causes as Weibull distributed. The functional form for non-recall probability is considered of exponential form. In the classical framework, an equivalent quantity-based approach is used and an efficient algorithm based on the E-M is developed for finding the point estimates of unknown parameters. For the construction of the Fisher information matrix, the missing information principle is used. Further, the study is extended to the Bayesian paradigm. In the case of the Bayesian approach, a data augmentation technique is used for sample generation from conditional posteriors. An extensive simulation study is performed for distinct proportions of missing data by taking uniform and exponential as two different patterns of monitoring points. The results under both monitoring points are satisfactory. Finally, considering the first and later born as competing causes for the occurrence of menarche, the age at menarche is estimated using real data.

4.8 Appendix

4.8.1 A-1

In this chapter, we develop mathematical model for recall-based data under competing risks having K components. Here, we give the calculations for two causes setup. Now, under two causes setup, the likelihood function (4.1) can be written as below

$$\begin{aligned} L(\theta|d) = & \prod_{i=1}^{n_{r_1}} \Big[f(t_i,\theta_1)\bar{F}(t_i,\theta_2)\{1-\psi_1(s_i,t_i)\} \Big] \prod_{i=1}^{n_{r_2}} \Big[f(t_i,\theta_2)\bar{F}(t_i,\theta_1)\{1-\psi_2(s_i,t_i)\} \Big] \\ & \prod_{i=1}^{n_{nr}} \Big[\int_0^{s_i} f(u,\theta_1)\bar{F}(u,\theta_2)\psi_1(s_i,u)du + \int_0^{s_i} f(u,\theta_2)\bar{F}(u,\theta_1)\psi_2(s_i,u)du \Big] \\ & \prod_{i=1}^{n_c} \bar{F}(s_i;\theta_1,\theta_2). \end{aligned} \tag{4.23}$$

The symbols n_{r_1} and n_{r_2} denotes the number of individuals who had experienced the event of interest before monitoring time and recalled the exact causes of occurrence of the event as 1 and 2 respectively.

Assume $T \sim \mathcal{W}(\alpha_k, \beta_k)$; $k = 1, 2$ for causes 1 and 2 and denote the parameter vector by $\Theta = (\alpha_1, \beta_1, \lambda_1, \alpha_2, \beta_2, \lambda_2)$. Putting all the functional forms of the densities in (4.23), the likelihood function can be written as below

$$\begin{aligned} L(\Theta|d) \;=\; & \alpha_1^{n_{r_1}} \beta_1^{n_{r_1}} \alpha_2^{n_{r_2}} \beta_2^{n_{r_2}} \prod_{i=1}^{n_{r_1}} t_i^{\alpha_1-1} \exp\{-\beta_1 t_i^{\alpha_1} - \beta_2 t_i^{\alpha_2} - \lambda_1(s_i-t_i)\} \\ & \prod_{i=1}^{n_{r_2}} t_i^{\alpha_2-1} \exp\{-\beta_1 t_i^{\alpha_1} - \beta_2 t_i^{\alpha_2} - \lambda_2(s_i-t_i)\} \prod_{i=1}^{n_c} \exp\{-\beta_1 s_i^{\alpha_1} - \beta_2 s_i^{\alpha_2}\} \\ & \prod_{i=1}^{n_{nr}} \Bigg[\int_0^{s_i} \alpha_1\beta_1 u^{\alpha_1-1} \exp\{-\beta_1 u^{\alpha_1} - \beta_2 u^{\alpha_2}\} \Big(1 - \exp\{-\lambda_1(s_i-u)\}\Big) du \\ & + \int_0^{s_i} \alpha_2\beta_2 u^{\alpha_2-1} \exp\{-\beta_1 u^{\alpha_1} - \beta_2 u^{\alpha_2}\} \Big(1 - \exp\{-\lambda_2(s_i-u)\}\Big) du \Bigg] \end{aligned} \tag{4.24}$$

4.8.2 A-2

The likelihood function (4.24) is incompatible to deal further. Since for the non-recall observations, partial information is available on time to event and causes of the event is also

unknown. We treat it as a missing data problem and apply the E-M algorithm for point estimation.

Expectation Maximization Algorithm

Since any observation under non-recall is exposed to two cause, hence a latent variable Z_i following Bernoulli distribution with probability of success P_i is introduced. The probability of success $(Z_i = 1)$ is defined as

$$P_i = \frac{I_k(s_i)}{\sum\limits_k I_k(s_i)}; \quad i = 1, 2, \ldots, n_{nr}, \ k = 1, 2, \tag{4.25}$$

where, $I_1(.)$ and $I_2(.)$ is given by

$$I_1(s_i) = \int_0^{s_i} \alpha_1\beta_1 u^{\alpha_1 - 1} \exp\{-\beta_1 u^{\alpha_1} - \beta_2 u^{\alpha_2}\}\Big(1 - \exp\{-\lambda_1(s_i - u)\}\Big)du$$

and

$$I_2(s_i) = \int_0^{s_i} \alpha_2\beta_2 u^{\alpha_2 - 1} \exp\{-\beta_1 u^{\alpha_1} - \beta_2 u^{\alpha_2}\}\Big(1 - \exp\{-\lambda_2(s_i - u)\}\Big)du.$$

With the help of these probabilities and latent variable Z_i, we assign exact cause k to the non-recall observations. Now, after introducing the latent variable Z, the likelihood can be written as

$$\begin{aligned} L(\Theta|d) = {} & \alpha_1^{n_{r_1}}\beta_1^{n_{r_1}}\alpha_2^{n_{r_2}}\beta_2^{n_{r_2}} \prod_{i=1}^{n_{r_1}} t_i^{\alpha_1 - 1} \exp\{-\beta_1 t_i^{\alpha_1} - \beta_2 t_i^{\alpha_2} - \lambda_1(s_i - t_i)\} \\ & \prod_{i=1}^{n_{r_2}} t_i^{\alpha_2 - 1} \exp\{-\beta_1 t_i^{\alpha_1} - \beta_2 t_i^{\alpha_2} - \lambda_2(s_i - t_i)\} \prod_{i=1}^{n_c} \exp\{-\beta_1 s_i^{\alpha_1} - \beta_2 s_i^{\alpha_2}\} \\ & \prod_{i=1}^{n_{nr}} \left[\int_0^{s_i} \alpha_1\beta_1 u^{\alpha_1 - 1} \exp\{-\beta_1 u^{\alpha_1} - \beta_2 u^{\alpha_2}\}\Big(1 - \exp\{-\lambda_1(s_i - u)\}\Big)du\right]^{z_i} \\ & \prod_{i=1}^{n_{nr}} \left[\int_0^{s_i} \alpha_2\beta_2 u^{\alpha_2 - 1} \exp\{-\beta_1 u^{\alpha_1} - \beta_2 u^{\alpha_2}\}\Big(1 - \exp\{-\lambda_2(s_i - u)\}\Big)du\right]^{1 - z_i}. \end{aligned} \tag{4.26}$$

Following equivalent quantity approach, in light of given causes, quantity T^* is introduced, which lies in the interval $(0, S)$. Now the likelihood function (4.26) in the light of conditional density can be re-written as

$$\begin{aligned} L(\Theta|d) = \alpha_1^{n_{r_1}}\beta_1^{n_{r_1}}\alpha_2^{n_{r_2}}\beta_2^{n_{r_2}}\prod_{i=1}^{n_{r_1}} t_i^{\alpha_1-1}\exp\{-\beta_1 t_i^{\alpha_1}-\beta_2 t_i^{\alpha_2}-\lambda_1(s_i-t_i)\} \\ \prod_{i=1}^{n_{r_2}} t_i^{\alpha_2-1}\exp\{-\beta_1 t_i^{\alpha_1}-\beta_2 t_i^{\alpha_2}-\lambda_2(s_i-t_i)\}\prod_{i=1}^{n_c}\exp\{-\beta_1 s_i^{\alpha_1}-\beta_2 s_i^{\alpha_2}\} \\ \prod_{i=1}^{n_{nr}}\Big[\alpha_1\beta_1 t_i^{*\alpha_1-1}\exp\{-\beta_1 t_i^{*\alpha_1}-\beta_2 t_i^{*\alpha_2}\}\Big(1-\exp\{-\lambda_1(s_i-t_i^*)\}\Big)\Big]^{z_i} \\ \prod_{i=1}^{n_{nr}}\Big[\alpha_2\beta_2 t_i^{*\alpha_2-1}\exp\{-\beta_1 t_i^{*\alpha_1}-\beta_2 t_i^{*\alpha_2}\}\Big(1-\exp\{-\lambda_2(s_i-t_i^*)\}\Big)\Big]^{1-z_i}. \quad (4.27) \end{aligned}$$

Also, for making the functional form of non-recall probabilities compatible, we introduce two latent variables U^* and V^* corresponding to causes 1 and 2 following exponential distribution with mean $\frac{1}{\lambda_1}$ and $\frac{1}{\lambda_2}$ truncated at $(S-T^*)$. For non-recall observations, let us denote the missing data by $\underline{D}^* = (Z, T^*, U^*, V^*)$. For i^{th} observation belonging to non-recall category, denote the missing data by $d^* = (z_i, t_i^*, u_i^*, v_i^*)$. The complete data can be written as (d, d^*). Now under the complete data the likelihood function (4.27) can be written as

$$\begin{aligned} L_c(\Theta|d,d^*) = \alpha_1^{n_{r_1}+Z^*}\beta_1^{n_{r_1}+Z^*}\lambda_1^{Z^*}\alpha_2^{n_{r_2}+n_{nr}-Z^*}\beta_2^{n_{r_2}+n_{nr}-Z^*}\lambda_2^{n_{nr}-Z^*} \\ \prod_{i=1}^{n_{r_1}} t_i^{\alpha_1-1}\exp\{-\beta_1 t_i^{\alpha_1}-\beta_2 t_i^{\alpha_2}-\lambda_1(s_i-t_i)\} \\ \prod_{i=1}^{n_{r_2}} t_i^{\alpha_2-1}\exp\{-\beta_1 t_i^{\alpha_1}-\beta_2 t_i^{\alpha_2}-\lambda_2(s_i-t_i)\} \\ \prod_{i=1}^{n_{nr}}\big[t_i^{*\alpha_1-1}\exp\{-\beta_1 t_i^{*\alpha_1}-\beta_2 t_i^{*\alpha_2}-\lambda_1 u_i^*\}\big]^{z_i} \\ \prod_{i=1}^{n_{nr}}\big[t_i^{*\alpha_2-1}\exp\{-\beta_1 t_i^{*\alpha_1}-\beta_2 t_i^{*\alpha_2}-\lambda_2 v_i^*\}\big]^{1-z_i}\prod_{i=1}^{n_c}\exp\{-\beta_1 s_i^{\alpha_1}-\beta_2 s_i^{\alpha_2}\}, \end{aligned} \quad (4.28)$$

Taking natural logarithm of (4.28), the log-likelihood function can be written as

$$
\begin{aligned}
l_c(\Theta|d,d^*) &= (n_{r_1}+Z^*)\ln(\alpha_1)+(n_{r_1}+Z^*)\ln(\beta_1)+Z^*\ln(\lambda_1)+(n_{r_2}+n_{nr}-Z^*)\ln(\alpha_2)\\
&+(n_{r_2}+n_{nr}-Z^*)\ln(\beta_2)+(n_{nr}-Z^*)\ln(\lambda_2)+(\alpha_1-1)\sum_{i=1}^{n_{r_1}}\ln(t_i)-\beta_1\sum_{i=1}^{n_{r_1}}t_i^{\alpha_1}\\
&-\beta_2\sum_{i=1}^{n_{r_1}}t_i^{\alpha_2}-\lambda_1\sum_{i=1}^{n_{r_1}}(s_i-t_i)+(\alpha_2-1)\sum_{i=1}^{n_{r_2}}\ln(t_i)-\beta_1\sum_{i=1}^{n_{r_2}}t_i^{\alpha_1}-\beta_2\sum_{i=1}^{n_{r_2}}t_i^{\alpha_2}\\
&-\lambda_2\sum_{i=1}^{n_{r_2}}(s_i-t_i)+(\alpha_1-1)\sum_{i=1}^{n_{nr}}z_i\ln(t_i^*)-\beta_1\sum_{i=1}^{n_{nr}}z_i t_i^{*\alpha_1}-\beta_2\sum_{i=1}^{n_{nr}}z_i t_i^{*\alpha_2}-\lambda_1\sum_{i=1}^{n_{nr}}z_i u_i^*\\
&+(\alpha_2-1)\sum_{i=1}^{n_{nr}}(1-z_i)\ln(t_i^*)-\beta_1\sum_{i=1}^{n_{nr}}(1-z_i)t_i^{*\alpha_1}-\beta_2\sum_{i=1}^{n_{nr}}(1-z_i)t_i^{*\alpha_2}\\
&-\lambda_2\sum_{i=1}^{n_{nr}}(1-z_i)v_i^*-\beta_1\sum_{i=1}^{n_c}s_i^{\alpha_1}-\beta_2\sum_{i=1}^{n_c}s_i^{\alpha_2}. \qquad (4.29)
\end{aligned}
$$

We define the quantity Q as

$$
\begin{aligned}
Q\Big(\Theta|\hat{\Theta}^{(m)}\Big) &= (n_{r_1}+P)\ln(\alpha_1)+(n_{r_1}+P)\ln(\beta_1)+P\ln(\lambda_1)+(n_{r_2}+n_{nr}-P)\ln(\alpha_2)\\
&+(n_{r_2}+n_{nr}-P)\ln(\beta_2)+(n_{nr}-P)\ln(\lambda_2)+(\alpha_1-1)\sum_{i=1}^{n_{r_1}}\ln(t_i)-\beta_1\sum_{i=1}^{n_{r_1}}t_i^{\alpha_1}\\
&-\beta_2\sum_{i=1}^{n_{r_1}}t_i^{\alpha_2}-\lambda_1\sum_{i=1}^{n_{r_1}}(s_i-t_i)+(\alpha_2-1)\sum_{i=1}^{n_{r_2}}\ln(t_i)-\beta_1\sum_{i=1}^{n_{r_2}}t_i^{\alpha_1}-\beta_2\sum_{i=1}^{n_{r_2}}t_i^{\alpha_2}\\
&-\lambda_2\sum_{i=1}^{n_{r_2}}(s_i-t_i)+(\alpha_1-1)\sum_{i=1}^{n_{nr}}P_iE\Big[\ln(t_i^*)\mid t_i^*<s_i\Big]-\beta_1\sum_{i=1}^{n_{nr}}P_iE\Big[t_i^{*\alpha_1}\mid t_i^*<s_i\Big]\\
&-\beta_2\sum_{i=1}^{n_{nr}}P_iE\Big[t_i^{*\alpha_2}\mid t_i^*<s_i\Big]-\lambda_1\sum_{i=1}^{n_{nr}}P_iE\Big[u_i^*\mid u_i^*<(s_i-t_i^*)\Big]\\
&+(\alpha_2-1)\sum_{i=1}^{n_{nr}}(1-P_i)E\Big[\ln(t_i^*)\mid t_i^*<s_i\Big]-\beta_1\sum_{i=1}^{n_{nr}}(1-P_i)E\Big[t_i^{*\alpha_1}\mid t_i^*<s_i\Big]\\
&-\beta_2\sum_{i=1}^{n_{nr}}(1-P_i)E\Big[t_i^{*\alpha_2}\mid t_i^*<s_i\Big]-\lambda_2\sum_{i=1}^{n_{nr}}(1-P_i)E\Big[v_i^*\mid v_i^*<(s_i-t_i^*)\Big]\\
&-\beta_1\sum_{i=1}^{n_c}s_i^{\alpha_1}-\beta_2\sum_{i=1}^{n_c}s_i^{\alpha_2}. \qquad (4.30)
\end{aligned}
$$

For finding the estimators of unknown parameters, differentiate (4.29) with respect to the parameters and we get

$$\frac{\partial Q}{\partial \alpha_1} = \frac{(n_{r_1}+P)}{\alpha_1} + \sum_{i=1}^{n_{r_1}} \ln(t_i) - \beta_1 \sum_{i=1}^{n_{r_1}} t_i^{\alpha_1} \ln(t_i) - \beta_1 \sum_{i=1}^{n_{r_2}} t_i^{\alpha_1} \ln(t_i) + \sum_{i=1}^{n_{nr}} P_i E\Big[\ln(t_i^*) \mid t_i^* < s_i\Big]$$
$$- \beta_1 \sum_{i=1}^{n_{nr}} P_i E\Big[t_i^{*\alpha_1} \ln(t_i^*) \mid t_i^* < s_i\Big] - \beta_1 \sum_{i=1}^{n_{nr}} (1-P_i) E\Big[t_i^{*\alpha_1} \ln(t_i^*) \mid t_i^* < s_i\Big]$$
$$- \beta_1 \sum_{i=1}^{n_c} s_i^{\alpha_1} \ln(s_i), \tag{4.31}$$

$$\frac{\partial Q}{\partial \beta_1} = \frac{(n_{r_1}+P)}{\beta_1} - \sum_{i=1}^{n_{r_1}} t_i^{\alpha_1} - \sum_{i=1}^{n_{r_2}} t_i^{\alpha_1} - \sum_{i=1}^{n_{nr}} P_i E\Big[t_i^{*\alpha_1} \mid t_i^* < s_i\Big] - \sum_{i=1}^{n_{nr}} (1-P_i) E\Big[t_i^{*\alpha_1} \mid t_i^* < s_i\Big]$$
$$- \sum_{i=1}^{n_c} s_i^{\alpha_1}, \tag{4.32}$$

$$\frac{\partial Q}{\partial \lambda_1} = \frac{P}{\lambda_1} - \sum_{i=1}^{n_{r_1}} (s_i - t_i) - \sum_{i=1}^{n_{nr}} P_i E\Big[u_i^* \mid u_i^* < (s_i - t_i^*)\Big], \tag{4.33}$$

$$\frac{\partial Q}{\partial \alpha_2} = \frac{(n_{r_2}+n_{nr}-P)}{\alpha_2} + \sum_{i=1}^{n_{r_2}} \ln(t_i) - \beta_2 \sum_{i=1}^{n_{r_1}} t_i^{\alpha_2} \ln(t_i) - \beta_2 \sum_{i=1}^{n_{r_2}} t_i^{\alpha_2} \ln(t_i)$$
$$- \beta_2 \sum_{i=1}^{n_{nr}} P_i E\Big[t_i^{*\alpha_2} \ln(t_i^*) \mid t_i^* < s_i\Big] + \sum_{i=1}^{n_{nr}} (1-P_i) E\Big[\ln(t_i^*) \mid t_i^* < s_i\Big]$$
$$- \beta_2 \sum_{i=1}^{n_{nr}} (1-P_i) E\Big[t_i^{*\alpha_2} \ln(t_i^*) \mid t_i^* < s_i\Big] - \beta_2 \sum_{i=1}^{n_c} s_i^{\alpha_2} \ln(s_i), \tag{4.34}$$

$$\frac{\partial Q}{\partial \beta_2} = \frac{(n_{r_2}+n_{nr}-P)}{\beta_2} - \sum_{i=1}^{n_{r_1}} t_i^{\alpha_1} - \sum_{i=1}^{n_{r_2}} t_i^{\alpha_2} - \sum_{i=1}^{n_{nr}} P_i E\Big[t_i^{*\alpha_2} \mid t_i^* < s_i\Big]$$
$$- \sum_{i=1}^{n_{nr}} (1-P_i) E\Big[t_i^{*\alpha_2} \mid t_i^* < s_i\Big] - \sum_{i=1}^{n_c} s_i^{\alpha_2}, \tag{4.35}$$

$$\frac{\partial Q}{\partial \lambda_2} = \frac{(n_{nr}-P)}{\lambda_2} - \sum_{i=1}^{n_{r_2}} (s_i - t_i) - \sum_{i=1}^{n_{nr}} (1-P_i) E\Big[v_i^* \mid v_i^* < (s_i - t_i^*)\Big], \tag{4.36}$$

Since $z_i \sim \mathcal{B}(1, P_i)$, hence $E(z_i) = P_i$ and $E(Z^*) = E(\sum_{i=1}^{n_{nr}} z_i) = P$. The conditional densities of U^* and V^* can be calculated by using density below

$$h_Y\Big(y_i | y_i < (s_i - t_i^*), z_i, \lambda_k\Big) = \frac{h(y_i; \lambda_k)}{1 - \bar{H}\Big((s_i - t_i^*); \lambda_k\Big)} = \frac{\lambda_k \exp\{-\lambda_k y_i\}}{1 - \exp\{-\lambda_k (s_i - t_i^*)\}} \tag{4.37}$$

where $h(.)$ and $\bar{H}(.)$ denotes the truncated density and survival functions of Y at point $(S - T^*)$.

The expectation terms used in E-step of the E-M algorithm can be calculated from

conditional densities defined in (4.5) and (4.37) and given below

$$\xi_1(t_i^*;\alpha_1,\beta_1) = E\Big[t_i^{*\alpha_1}\Big|t_i^* < s_i, k=1,\alpha_1,\beta_1\Big] = \frac{\int_0^{s_i} u^{2\alpha_1-1}\alpha_1\beta_1\exp\{-\beta_1 u^{\alpha_1}\}du}{1-\exp\{-\beta_1 s_i^{\alpha_1}\}}$$
$$= \frac{1-(1+\beta_1 s_i^{\alpha_1})\exp\{-\beta_1 s_i^{\alpha_1}\}}{\beta_1\Big(1-\exp\{-\beta_1 s_i^{\alpha_1}\}\Big)},$$

$$\xi_2(t_i^*;\alpha_1,\beta_1,\alpha_2) = E\Big[t_i^{*\alpha_2}\Big|t_i^* < s_i, k=1,\alpha_1,\beta_1\Big] = \frac{\int_0^{s_i} u^{\alpha_1+\alpha_2-1}\alpha_1\beta_1\exp\{-\beta_1 u^{\alpha_1}\}du}{1-\exp\{-\beta_1 s_i^{\alpha_1}\}}$$
$$= \frac{\Gamma\left(\frac{\alpha_1+\alpha_2}{\alpha_1}\right)-\Gamma\left(\frac{\alpha_1+\alpha_2}{\alpha_1},\beta_1 s_i^{\alpha_1}\right)}{(\beta_1)^{\alpha_2/\alpha_1}\Big(1-\exp\{-\beta_1 s_i^{\alpha_1}\}\Big)},$$

$$\xi_3(t_i^*;\alpha_1,\beta_1) = E\Big[\ln(t_i^*)\Big|t_i^* < s_i, k=1,\alpha_1,\beta_1\Big] = \frac{\int_0^{s_i} u^{\alpha_1-1}\ln(u)\alpha_1\beta_1\exp\{-\beta_1 u^{\alpha_1}\}du}{1-\exp\{-\beta_1 s_i^{\alpha_1}\}}$$
$$= \frac{K_1(s_i)-\ln(\beta_1)\Big(1-\exp\{-\beta_1 s_i^{\alpha_1}\}\Big)}{\alpha_1\Big(1-\exp\{-\beta_1 s_i^{\alpha_1}\}\Big)},$$

$$\xi_4(t_i^*;\alpha_1,\beta_1) = E\Big[t_i^{*\alpha_1}\ln(t_i^*)\Big|t_i^* < s_i, k=1,\alpha_1,\beta_1\Big] = \frac{\int_0^{s_i} u^{2\alpha_1-1}\ln(u)\alpha_1\beta_1\exp\{-\beta_1 u^{\alpha_1}\}du}{1-\exp\{-\beta_1 s_i^{\alpha_1}\}}$$
$$= \frac{K_2(s_i)-\ln(\beta_1)\Big(1-(1+\beta_1 s_i^{\alpha_1})\exp\{-\beta_1 s_i^{\alpha_1}\}\Big)}{\alpha_1\beta_1\Big(1-\exp\{-\beta_1 s_i^{\alpha_1}\}\Big)},$$

$$\xi_5(t_i^*;\alpha_1,\beta_1,\alpha_2) = E\Big[t_i^{*\alpha_2}\ln(t_i^*)|t_i^* < s_i, k=1,\alpha_1,\beta_1\Big]$$
$$= \frac{\int_0^{s_i} u^{\alpha_1+\alpha_2-1}\ln(u)\alpha_1\beta_1\exp\{-\beta_1 u^{\alpha_1}\}du}{1-\exp\{-\beta_1 s_i^{\alpha_1}\}}$$
$$= \frac{K_3(s_i)-\ln(\beta_1)\left\{\Gamma\left(\frac{\alpha_1+\alpha_2}{\alpha_1}\right)-\Gamma\left(\frac{\alpha_1+\alpha_2}{\alpha_1},\beta_1 s_i^{\alpha_1}\right)\right\}}{\alpha_1(\beta_1)^{\alpha_2/\alpha_1}\Big(1-\exp\{-\beta_1 s_i^{\alpha_1}\}\Big)},$$

$$\xi_6(u_i^*;\lambda_1) = E\Big[u_i^*|u_i^* < (s_i-t_i^*), k=1,\lambda_1\Big] = \frac{\int_0^{s_i-t_i^*} u\lambda_1\exp\{-\lambda_1 u\}}{1-\exp\{-\lambda_1(s_i-t_i^*)\}}$$
$$= \frac{1}{\lambda_1}\left[\frac{1-\{1+\lambda_1(s_i-t_i^*)\}\exp\{-\lambda_1(s_i-t_i^*)\}}{1-\exp\{-\lambda_1(s_i-t_i^*)\}}\right],$$

$$\xi_7(t_i^*;\alpha_1,\beta_1) = E\Big[t_i^{*\alpha_1}\Big(\ln(t_i^*)\Big)^2\Big|t_i^* < s_i, k=1,\alpha_1,\beta_1\Big]$$
$$= \frac{\int_0^{s_i} u^{2\alpha_1-1}\Big(\ln(u)\Big)^2\alpha_1\beta_1\exp\{-\beta_1 u^{\alpha_1}\}du}{1-\exp\{-\beta_1 s_i^{\alpha_1}\}}$$
$$= \frac{K_4(s_i)+\Big(\ln(\beta_1)\Big)^2\Big(1-(1+\beta_1 s_i^{\alpha_1})\exp\{-\beta_1 s_i^{\alpha_1}\}\Big)-2\ln(\beta_1)K_2(s_i)}{\alpha_1^2\beta_1\Big(1-\exp\{-\beta_1 s_i^{\alpha_1}\}\Big)},$$

$$\xi_8(t_i^*;\alpha_1,\beta_1,\alpha_2) = E\Big[t_i^{*\alpha_2}\Big(\ln(t_i^*)\Big)^2\Big|t_i^* < s_i, k=1,\alpha_1,\beta_1\Big]$$
$$= \frac{\int_0^{s_i} u^{\alpha_1+\alpha_2-1}\Big(\ln(u)\Big)^2\alpha_1\beta_1\exp\{-\beta_1 u^{\alpha_1}\}du}{1-\exp\{-\beta_1 s_i^{\alpha_1}\}}$$

$$= \frac{K_5(s_i) + \Big(\ln(\beta_1)\Big)^2 \Big\{\Gamma\Big(\frac{\alpha_1+\alpha_2}{\alpha_1}\Big) - \Gamma\Big(\frac{\alpha_1+\alpha_2}{\alpha_1}, \beta_1 s_i^{\alpha_1}\Big)\Big\} - 2\ln(\beta_1)K_3(s_i)}{\alpha_1^2(\beta_1)^{\alpha_2/\alpha_1}\Big(1-\exp\{-\beta_1 s_i^{\alpha_1}\}\Big)},$$

$$\xi_9(t_i^*;\alpha_1,\alpha_2,\beta_2) = E\Big[t_i^{*\alpha_1}\Big|t_i^* < s_i, k=2,\alpha_2,\beta_2\Big] = \frac{\int_0^{s_i} u^{\alpha_1+\alpha_2-1}\alpha_2\beta_2\exp\{-\beta_2 u^{\alpha_2}\}du}{1-\exp\{-\beta_2 s_i^{\alpha_2}\}}$$

$$= \frac{\Gamma\Big(\frac{\alpha_1+\alpha_2}{\alpha_2}\Big) - \Gamma\Big(\frac{\alpha_1+\alpha_2}{\alpha_2}, \beta_2 s_i^{\alpha_2}\Big)}{(\beta_2)^{\alpha_1/\alpha_2}\Big(1-\exp\{-\beta_2 s_i^{\alpha_2}\}\Big)},$$

$$\xi_{10}(t_i^*;\alpha_2,\beta_2) = E\Big[t_i^{*\alpha_2}\Big|t_i^* < s_i, k=2,\alpha_2,\beta_2\Big] = \frac{\int_0^{s_i} u^{2\alpha_2-1}\alpha_2\beta_2\exp\{-\beta_2 u^{\alpha_2}\}du}{1-\exp\{-\beta_2 s_i^{\alpha_2}\}}$$

$$= \frac{1-(1+\beta_2 s_i^{\alpha_2})\exp\{-\beta_2 s_i^{\alpha_2}\}}{\beta_2\Big(1-\exp\{-\beta_2 s_i^{\alpha_2}\}\Big)},$$

$$\xi_{11}(t_i^*;\alpha_2,\beta_2) = E\Big[\ln(t_i^*)\Big|t_i^* < s_i, k=2,\alpha_2,\beta_2\Big] = \frac{\int_0^{s_i} u^{\alpha_2-1}\ln(u)\alpha_2\beta_2\exp\{-\beta_2 u^{\alpha_2}\}du}{1-\exp\{-\beta_2 s_i^{\alpha_2}\}}$$

$$= \frac{K_6(s_i) - \ln(\beta_2)\Big(1-\exp\{-\beta_2 s_i^{\alpha_2}\}\Big)}{\alpha_2\Big(1-\exp\{-\beta_2 s_i^{\alpha_2}\}\Big)},$$

$$\xi_{12}(t_i^*;\alpha_1,\alpha_2,\beta_2) = E\Big[t_i^{*\alpha_1}\ln(t_i^*)\Big|t_i^* < s_i, k=2,\alpha_2,\beta_2\Big] = \frac{\int_0^{s_i} u^{\alpha_1+\alpha_2-1}\ln(u)\alpha_2\beta_2\exp\{-\beta_2 u^{\alpha_2}\}du}{1-\exp\{-\beta_2 s_i^{\alpha_2}\}}$$

$$= \frac{K_7(s_i) - \ln(\beta_2)\Big\{\Gamma\Big(\frac{\alpha_1+\alpha_2}{\alpha_2}\Big) - \Gamma\Big(\frac{\alpha_1+\alpha_2}{\alpha_2}, \beta_2 s_i^{\alpha_2}\Big)\Big\}}{\alpha_2(\beta_2)^{\alpha_1/\alpha_2}\Big(1-\exp\{-\beta_2 s_i^{\alpha_2}\}\Big)},$$

$$\xi_{13}(t_i^*;\alpha_2,\beta_2) = E\Big[t_i^{*\alpha_2}\ln(t_i^*)\Big|t_i^* < s_i, k=2,\alpha_2,\beta_2\Big] = \frac{\int_0^{s_i} u^{2\alpha_2-1}\ln(u)\alpha_2\beta_2\exp\{-\beta_2 u^{\alpha_2}\}du}{1-\exp\{-\beta_2 s_i^{\alpha_2}\}}$$

$$= \frac{K_8(s_i) - \ln(\beta_2)\Big(1-(1+\beta_2 s_i^{\alpha_2})\exp\{-\beta_2 s_i^{\alpha_2}\}\Big)}{\alpha_2\beta_2\Big(1-\exp\{-\beta_2 s_i^{\alpha_2}\}\Big)},$$

$$\xi_{14}(v_i^*;\lambda_2) = E\Big[v_i^*\Big|v_i^* < (s_i - t_i^*), k=2,\lambda_2\Big] = \frac{\int_0^{s_i-t_i^*} u\lambda_2\exp\{-\lambda_2 u\}}{1-\exp\{-\lambda_2(s_i-t_i^*)\}}$$

$$= \frac{1}{\lambda_2}\left[\frac{1-\Big(1+\lambda_2(s_i-t_i^*)\Big)\exp\{-\lambda_2(s_i-t_i^*)\}}{1-\exp\{-\lambda_2(s_i-t_i^*)\}}\right],$$

$$\xi_{15}(t_i^*;\alpha_1,\alpha_2,\beta_2) = E\Big[t_i^{*\alpha_1}\Big(\ln(t_i^*)\Big)^2\Big|t_i^* < s_i, k=2,\alpha_2,\beta_2\Big]$$

$$= \frac{\int_0^{s_i} u^{\alpha_1+\alpha_2-1}\Big(\ln(u)\Big)^2\alpha_2\beta_2\exp\{-\beta_2 u^{\alpha_2}\}du}{1-\exp\{-\beta_2 s_i^{\alpha_2}\}}$$

$$= \frac{K_9(s_i) + \Big(\ln(\beta_2)\Big)^2\Big\{\Gamma\Big(\frac{\alpha_1+\alpha_2}{\alpha_2}\Big) - \Gamma\Big(\frac{\alpha_1+\alpha_2}{\alpha_2}, \beta_2 s_i^{\alpha_2}\Big)\Big\} - 2\ln(\beta_2)K_7(s_i)}{\alpha_2^2(\beta_2)^{\alpha_1/\alpha_2}\Big(1-\exp\{-\beta_2 s_i^{\alpha_2}\}\Big)},$$

$$\xi_{16}(t_i^*;\alpha_2,\beta_2) = E\Big[t_i^{*\alpha_2}\Big(\ln(t_i^*)\Big)^2\Big|t_i^* < s_i, k=2,\alpha_2,\beta_2\Big]$$

$$= \frac{\int_0^{s_i} u^{2\alpha_2-1}\Big(\ln(u)\Big)^2\alpha_2\beta_2\exp\{-\beta_2 u^{\alpha_2}\}du}{1-\exp\{-\beta_2 s_i^{\alpha_2}\}}$$

$$= \frac{K_{10}(s_i) + \Big(\ln(\beta_2)\Big)^2\Big(1-(1+\beta_2 s_i^{\alpha_2})\exp\{-\beta_2 s_i^{\alpha_2}\}\Big) - 2\ln(\beta_2)K_8(s_i)}{\alpha_2^2\beta_2\Big(1-\exp\{-\beta_2 s_i^{\alpha_2}\}\Big)},$$

where,

$$K_1(y) = \int_0^{\beta_1 y^{\alpha_1}} \ln(y)\exp\{-y\}dy, \quad K_2(y) = \int_0^{\beta_1 y^{\alpha_1}} y\ln(y)\exp\{-y\}dy,$$
$$K_3(y) = \int_0^{\beta_1 y^{\alpha_1}} y^{\alpha_2/\alpha_1}\ln(y)\exp\{-y\}dz, \quad K_4(y) = \int_0^{\beta_1 y^{\alpha_1}} y\Big(\ln(y)\Big)^2\exp\{-y\}dy$$
$$K_5(z) = \int_0^{\beta_1 y^{\alpha_1}} y^{\alpha_2/\alpha_1}\Big(\ln(y)\Big)^2\exp\{-y\}dy, \quad K_6(y) = \int_0^{\beta_2 y^{\alpha_2}} \ln(y)\exp\{-y\}dy,$$
$$K_7(y) = \int_0^{\beta_2 y^{\alpha_2}} y^{\alpha_1/\alpha_2}\ln(y)\exp\{-y\}dy, \quad K_8(y) = \int_0^{\beta_2 y^{\alpha_2}} y\ln(y)\exp\{-y\}dy$$
$$K_9(y) = \int_0^{\beta_2 y^{\alpha_2}} y^{\alpha_1/\alpha_2}\Big(\ln(y)\Big)^2\exp\{-y\}dy, \quad K_{10}(y) = \int_0^{\beta_2 y^{\alpha_2}} y\Big(\ln(y)\Big)^2\exp\{-y\}dy.$$

After updating the missing data with help of expectation terms calculated in the E-step, the log-likelihood is maximized in the M-step and let at the m^{th} iteration $\hat{\Theta}^{(m)}$ be the vector of unknown parameter's estimate of Θ. Then the updated value of parameters at $(m+1)^{th}$ iteration can be computed by using the following expressions

$$\hat{\alpha}_1^{(m+1)} = \frac{\Big(n_{r_1} + P^{(m)}\Big)}{\left[\begin{array}{l} -\sum_{i=1}^{n_{r_1}}\ln(t_i) + \hat{\beta}_1^{(m)}\sum_{i=1}^{n_{r_1}} t_i^{\hat{\alpha}_1^{(m)}}\ln(t_i) + \hat{\beta}_1^{(m)}\sum_{i=1}^{n_{r_2}} t_i^{\hat{\alpha}_1^{(m)}}\ln(t_i) - \sum_{i=1}^{n_{nr}} P_i{}^{(m)}\xi_3\Big(t_i^*;\alpha_1{}^{(m)},\beta_1{}^{(m)}\Big) \\ \quad + \hat{\beta}_1^{(m)}\sum_{i=1}^{n_{nr}} P_i{}^{(m)}\xi_4\Big(t_i^*;\alpha_1{}^{(m)},\beta_1{}^{(m)}\Big) + \hat{\beta}_1^{(m)}\sum_{i=1}^{n_{nr}}\Big(1-P_i{}^{(m)}\Big)\xi_{12}\Big(t_i^*;\alpha_1{}^{(m)},\alpha_2{}^{(m)},\beta_2{}^{(m)}\Big) + \hat{\beta}_1^{(m)}\sum_{i=1}^{n_c} s_i^{\hat{\alpha}_1^{(m)}}\ln(s_i) \end{array}\right]} \tag{4.38}$$

$$\hat{\alpha}_2^{(m+1)} = \frac{\Big(n_{r_2} + n_{nr} - P^{(m)}\Big)}{\left[\begin{array}{l} -\sum_{i=1}^{n_{r_2}}\ln(t_i) + \hat{\beta}_2^{(m)}\sum_{i=1}^{n_{r_1}} t_i^{\hat{\alpha}_2^{(m)}}\ln(t_i) + \hat{\beta}_2^{(m)}\sum_{i=1}^{n_{r_2}} t_i^{\hat{\alpha}_2^{(m)}}\ln(t_i) + \hat{\beta}_2^{(m)}\sum_{i=1}^{n_{nr}} P_i{}^{(m)}\xi_5\Big(t_i^*;\alpha_1{}^{(m)},\alpha_2{}^{(m)},\beta_2{}^{(m)}\Big) \\ \quad - \sum_{i=1}^{n_{nr}}\Big(1-P_i{}^{(m)}\Big)\xi_{11}\Big(t_i^*;\alpha_2{}^{(m)},\beta_2{}^{(m)}\Big) + \hat{\beta}_2^{(m)}\sum_{i=1}^{n_{nr}}(1-P_i{}^{(m)})\xi_{13}\Big(t_i^*;\alpha_2{}^{(m)},\beta_2{}^{(m)}\Big) + \hat{\beta}_2^{(m)}\sum_{i=1}^{n_c} s_i^{\hat{\alpha}_2^{(m)}}\ln(s_i) \end{array}\right]} \tag{4.39}$$

$$\hat{\beta}_1^{(m+1)} = \frac{\Big(n_{r_1} + P^{(m)}\Big)}{\left[\begin{array}{l} \sum_{i=1}^{r_1} t_i^{\hat{\alpha}_1^{(m+1)}} + \sum_{i=1}^{r_2} t_i^{\hat{\alpha}_1^{(m+1)}} + \sum_{i=1}^{n_{nr}} P_i{}^{(m)}\xi_1\Big(t_i^*;\alpha_1{}^{(m)},\beta_1{}^{(m)}\Big) \\ \quad + \sum_{i=1}^{n_{nr}}\Big(1-P_i{}^{(m)}\Big)\xi_9\Big(t_i^*;\alpha_1{}^{(m)},\alpha_2{}^{(m)},\beta_2{}^{(m)}\Big) + \sum_{i=1}^{n_c} s_i^{\hat{\alpha}_1^{(m+1)}} \end{array}\right]} \tag{4.40}$$

$$\hat{\lambda}_1^{(m+1)} = \frac{P^{(m)}}{\sum_{i=1}^{n_{r_1}}(s_i - t_i) + \sum_{i=1}^{n_{nr}} P_i{}^{(m)}\xi_6\Big(u_i^*;\lambda_1{}^{(m)}\Big)} \tag{4.41}$$

$$\hat{\beta}_2^{(m+1)} = \frac{\Big(n_{r_2} + n_{nr} - P^{(m)}\Big)}{\left[\begin{array}{l} \sum_{i=1}^{r_1} t_i^{\hat{\alpha}_2^{(m+1)}} + \sum_{i=1}^{r_2} t_i^{\hat{\alpha}_2^{(m+1)}} + \sum_{i=1}^{n_{nr}} P_i{}^{(m)}\xi_2\Big(t_i^*;\alpha_1{}^{(m)},\beta_1{}^{(m)},\alpha_2{}^{(m)}\Big) \\ \quad + \sum_{i=1}^{n_{nr}}\Big(1-P_i{}^{(m)}\Big)\xi_{10}\Big(t_i^*;\alpha_2{}^{(m)},\beta_2{}^{(m)}\Big) + \sum_{i=1}^{n_c} s_i^{\hat{\alpha}_2^{(m+1)}} \end{array}\right]} \tag{4.42}$$

$$\hat{\lambda}_2^{(m+1)} = \frac{\Big(n_{nr} - P^{(m)}\Big)}{\sum_{i=1}^{n_{r_2}}(s_i - t_i) + \sum_{i=1}^{n_{nr}}\Big(1-P_i{}^{(m)}\Big)\xi_{14}\Big(v_i^*;\lambda_2{}^{(m)}\Big)} \tag{4.43}$$

The estimates of parameters can be computed by iterating (4.38) to (4.43) and iterations can be terminated when $|\hat{\Theta}^{(m+1)} - \hat{\Theta}^{(m)}| < \epsilon$, where $\epsilon > 0$ is a sufficiently small real number.

Observed Fisher Information Matrix

Te structure of observed Fisher information matrix is given by

$$I(\hat{\Theta}) = \left[\begin{array}{cccc:cccc} -\frac{\partial^2 l_c}{\partial\alpha_1^2} - \left(\frac{\partial Q}{\partial\alpha_1}\right)^2 & \cdots & -\frac{\partial^2 Q}{\partial\alpha_1\partial\lambda_1} - \frac{\partial Q}{\partial\alpha_1}\frac{\partial Q}{\partial\lambda_1} & & -\frac{\partial^2 Q}{\partial\alpha_1\partial\alpha_2} - \frac{\partial Q}{\partial\alpha_1}\frac{\partial Q}{\partial\alpha_2} & \cdots & -\frac{\partial^2 Q}{\partial\alpha_1\partial\lambda_2} - \frac{\partial Q}{\partial\alpha_1}\frac{\partial Q}{\partial\lambda_2} \\ -\frac{\partial^2 Q}{\partial\beta_1\partial\alpha_1} - \frac{\partial Q}{\partial\beta_1}\frac{\partial Q}{\partial\alpha_1} & \cdots & -\frac{\partial^2 Q}{\partial\beta_1\partial\lambda_1} - \frac{\partial Q}{\partial\beta_1}\frac{\partial Q}{\partial\lambda_1} & & -\frac{\partial^2 Q}{\partial\beta_1\partial\alpha_2} - \frac{\partial Q}{\partial\beta_1}\frac{\partial Q}{\partial\alpha_2} & \cdots & -\frac{\partial^2 Q}{\partial\beta_1\partial\lambda_2} - \frac{\partial Q}{\partial\beta_1}\frac{\partial Q}{\partial\lambda_2} \\ -\frac{\partial^2 Q}{\partial\lambda_1\partial\alpha_1} - \frac{\partial Q}{\partial\lambda_1}\frac{\partial Q}{\partial\alpha_1} & \cdots & -\frac{\partial^2 Q}{\partial\lambda_1^2} - \left(\frac{\partial Q}{\partial\lambda_1}\right)^2 & & -\frac{\partial^2 Q}{\partial\lambda_1\partial\alpha_2} - \frac{\partial Q}{\partial\lambda_1}\frac{\partial Q}{\partial\alpha_2} & \cdots & -\frac{\partial^2 Q}{\partial\lambda_1\partial\lambda_2} - \frac{\partial Q}{\partial\lambda_1}\frac{\partial Q}{\partial\lambda_2} \\ \hdashline -\frac{\partial^2 Q}{\partial\alpha_2\partial\alpha_1} - \frac{\partial Q}{\partial\alpha_2}\frac{\partial Q}{\partial\alpha_1} & \cdots & -\frac{\partial^2 Q}{\partial\alpha_2\partial\lambda_1} - \frac{\partial Q}{\partial\alpha_2}\frac{\partial Q}{\partial\lambda_1} & & -\frac{\partial^2 Q}{\partial\alpha_2^2} - \left(\frac{\partial Q}{\partial\alpha_2}\right)^2 & \cdots & -\frac{\partial^2 Q}{\partial\alpha_2\partial\lambda_2} - \frac{\partial Q}{\partial\alpha_2}\frac{\partial Q}{\partial\lambda_2} \\ -\frac{\partial^2 Q}{\partial\beta_2\partial\alpha_1} - \frac{\partial Q}{\partial\beta_2}\frac{\partial Q}{\partial\alpha_1} & \cdots & -\frac{\partial^2 Q}{\partial\beta_2\partial\lambda_1} - \frac{\partial Q}{\partial\beta_2}\frac{\partial Q}{\partial\lambda_1} & & -\frac{\partial^2 Q}{\partial\beta_2\partial\alpha_2} - \frac{\partial Q}{\partial\beta_2}\frac{\partial Q}{\partial\alpha_2} & \cdots & -\frac{\partial^2 Q}{\partial\beta_2\partial\lambda_2} - \frac{\partial Q}{\partial\beta_2}\frac{\partial Q}{\partial\lambda_2} \\ -\frac{\partial^2 Q}{\partial\lambda_2\partial\alpha_1} - \frac{\partial Q}{\partial\lambda_2}\frac{\partial Q}{\partial\alpha_1} & \cdots & -\frac{\partial^2 Q}{\partial\lambda_2\partial\lambda_1} - \frac{\partial Q}{\partial\lambda_2}\frac{\partial Q}{\partial\lambda_1} & & -\frac{\partial^2 Q}{\partial\lambda_2\partial\alpha_2} - \frac{\partial Q}{\partial\lambda_2}\frac{\partial Q}{\partial\alpha_2} & \cdots & -\frac{\partial^2 Q}{\partial\lambda_2^2} - \left(\frac{\partial Q}{\partial\lambda_2}\right)^2 \end{array}\right]\Bigg|_{\Theta=\hat{\Theta}}$$

The terms used for the construction of observed Fisher information matrix are given below

$$\frac{\partial^2 Q}{\partial\alpha_1^2} = -\frac{(n_{r_1}+P)}{\alpha_1^2} - \beta_1\sum_{i=1}^{n_{r_1}} t_i^{\alpha_1}\Big(\ln(t_i)\Big)^2 - \beta_1\sum_{i=1}^{n_{r_2}} t_i^{\alpha_1}\Big(\ln(t_i)\Big)^2 - \beta_1\sum_{i=1}^{n_{nr}} P_i\xi_7\Big(t_i^*;\alpha_1,\beta_1\Big)$$
$$- \beta_1\sum_{i=1}^{n_{nr}}(1-P_i)\xi_{15}\Big(t_i^*;\alpha_1,\alpha_2,\beta_2\Big) - \beta_1\sum_{i=1}^{n_c} s_i^{\alpha_1}\Big(\ln(s_i)\Big)^2,$$

$$\frac{\partial^2 Q}{\partial\beta_1^2} = -\frac{\Big(n_{r_1}+P\Big)}{\beta_1^2}; \quad \frac{\partial^2 Q}{\partial\beta_2^2} = -\frac{(n_{r_2}+n_{nr}-P)}{\beta_2^2},$$

$$\frac{\partial^2 Q}{\partial\alpha_2^2} = -\frac{\Big(n_{r_2}+n_{nr}-P\Big)}{\alpha_2^2} - \beta_2\sum_{i=1}^{n_{r_1}} t_i^{\alpha_2}\Big(\ln(t_i)\Big)^2 - \beta_2\sum_{i=1}^{n_{r_2}} t_i^{\alpha_2}\Big(\ln(t_i)\Big)^2$$
$$- \beta_2\sum_{i=1}^{n_{nr}} P_i\xi_8\Big(t_i^*;\alpha_1,\beta_1,\alpha_2\Big) - \beta_2\sum_{i=1}^{n_{nr}}(1-P_i)\xi_{16}\Big(t_i^*;\alpha_2,\beta_2\Big) - \beta_2\sum_{i=1}^{n_c} s_i^{\alpha_2}\Big(\ln(s_i)\Big)^2,$$

$$\frac{\partial^2 Q}{\partial\lambda_1^2} = -\frac{P}{\lambda_1^2}; \quad \frac{\partial^2 Q}{\partial\lambda_2^2} = -\frac{\Big(n_{nr}-P\Big)}{\lambda_2^2},$$

$$\frac{\partial^2 Q}{\partial\beta_1\partial\alpha_1} = -\sum_{i=1}^{n_{r_1}} t_i^{\alpha_1}\ln(t_i) - \sum_{i=1}^{n_{r_2}} t_i^{\alpha_1}\ln(t_i) - \sum_{i=1}^{n_{nr}} P_i\xi_4\Big(t_i^*;\alpha_1,\beta_1\Big) - \sum_{i=1}^{n_c} s_i^{\alpha_1}\ln(s_i)$$
$$- \sum_{i=1}^{n_{nr}}\Big(1-P_i\Big)\xi_{12}\Big(t_i^*;\alpha_1,\alpha_2,\beta_2\Big) = \frac{\partial^2 Q}{\partial\alpha_1\partial\beta_1},$$

$$\frac{\partial^2 Q}{\partial\beta_2\partial\alpha_2} = -\sum_{i=1}^{n_{r_1}} t_i^{\alpha_2}\ln(t_i) - \sum_{i=1}^{n_{r_2}} t_i^{\alpha_2}\ln(t_i) - \sum_{i=1}^{n_{nr}} P_i\xi_5\Big(t_i^*;\alpha_1,\beta_1,\alpha_2\Big) - \sum_{i=1}^{n_{nr}}\Big(1-P_i\Big)\xi_{13}\Big(t_i^*;\alpha_2,\beta_2\Big)$$
$$- \sum_{i=1}^{n_c} s_i^{\alpha_2}\ln(s_i) = \frac{\partial^2 Q}{\partial\alpha_2\partial\beta_2}.$$

These terms are evaluated at estimated values of parameters.

4.8.3 A-3

All the inferences in the Bayesian paradigm are drawn from the posterior distribution. The likelihood in (4.24) indicates that it is difficult to study the properties of posterior analytically. So, we use likelihood given in (4.28) in this section for further study.

Prior Distribution

Consider α_k, β_k, λ_k; $(k = 1, 2)$ as Gamma distributed with parameters (a_l, b_l); $l = 1, 2, \ldots, 6$. Assuming the independence of priors the joint prior density is written up to proportionality constants as below

$$\pi(\Theta) \propto {\alpha_1}^{a_1-1}{\beta_1}^{a_2-1}{\lambda_1}^{a_3-1}{\alpha_2}^{a_4-1}{\beta_2}^{a_5-1}{\lambda_2}^{a_6-1} \exp\{-(\alpha_1 b_1 + \beta_1 b_2 + \lambda_1 b_3 + \alpha_2 b_4 + \beta_2 b_5 + \lambda_2 b_6)\} \tag{4.44}$$

Gibbs Sampling

For applying the Gibbs Sampling the posterior distribution is required which can be obtained by merging complete likelihood (4.28) and the joint prior (4.44). The posterior distribution up-to proportionality constant is written below

$$\begin{aligned}
\Pi(\Theta|d, d^*) &\propto L_c(\Theta|d, d^*)\pi(\Theta) \\
&\propto \alpha_1^{(n_{r_1}+Z^*+a_1-1)}\beta_1^{(n_{r_1}+Z^*+a_2-1)}\lambda_1^{(Z^*+a_3-1)}\alpha_2^{(n_{r_2}+n_{nr}-Z^*+a_4-1)}\beta_2^{(n_{r_2}+n_{nr}-Z^*+a_5-1)} \\
&\lambda_2^{(n_{nr}-Z^*+a_6-1)}\prod_{i=1}^{n_{r_1}} t_i^{\alpha_1-1}\exp\{-\beta_1 t_i^{\alpha_1} - \beta_2 t_i^{\alpha_2} - \lambda_1(s_i - t_i)\}\prod_{i=1}^{n_c}\exp\{-\beta_1 s_i^{\alpha_1} - \beta_2 s_i^{\alpha_2}\} \\
&\prod_{i=1}^{n_{r_2}} t_i^{\alpha_2-1}\exp\{-\beta_1 t_i^{\alpha_1} - \beta_2 t_i^{\alpha_2} - \lambda_2(s_i - t_i)\} \\
&\prod_{i=1}^{n_{nr}}\left[t_i^{*\alpha_1-1}\exp\{-\beta_1 t_i^{*\alpha_1} - \beta_2 t_i^{*\alpha_2} - \lambda_1 u_i^*\}\right]^{z_i} \\
&\prod_{i=1}^{n_{nr}}\left[t_i^{*\alpha_2-1}\exp\{-\beta_1 t_i^{*\alpha_1} - \beta_2 t_i^{*\alpha_2} - \lambda_2 v_i^*\}\right]^{1-z_i} \\
&\exp\{-\alpha_1 b_1 - \beta_1 b_2 - \lambda_1 b_3 - \alpha_2 b_4 - \beta_2 b_5 - \lambda_2 b_6\}
\end{aligned} \tag{4.45}$$

In order to draw the samples from posterior, we write the full conditionals for parameters based on the posterior distribution (4.45) which are given below

$$\alpha_1|\Theta_{(-\alpha_1)} \propto \alpha_1^{(n_{r_1}+Z^*+a_1-1)} \left(\prod_{i=1}^{n_{r_1}} t_i^{\alpha_1-1}\right) \left(\prod_{i=1}^{n_{nr}} t_i^{*z_i(\alpha_1-1)}\right) \exp\left(-\beta_1 \sum_{i=1}^{n_{r_1}} t_i^{\alpha_1} - \beta_1 \sum_{i=1}^{n_{r_2}} t_i^{\alpha_1} - \beta_1 \sum_{i=1}^{n_{nr}} z_i t_i^{*\alpha_1}\right)$$
$$\exp\left(-\beta_1 \sum_{i=1}^{n_{nr}} (1-z_i) t_i^{*\alpha_1} - \beta_1 \sum_{i=1}^{n_c} s_i^{\alpha_1} - \alpha_1 b_1\right), \tag{4.46}$$

$$\alpha_2|\Theta_{(-\alpha_2)} \propto \alpha_2^{(n_{r_2}+n_{nr}-Z^*+a_4-1)} \left(\prod_{i=1}^{n_{r_2}} t_i^{\alpha_2-1}\right) \left(\prod_{i=1}^{n_{nr}} t_i^{*(1-z_i)(\alpha_2-1)}\right) \exp\left(-\beta_2 \sum_{i=1}^{n_{r_1}} t_i^{\alpha_2} - \beta_2 \sum_{i=1}^{n_{r_2}} t_i^{\alpha_2}\right)$$
$$\exp\left(-\beta_2 \sum_{i=1}^{n_{nr}} z_i t_i^{*\alpha_2} - \beta_2 \sum_{i=1}^{n_{nr}} (1-z_i) t_i^{*\alpha_2} - \beta_2 \sum_{i=1}^{n_c} s_i^{\alpha_2} - \alpha_2 b_4\right), \tag{4.47}$$

$$\beta_1|\Theta_{(-\beta_1)} \sim \mathcal{G}\left(n_{r_1} + Z^* + a_2,\ \sum_{i=1}^{n_{r_1}} t_i^{\alpha_1} + \sum_{i=1}^{n_{r_2}} t_i^{\alpha_1} + \sum_{i=1}^{n_{nr}} z_i t_i^{*\alpha_1} + \sum_{i=1}^{n_{nr}} (1-z_i) t_i^{*\alpha_1} + \sum_{i=1}^{n_c} s_i^{\alpha_1} + b_2\right), \tag{4.48}$$

$$\lambda_1|\Theta_{(-\lambda_1)} \sim \mathcal{G}\left(Z^* + a_3,\ \sum_{i=1}^{n_{r_1}} (s_i - t_i) + \sum_{i=1}^{n_{nr}} z_i u_i^* + b_3\right), \tag{4.49}$$

$$\beta_2|\Theta_{(-\beta_2)} \sim \mathcal{G}\left(n_{r_2} + n_{nr} - Z^* + a_5,\ \sum_{i=1}^{n_{r_1}} t_i^{\alpha_2} + \sum_{i=1}^{n_{r_2}} t_i^{\alpha_2} + \sum_{i=1}^{n_{nr}} z_i t_i^{*\alpha_2} + \sum_{i=1}^{n_{nr}} (1-z_i) t_i^{*\alpha_2} + \sum_{i=1}^{n_c} s_i^{\alpha_2} + b_5\right), \tag{4.50}$$

$$\lambda_2|\Theta_{(-\lambda_2)} \sim \mathcal{G}\left(n_{nr} - Z^* + a_6,\ \sum_{i=1}^{n_{r_2}} (s_i - t_i) + \sum_{i=1}^{n_{nr}} (1-z_i) v_i^* + b_6\right), \tag{4.51}$$

From equations (4.46) to (4.51), we can see that the full conditionals of α_1 and α_2 are not in standard density form. So, samples from these full conditionals are generated by using Metropolis-Hasting (M-H) (see Metropolis and Ulam (1949) and Hastings (1970)) algorithm taking Normal distribution as proposal density. Further full conditionals of β_1, λ_1, β_2 and λ_2 are following Gamma distribution with varying shapes and scales and observations from these can be drawn directly for given values of latent variables z, t^*, u^* and v^* and hyper-parameters.

Data Augmentation Algorithm

The algorithm used for generating samples from full conditionals using three-stage Gibbs sampling is given below:

1. Step 1: Set initial values of parameters, say $\Theta^{(0)} = \left(\alpha_1^{(0)}, \beta_1^{(0)}, \lambda_1^{(0)}, \alpha_2^{(0)}, \beta_2^{(0)}, \lambda_2^{(0)}\right)$ and generate $z_i \sim \mathcal{B}(1, P_i);\ i = 1, 2, \cdots, n_{nr}$.

2. Step 2: For given values of z_i and Θ, generate observations on t_i^* using expression

$$t_i^* = \left[-\frac{1}{\beta_k}\ln\left\{1 - w_i\left(1 - \exp\{-\beta_k s_i^{\alpha_k}\}\right)\right\}\right]^{1/\alpha_k};\ i = 1, 2, \ldots, n_{nr},\ k = 1, 2, \quad (4.52)$$

where $w \sim \mathcal{U}(0, 1)$.

3. Step 3: Based on generated values of z_i, t_i^* and Θ from previous steps, observations on u_i^* and v_i^* can be generated by using expression

$$y_i^* = -\frac{1}{\lambda_k}\ln\left[1 - w_i\Big(1 - \exp\{-\lambda_k(s_i - t_i^*)\}\Big)\right];\ i = 1, 2, \ldots, n_{nr},\ k = 1, 2. \quad (4.53)$$

4. Step 4: For given values of z_i, t_i^*, u_i^* and v_i^* generate observations on $\alpha_1^{(1)}$ and $\alpha_2^{(1)}$ using M-H algorithm taking normal as proposal density.

5. Step 5: In the next step observations on $\beta_1^{(1)}$, $\lambda_1^{(1)}$, $\beta_2^{(1)}$ and $\lambda_2^{(1)}$ can be generated for given $\alpha_1^{(1)}$ and $\alpha_2^{(1)}$ using (4.48) to (4.51) respectively.

Now, the current state is $\Theta^{(1)} = \left(\alpha_1^{(1)}, \beta_1^{(1)}, \lambda_1^{(1)}, \alpha_2^{(1)}, \beta_2^{(1)}, \lambda_2^{(1)}\right)$. Steps 1 to 5 are replicated M times to obtain a sequence of random variables $\left(\Theta^{(1)}, \Theta^{(2)}, \cdots, \Theta^{(M)}\right)$. After discarding the burn-in from generated chains, the stationarity of the chains are checked by trace plot and Gelman and Rubin's test statistics (see Gelman and Rubin (1992)), we left out with a reduced chain of length M'. All the Bayesian inferences are drawn based on these samples.

Chapter 5

Modeling of Recall-Based Competing Risk Data using Exponential and Rayleigh Distributions

5.1 Introduction

In real life situations it is quite common that a system or individual may be exposed to more than one risks at any time point and the breakdown/death is due to any of these causes which appears first. The occurrence of time due to the event of interest is affected by the presence of other causes. In competing risks setup, we observe the time to event and associated causes. Thus, the modeling of competing risks differs from the usual time to event analysis. For details and basics about competing risk one may refer David and Moeschberger (1978) and Crowder (2001). In some cases, the time to event is observed but the cause of occurrence of the event is unknown, known as masked data in the literature. For more details about the masked data one may refer to Usher and Hodgson (1988), Kuo and Yang (2000), Sarhan (2003) and Tomer et al. (2014).

In most of the literature, it is seen that the causes responsible for death/failure of an individual/system are taken from the same family. But in real life situations, it is not necessary and causes responsible for event may follow different distributions. Friedman and Gertsbakh (1980) studies the properties and conditions for the existence of MLEs formulated a model based on minimum of exponential and Weibull distributions. The authors have shown when all parameters are unknown the MLE does not exist but when the shape parameter for

Weibull is known the log-likelihood is concave function and MLEs exist. Jiang and Murthy (1997) studied the competing risk model using minimum of two Weibull distributions under parametric setup. Bousquet et al. (2006) proposed a distribution based on minimum of the exponential and Weibull distributions for modeling the accidental failure and failure due to aging in a competing risks problem. Ranjan et al. (2015) proposed a competing risk model based on minimum of gamma and exponential distribution to model the aging and accidental failures. The proposed model is analyzed under Bayesian framework using proper but no-informative priors. The study is illustrated with extensive simulation procedure. Ranjan and Upadhyay (2016) analyzed a competing risk model based on minimum of exponential and gamma under classical and Bayesian approaches. Under classical inference the E-M algorithm is applied to find out the point estimates whereas samples from posterior density is drawn using importance sampling. The model is illustrated through simulation and vehicle failures dataset. In another study, Gupta et al. (2018) studied the competing risks problem considering an increasing and decreasing hazard Weibull distribution for two different causes. The proposed model is illustrated through a real dataset based on failure times of 58 electrodes which were put on a high-stress voltage endurance life test.

These studies motivated us to model recall-based data under competing risk scenario taking two different distributions for associated to causes. The objective of this study is to develop a mathematical model for recall-based current status data under competing risks setup when the causes associated to two risks are assumed to follow exponential and Rayleigh distributions respectively. Section 5.2 is dedicated to defining data structure and formation of likelihood. The non-recall probability is taken as a function of elapsed time between monitoring time and time to event as discussed in previous chapters. In Section 5.3, an algorithm based on the E-M is proposed for point estimation. The Fisher information matrix is also constructed using the missing information principle for interval estimation under both cases. In Section 5.3, we study the same setup under Bayesian paradigm by choosing conjugate priors for the two distributions. The samples from posterior densities are generated using Gibbs Sampling based algorithm. In Section 5.4 to check the performance of proposed estimators, an extensive simulation study is done for different proportions of non-

recall and censoring with varying sample sizes under uniform and exponential monitoring points. Finally, we concluded the chapter in Section 5.5.

5.2 Data Structure and Likelihood

In Chapters 3 and 4, we have considered the competing risks scenario for recall-based data. In Chapter 3, the causes responsible for the occurrence of events are assumed to follow exponential distribution while in Chapter 4 the causes responsible for the occurrence of events are assumed to follow Weibull distribution. We developed the methodologies for point and interval estimation under classical and Bayesian approaches. Here, we follow the same data setup under competing risk scenario for $K = 2$, i.e., there are only two causes responsible for occurrence of time to event. We assume latent time to events for two causes as exponential and Rayleigh, i.e., $T_1 \sim \mathcal{E}(\theta_1)$ and $T_2 \sim \mathcal{R}(\theta_2)$. This indicates that the resulting time to event is observed as a minimum of exponential and Rayleigh distributions. Also, the non-recall probability parameters are denoted by λ_1 and λ_2 corresponding to two causes respectively. The density and cdf for the latent time (t) associated to j^{th} cause under exponential distribution can be defined as below

$$f(t;\theta_j) = \theta_j \exp\{-\theta_j t\};\ \theta_j > 0,\ t > 0,$$
$$\text{and} \quad F(t;\theta_j) = 1 - \exp\{-\theta_j t\};\ \theta_j > 0,\ t > 0.$$

Similarly, the density and cdf for the latent time associated to l^{th} cause under Rayleigh distribution can be defined as below

$$f(t;\theta_l) = \frac{t}{\theta_l^2} \exp\left\{-\frac{t^2}{2\theta_l^2}\right\};\ \theta_l > 0,\ t > 0,$$
$$\text{and} \quad F(t;\theta_l) = 1 - \exp\left\{-\frac{t^2}{2\theta_l^2}\right\};\ \theta_l > 0,\ t > 0.$$

Let us denote the set of unknown parameters by $\Theta = (\theta_1, \theta_2, \lambda_1, \lambda_2)$. Under this assumption, we get the likelihood function as below

$$\begin{aligned} L(\Theta|d) &= \theta_1^{n_{r_1}} \theta_2^{-2n_{r_2}} \prod_{i=1}^{n_{r_1}} \exp\left\{-t_i\theta_1 - \frac{t_i^2}{2\theta_2^2}\right\} \exp\{-(s_i - t_i)\lambda_1\} \\ &\quad \prod_{i=1}^{n_{r_2}} t_i \exp\left\{-t_i\theta_1 - \frac{t_i^2}{2\theta_2^2}\right\} \exp\{-(s_i - t_i)\lambda_2\} \\ &\quad \prod_{i=1}^{n_{nr}} \left[\int_0^{s_i} \theta_1 \exp\left\{-u\theta_1 - \frac{u^2}{2\theta_2^2}\right\}\Big(1 - \exp\{-(s_i - u)\lambda_1\}\Big)du \right. \\ &\quad \left. + \int_0^{s_i} \frac{u}{\theta_2^2} \exp\left\{-u\theta_1 - \frac{u^2}{2\theta_2^2}\right\}\Big(1 - \exp\{-(s_i - u)\lambda_2\}\Big)du\right] \\ &\quad \prod_{i=1}^{n_c} \exp\left\{-s_i\theta_1 - \frac{s_i^2}{2\theta_2^2}\right\}. \end{aligned} \tag{5.1}$$

The equation given in (5.1) is not in a convenient form and it is not easy to draw further inferences with it. Due to the presence of partial information on time to event and unknown causes of occurrence of events for non-recall observations, we treat it as missing data problem and apply E-M algorithm for point estimation.

5.3 Classical Estimation

5.3.1 Estimation using Expectation-Maximization

The E-M was proposed by (Dempster et al., 1977) which is applicable in case of missing data such as censoring, truncation, mixture models and clustering. Here, we apply the E-M algorithm to deal with time to event and causes of occurrence of the events. Keeping the complexity of likelihood (5.1), first we introduce a latent variable z_i for the non-recall observations having causes of occurrence of the event as not recall. Thus, for i^{th} observation under non-recall with $k = \{1, 2\}$, denote $z_i|(i = 1, 2, \ldots, n_{nr})$, a Bernoulli variable with probability of success $(z_i = 1)$. The probability of success is defined as

$$P_i = \frac{I_1(s_i)}{I_1(s_i) + I_2(s_i)}, \tag{5.2}$$

where, the functional form of $I_1(.)$ and $I_2(.)$, is given by:

$$I_1(s_i) = \int_0^{s_i} \theta_1 \exp\left\{-u\theta_1 - \frac{u^2}{2\theta_2^2}\right\}\Big(1 - \exp\{-(s_i - u)\lambda_1\}\Big)du,$$

$$I_2(s_i) = \int_0^{s_i} \frac{u}{\theta_2^2} \exp\left\{-u\theta_1 - \frac{u^2}{2\theta_2^2}\right\}\Big(1 - \exp\{-(s_i - u)\lambda_2\}\Big)du.$$

Once, we introduce z_i, we are able to assign causes for non-recall observations. Along with z_i, an equivalent quantity, say, T_i^*, is introduced for non-recall observations to deal with the integral. The density of t_i^* can be calculated using the concept of truncation. Thus, for given causes 1 and 2 we can calculate the conditional density of t_i^* given information that it lies in the interval $(0, s_i)$.

Further, it can also be observed that non-recall probabilities are not in a convenient form. So, to make the functional form of non-recall probabilities into convenient form, two latent variables u_i^* and v_i^* are introduced corresponding to causes 1 and 2 respectively. The idea is taken from (Kuo and Yang, 2000), who had described the same approach in the case of masked data under competing risks in survival analysis to find the Bayes estimates. We can generate u_i^* independently from an exponential distribution random variable with mean $\frac{1}{\lambda_1}$ truncated at $(s_i - t_i^*)$ for non-recall observation with $z_i = 1$ and similarly v_i^* can be also generated independently from an exponential distribution random variable with mean $\frac{1}{\lambda_2}$ truncated at $(s_i - t_i^*)$ for non-recall observation with $z_i = 0$.

After introducing the latent variables z_i, u_i^* and v_i^*, we set vector for missing data as $\underline{D}^* = (Z, T^*, U^*, V^*)$. Here, the symbol $d_i^* = (z_i, t_i^*, u_i^*, v_i^*)$ denotes the realization of vector $\underline{D}^*$ corresponding to i^{th} individual belonging to non-recall category. Combining the observed and missing data, the complete data can be denoted by (d, d^*). Now, the complete likelihood can be written as

$$L_c(\Theta|d, d^*) \propto \theta_1^{(n_{r_1}+Z^*)}\theta_2^{-2(n_{r_2}+n_{nr}-Z^*)}\lambda_1^{Z^*}\lambda_2^{(n_{nr}-Z^*)}\prod_{i=1}^{n_{r_1}}\exp\left\{-t_i\theta_1 - \frac{t_i^2}{2\theta_2^2}\right\}$$
$$\prod_{i=1}^{n_{r_1}}\exp\{-(s_i - t_i)\lambda_1\}\prod_{i=1}^{n_{r_2}}\exp\left\{-t_i\theta_1 - \frac{t_i^2}{2\theta_2^2}\right\}\exp\{-(s_i - t_i)\lambda_2\}$$

$$\prod_{i=1}^{n_c}\exp\Big\{-s_i\theta_1-\frac{s_i^2}{2\theta_2^2}\Big\}\prod_{i=1}^{n_{nr}}\Big[\exp\Big\{-t_i^*\theta_1-\frac{t_i^{*2}}{2\theta_2^2}\Big\}\exp\{-u_i^*\lambda_1\}\Big]^{z_i}$$
$$\prod_{i=1}^{n_{nr}}\Big[\exp\Big\{-t_i^*\theta_1-\frac{t_i^{*2}}{2\theta_2^2}\Big\}\exp\{-v_i^*\lambda_2\}\Big]^{1-z_i}, \tag{5.3}$$

After taking the natural logarithm of complete likelihood, we get the simplified log-likelihood as

$$\begin{aligned}
l_c(\Theta|d,d^*) \propto\ & (n_{r_1}+Z^*)\ln(\theta_1)-2(n_{r_2}+n_{nr}-Z^*)\ln(\theta_2)+Z^*\ln(\lambda_1)\\
&+(n_{nr}-Z^*)\ln(\lambda_2)-\theta_1\sum_{i=1}^{n_{r_1}}t_i-\frac{1}{2\theta_2^2}\sum_{i=1}^{n_{r_1}}t_i^2-\lambda_1\sum_{i=1}^{n_{r_1}}(s_i-t_i)-\theta_1\sum_{i=1}^{n_{r_2}}t_i\\
&-\frac{1}{2\theta_2^2}\sum_{i=1}^{n_{r_2}}t_i^2-\lambda_2\sum_{i=1}^{n_{r_2}}(s_i-t_i)-\theta_1\sum_{i=1}^{n_{nr}}z_it_i^*-\frac{1}{2\theta_2^2}\sum_{i=1}^{n_{nr}}z_it_i^{*2}-\lambda_1\sum_{i=1}^{n_{nr}}z_iu_i^*\\
&-\theta_1\sum_{i=1}^{n_{nr}}(1-z_i)t_i^*-\frac{1}{2\theta_2^2}\sum_{i=1}^{n_{nr}}(1-z_i)t_i^{*2}-\lambda_2\sum_{i=1}^{n_{nr}}(1-z_i)v_i^*-\theta_1\sum_{i=1}^{n_c}s_i\\
&-\frac{1}{2\theta_2^2}\sum_{i=1}^{n_c}s_i^2.
\end{aligned} \tag{5.4}$$

Let us denote the estimate of unknown parameter vectors by $\hat{\Theta}^{(m)}$ at m^{th} iteration. Then in order to apply the E step of E-M algorithm let us define the quantity

$$Q\Big(\Theta|\hat{\Theta}^{(m)}\Big)=E\Big[l_c(\Theta\big|d,d^*)\big|d,\hat{\Theta}^{(m)}\Big].$$

The expectation is taken with respect to missing data. Now, the quantity $Q(.)$ can be written as below

$$\begin{aligned}
Q(\Theta|\hat{\Theta}^{(m)}) \propto\ & (n_{r_1}+P)\ln(\theta_1)-2(n_{r_2}+n_{nr}-P)\ln(\theta_2)+P\ln(\lambda_1)+(n_{nr}-P)\ln(\lambda_2)\\
&-\theta_1\sum_{i=1}^{n_{r_1}}t_i-\frac{1}{2\theta_2^2}\sum_{i=1}^{n_{r_1}}t_i^2-\lambda_1\sum_{i=1}^{n_{r_1}}(s_i-t_i)-\theta_1\sum_{i=1}^{n_{r_2}}t_i-\frac{1}{2\theta_2^2}\sum_{i=1}^{n_{r_2}}t_i^2-\lambda_2\sum_{i=1}^{n_{r_2}}(s_i-t_i)\\
&-\theta_1\sum_{i=1}^{n_{nr}}P_i\xi_1(t_i^*;\theta^{(m)})-\frac{1}{2\theta_2^2}\sum_{i=1}^{n_{nr}}P_i\xi_2(t_i^*;\theta^{(m)})-\lambda_1\sum_{i=1}^{n_{nr}}P_i\xi_3(u_i^*;\lambda^{(m)})\\
&-\theta_1\sum_{i=1}^{n_{nr}}(1-P_i)\xi_4(t_i^*;\theta^{(m)})-\frac{1}{2\theta_2^2}\sum_{i=1}^{n_{nr}}(1-P_i)\xi_5(t_i^*;\theta^{(m)})-\theta_1\sum_{i=1}^{n_c}s_i\\
&-\lambda_2\sum_{i=1}^{n_{nr}}(1-P_i)\xi_6(v_i^*;\lambda^{(m)})-\frac{1}{2\theta_2^2}\sum_{i=1}^{n_c}s_i^2.
\end{aligned} \tag{5.5}$$

Since, z_i is a Bernoulli variate, then, we have $E(z_i) = P_i$ and $E(Z^*) = E(\sum_{i=1}^{n_{nr}} z_i) = P$. Now in M step, we can find the parameter estimates by maximizing the function $Q(\Theta|\hat{\Theta}^{(m)})$. The expected values of truncated time points used at M step in E-M algorithm can be obtained by terms below:

$$\xi_1(t_i^*;\theta_1) = E\Big[t_i^*|t_i^* < s_i, k=1, \theta_1\Big] = \frac{\Big[1-\exp\{-\theta_1 s_i\}(\theta_1 s_i+1)\Big]}{\theta_1\Big[1-\exp\{-\theta_1 s_i\}\Big]},$$

$$\xi_2(t_i^*;\theta_1) = E\Big[t_i^{*2}|t_i^* < s_i, k=1, \theta_1\Big] = \frac{\Big[2+\exp\{-\theta_1 s_i\}\{-\theta_1 s_i(\theta_1 s_i+2)-2\}\Big]}{\theta_1^2\Big[1-\exp\{-\theta_1 s_i\}\Big]},$$

$$\xi_3(u_i^*;\lambda_1) = E\Big[u_i^*|u_i^* < (s_i - t_i^*), k=1, \lambda_1\Big] = \frac{\Big[1-\exp\{-\lambda_1(s_i-t_i^*)\}\{\lambda_1(s_i-t_i^*)+1\}\Big]}{\lambda_1\Big[1-\exp\{-\lambda_1(s_i-t_i^*)\}\Big]},$$

$$\xi_4(t_i^*;\theta_2) = E\Big[t_i^*|t_i^* < s_i, k=2, \theta_2\Big] = \frac{\sqrt{2}\theta_2\gamma\left(\frac{3}{2}, \frac{s_i^2}{2\theta_2^2}\right)}{1-\exp\{-\frac{s_i^2}{2\theta_2^2}\}},$$

$$\xi_5(t_i^*;\theta_2) = E\Big[t_i^{*2}|t_i^* < s_i, k=2, \theta_2\Big] = \frac{\Big[2\theta_2^2-\exp\{-\frac{s_i^2}{2\theta_2^2}\}(2\theta_2^2+s_i^2)\Big]}{1-\exp\{-\frac{s_i^2}{2\theta_2^2}\}},$$

$$\xi_6(v_i^*;\lambda_2) = E\Big[v_i^*|v_i^* < (s_i - t_i^*), k=2, \lambda_2\Big] = \frac{\Big[1-\exp\{-\lambda_2(s_i-t_i^*)\}\{\lambda_2(s_i-t_i^*)+1\}\Big]}{\lambda_2\Big[1-\exp\{-\lambda_2(s_i-t_i^*)\}\Big]},$$

where $\gamma(a,z) = \int_0^z t^{a-1}\exp(-t)dt$ denotes the lower incomplete gamma function. The estimators at $(m+1)^{th}$ step are given below

$$\hat{\theta}_1^{(m+1)} = \frac{(n_{r_1} + P^{(m)})}{\Big[\sum_{i=1}^{n_{r_1}} t_i + \sum_{i=1}^{n_{r_2}} t_i + \sum_{i=1}^{n_{nr}} P_i^{(m)}\xi_1\Big(t_i^*;\theta_1^{(m)}\Big) + \sum_{i=1}^{n_{nr}}\Big(1-P_i^{(m)}\Big)\xi_4\Big(t_i^*;\theta_2^{(m)}\Big) + \sum_{i=1}^{n_c} s_i\Big]}, \tag{5.6}$$

$$\hat{\theta}_2^{(m+1)} = \frac{\Big\{2\Big(n_{r_2} + n_{nr} - P^{(m)}\Big)\Big\}^{-1/2}}{\left[\sum_{i=1}^{n_{r_1}} t_i^2 + \sum_{i=1}^{n_{r_2}} t_i^2 + \sum_{i=1}^{n_{nr}} P_i^{(m)}\xi_2\Big(t_i^*;\theta_1^{(m)}\Big) + \sum_{i=1}^{n_{nr}}\Big(1-P_i^{(m)}\Big)\xi_5\Big(t_i^*;\theta_2^{(m)}\Big) + \sum_{i=1}^{n_c} s_i^2\right]^{-1/2}}. \tag{5.7}$$

$$\hat{\lambda}_1^{(m+1)} = \frac{P^{(m)}}{\Big[\sum_{i=1}^{n_{r_1}}(s_i - t_i) + \sum_{i=1}^{n_{nr}} P_i^{(m)}\xi_3\Big(u_i^*;\lambda_1^{(m)}\Big)\Big]}, \tag{5.8}$$

$$\hat{\lambda}_2^{(m+1)} = \frac{n_{nr} - P^{(m)}}{\Big[\sum_{i=1}^{n_{r_2}}(s_i - t_i) + \sum_{i=1}^{n_{nr}}(1-P_i^{(m)})\xi_6\Big(v_i^*;\lambda_2^{(m)}\Big)\Big]}, \tag{5.9}$$

The estimates of parameters can be computed by an iterative procedure using (5.6) to (5.9). The iterations can be terminated when $|\hat{\theta}_1^{(m+1)} - \hat{\theta}_1^{(m)}| + |\hat{\theta}_2^{(m+1)} - \hat{\theta}_2^{(m)}| + |\hat{\lambda}_1^{(m+1)} - \hat{\lambda}_1^{(m)}| + |\hat{\lambda}_2^{(m+1)} - \hat{\lambda}_2^{(m)}| < \epsilon$, where $\epsilon > 0$ is a sufficiently small real number.

5.3.2 Observed Fisher Information Matrix

For obtaining the Fisher information matrix, we have used the missing information principle proposed by Louis (1982). The structure of the observed Fisher information matrix can be written as

$$I(\hat{\Theta}) = \begin{bmatrix} -\frac{\partial^2 Q}{\partial\theta_1^2} - \left(\frac{\partial Q}{\partial\theta_1}\right)^2 & -\frac{\partial^2 Q}{\partial\theta_1\partial\theta_2} - \frac{\partial Q}{\partial\theta_1}\frac{\partial Q}{\partial\theta_2} & -\frac{\partial^2 Q}{\partial\theta_1\partial\lambda_1} - \frac{\partial Q}{\partial\theta_1}\frac{\partial Q}{\partial\lambda_1} & -\frac{\partial^2 Q}{\partial\theta_1\partial\lambda_2} - \frac{\partial Q}{\partial\theta_1}\frac{\partial Q}{\partial\lambda_2} \\ -\frac{\partial^2 Q}{\partial\theta_2\partial\theta_1} - \frac{\partial Q}{\partial\theta_2}\frac{\partial Q}{\partial\theta_1} & -\frac{\partial^2 Q}{\partial\theta_2^2} - \left(\frac{\partial Q}{\partial\theta_2}\right)^2 & -\frac{\partial^2 Q}{\partial\theta_2\partial\lambda_1} - \frac{\partial Q}{\partial\theta_2}\frac{\partial Q}{\partial\lambda_1} & -\frac{\partial^2 Q}{\partial\theta_2\partial\lambda_2} - \frac{\partial Q}{\partial\theta_2}\frac{\partial Q}{\partial\lambda_2} \\ -\frac{\partial^2 Q}{\partial\lambda_1\partial\theta_1} - \frac{\partial Q}{\partial\lambda_1}\frac{\partial Q}{\partial\theta_1} & -\frac{\partial^2 Q}{\partial\lambda_1\partial\theta_2} - \frac{\partial Q}{\partial\lambda_1}\frac{\partial Q}{\partial\theta_2} & -\frac{\partial^2 Q}{\partial\lambda_1^2} - \left(\frac{\partial Q}{\partial\lambda_1}\right)^2 & -\frac{\partial^2 Q}{\partial\lambda_1\partial\lambda_2} - \frac{\partial Q}{\partial\lambda_1}\frac{\partial Q}{\partial\lambda_2} \\ -\frac{\partial^2 Q}{\partial\lambda_2\partial\theta_1} - \frac{\partial Q}{\partial\lambda_2}\frac{\partial Q}{\partial\theta_1} & -\frac{\partial^2 Q}{\partial\lambda_2\partial\theta_2} - \frac{\partial Q}{\partial\lambda_2}\frac{\partial Q}{\partial\theta_2} & -\frac{\partial^2 Q}{\partial\lambda_2\partial\lambda_1} - \frac{\partial Q}{\partial\lambda_2}\frac{\partial Q}{\partial\lambda_1} & -\frac{\partial^2 Q}{\partial\lambda_2^2} - \left(\frac{\partial Q}{\partial\lambda_2}\right)^2 \end{bmatrix}\Bigg|_{\Theta=\hat{\Theta}}$$

Expressions which are used in construction of observed Fisher information matrix are given below

$$\frac{\partial^2 Q}{\partial\theta_1^2} = -\left(\frac{n_{r_1} + P}{\hat{\theta}_1^2}\right), \quad \frac{\partial^2 Q}{\partial\lambda_1^2} = -\frac{P}{\hat{\lambda}_1^2}, \quad \frac{\partial^2 Q}{\partial\lambda_2^2} = -\frac{(n_{nr} - P)}{\hat{\lambda}_2^2},$$

$$\frac{\partial^2 Q}{\partial\theta_2^2} = \frac{2(n_{r_2} + n_{nr} - P)}{\hat{\theta}_2^2} - \frac{3}{\hat{\theta}_2^4}\left\{\sum_{i=1}^{n_{r_1}} t_i^2 + \sum_{i=1}^{n_{r_2}} t_i^2 + \sum_{i=1}^{n_{nr}} P_i\xi_2(t_i^*;\theta_1) + \sum_{i=1}^{n_{nr}}(1-P_i)\xi_5(t_i^*;\theta_2)\right\}.$$

Rest of the derivatives comes out to be zero. Now, the variance-covariance matrix is obtained by inverting $I(\hat{\Theta})$ and standard error (SE) for parameters can be calculated by taking square root of diagonal elements of variance covariance matrix. Using asymptotic normality assumption of MLE, $100(1-\gamma)\%$ confidence limits for $\hat{\Theta}$ is given by $\hat{\Theta} \pm z_{\frac{\gamma}{2}} SE(\hat{\Theta})$, where $z_{(\frac{\gamma}{2})}$ is upper $100(\frac{\gamma}{2})^{th}$ percentile of standard normal variate.

5.4 Bayesian Inference

Classical estimation methods assume the parameters are fixed but known. In the Bayesian framework, it is assumed that unknown parameters are random and have certain distributions. The information on parameters is known as prior distribution and utilized to draw inferences under the Bayesian approach.

5.4.1 Choice of Prior Distribution

An important feature of the Bayesian analysis is the choice of the prior distribution. If the data have sufficient signal, even a bad prior will still not greatly influence the posterior. Usually, the prior information is subjective and is based on a person's own experience and judgment, a statement of one's degree of belief regarding the parameter. We can examine the impact of prior by observing the stability of posterior distribution related to different choices of priors. If the posterior distribution is highly dependent on the prior, then the data (the likelihood function) may not contain sufficient information. However, if the posterior is relatively stable over a choice of priors, then the data indeed contains significant information.

As we know, a conjugate prior for the exponential is gamma and for Rayleigh distribution a square root inverse gamma ($\mathcal{SQIG}$) as a conjugate prior is considered by Dey and Dey (2012). For study purposes, we have considered conjugate priors for all parameters, i.e., $\theta_1 \sim \mathcal{G}(a_1, b_1)$, $\theta_2 \sim \mathcal{SQIG}(a_2, b_2)$, $\lambda_1 \sim \mathcal{G}(a_3, b_3)$ and $\lambda_2 \sim \mathcal{G}(a_4, b_4)$ respectively. The density functions for different priors up-to proportionality constants are written as

$$\pi_1(\theta_1) \propto \theta_1^{a_1-1} exp\{-\theta_1 b_1\};\ a_1, b_1 > 0, \tag{5.10}$$

$$\pi_2(\theta_2) \propto \theta_2^{-(2b_2+1)} exp\left\{-\frac{a_2}{2\theta_2^2}\right\};\ a_2, b_2 > 0, \tag{5.11}$$

$$\pi_3(\lambda_1) \propto \lambda_1^{a_3-1} exp\{-\lambda_1 b_3\};\ a_3, b_3 > 0 \tag{5.12}$$

$$\pi_4(\lambda_2) \propto \lambda_2^{a_4-1} exp\{-\lambda_2 b_4\};\ a_4, b_4 > 0. \tag{5.13}$$

Considering $\theta_1, \theta_2, \lambda_1$ and λ_2 as independent and values of hyper-parameters to be known, the joint prior density for Θ is given by

$$\pi(\Theta) \propto \theta_1^{a_1-1}\theta_2^{-(2b_2+1)}\lambda_1^{a_3-1}\lambda_2^{a_4-1} exp\left\{-\theta_1 b_1 - \frac{a_2}{2\theta_2^2} - \lambda_1 b_3 - \lambda_2 b_4\right\};\ a_i, b_i > 0,\ i = 1, 2, 3, 4 \tag{5.14}$$

The values of hyper-parameters can be chosen by using moment matching criteria. This is done by matching the mean and variance of MLEs with the mean and variances of considered prior distribution. The Bayes estimates are obtained under the squared error loss function (SELF) which gives equal weight to over and under estimation.

5.4.2 Posterior Distribution

Using the complete likelihood given in (5.3), joint prior given in (5.14) and Bayes theorem, the posterior distribution is written as

$$\begin{aligned}\Pi(\Theta|d,d^*) \propto\ & \theta_1^{(n_{r_1}+Z^*+a_1-1)}\theta_2^{-\{2(n_{r_2}+n_{nr}-Z^*+b_2)+1\}}\lambda_1^{(Z^*+a_3-1)}\lambda_2^{(n_{nr}-Z^*+a_4-1)} \\ & \prod_{i=1}^{n_{r_1}}\exp\left\{-t_i\theta_1-\frac{t_i^2}{2\theta_2^2}-(s_i-t_i)\lambda_1\right\}\prod_{i=1}^{n_{r_2}}\exp\left\{-t_i\theta_1-\frac{t_i^2}{2\theta_2^2}-(s_i-t_i)\lambda_2\right\} \\ & \prod_{i=1}^{n_{nr}}\left[\exp\left\{-t_i^*\theta_1-\frac{t_i^{*2}}{2\theta_2^2}-u_i^*\lambda_1\right\}\right]^{z_i}\left[\exp\left\{-t_i^*\theta_1-\frac{t_i^{*2}}{2\theta_2^2}-v_i^*\lambda_2\right\}\right]^{1-z_i} \\ & \prod_{i=1}^{n_c}\left[\exp\left\{-s_i\theta_1-\frac{s_i^2}{2\theta_2^2}\right\}\right] exp\left\{-\theta_1 b_1-\frac{a_2}{2\theta_2^2}-\lambda_1 b_3-\lambda_2 b_4\right\} \end{aligned} \tag{5.15}$$

The full conditionals for θ_1, θ_2, λ_1 and λ_2 are obtained form posterior distribution (5.15) as given below

$$\theta_1|\Theta_{(-\theta_1)} \sim \mathcal{G}\left(n_{r_1}+Z^*+a_1,\ \sum_{i=1}^{n_{r_1}}t_i+\sum_{i=1}^{n_{r_2}}t_i+\sum_{i=1}^{n_{nr}}z_i t_i^*+\sum_{i=1}^{n_{nr}}(1-z_i)t_i^*+\sum_{i=1}^{n_c}s_i+b_1\right), \tag{5.16}$$

$$\theta_2|\Theta_{(-\theta_2)} \sim \mathcal{SQJG}\left(n_{r_2}+n_{nr}-Z^*+b_2,\ \sum_{i=1}^{n_{r_1}}t_i^2+\sum_{i=1}^{n_{r_2}}t_i^2+\sum_{i=1}^{n_{nr}}z_i t_i^{*2}+\sum_{i=1}^{n_{nr}}(1-z_i)t_i^{*2}+\sum_{i=1}^{n_c}s_i^2+a_2\right), \tag{5.17}$$

$$\lambda_1|\Theta_{(-\lambda_1)} \sim \mathcal{G}\left(Z^*+a_3,\ \sum_{i=1}^{n_{r_1}}(s_i-t_i)+\sum_{i=1}^{n_{nr}}z_i u_i^*+b_3\right) \tag{5.18}$$

$$\lambda_2|\Theta_{(-\lambda_2)} \sim \mathcal{G}\left(n_{nr}-Z^*+a_4,\ \sum_{i=1}^{n_{r_2}}(s_i-t_i)+\sum_{i=1}^{n_{nr}}(1-z_i)v_i^*+b_4\right), \tag{5.19}$$

5.4.3 Gibbs Sampling Based Algorithm

From the equations 5.16 to 5.19, we can see that full conditionals are in closed form. For generating the samples from these full conditionals, we use Gibbs Sampling algorithm discussed earlier in the Chapter 3 and Chapter 4. The steps for the sample generation can be summarized as below

1. We set the initial values of unknown parameters $\Theta^{(0)} = \left(\theta_1^{(0)}, \theta_2^{(0)}, \lambda_1^{(0)}, \lambda_2^{(0)}\right)$. Then,

for i^{th} observation under non-recall category, we generate $z_i \sim B(1, P_i)$ and compute $Z^* = \sum_{i=1}^{n_{nr}} z_i$.

2. For given values of $z_i = 1$, the samples from t_i^* is generated using the conditional densities of exponential using the expression

$$t_i^* = -\frac{1}{\theta_1} \ln \Big[1 - w_i \Big(1 - \exp\{-\theta_1 s_i\}\Big)\Big],\ i = 1, 2, \cdots, n_{nr}.$$

3. For given values of $z_i = 0$, the samples from t_i^* is generated using the conditional densities of Rayleigh using the expression

$$t_i^* = \left[-2\theta_2^2 \ln \left\{ 1 - w_i \Big[1 - \exp\left\{ -\frac{s_i^2}{2\theta_2^2} \right\} \Big] \right\} \right]^{1/2},\ i = 1, 2, \cdots, n_{nr}.$$

4. For given values of $\Theta^{(0)}$, $z_i = 1$ and t_i^*, observations on u_i^* can be generated using the expression

$$u_i^* = -\frac{1}{\lambda_1} \ln \Big[1 - w_i \Big(1 - \exp\{-\lambda_1 (s_i - t_i^*)\}\Big)\Big],\ i = 1, 2, \cdots, n_{nr}.$$

5. For given values of $\Theta^{(0)}$, $z_i = 0$ and t_i^*, observations on v_i^* can be generated using the expression

$$v_i^* = -\frac{1}{\lambda_2} \ln \Big[1 - w_i \Big(1 - \exp\{-\lambda_2 (s_i - t_i^*)\}\Big)\Big],\ i = 1, 2, \cdots, n_{nr}.$$

Here, $w \sim \mathcal{U}(0, 1)$. For given values of z_i, t_i^*, u_i^* and v_i^* the parameters θ_1, θ_2, λ_1 and λ_2 are conditionally independent from each other and updation can be done simultaneously using the expressions 5.16 to 5.19. The current state is $\Theta^{(1)}$. Now steps 1 to 5 are replicated M times to obtain a sequence of Markov Chain Monte Carlo (MCMC) chain of random variables $\left\{\Theta^{(1)}, \Theta^{(2)}, \cdots, \Theta^{(M)}\right\}$. Once, the samples are generated, after discarding initial burn-in sample and making chain stationary (based on cumsum and ACF plots) all posterior inferences can be drawn. Point estimates and Bayesian Credible Intervals (BCIs) are

reported based on these samples. Also, HPD intervals which give the shortest lengths are obtained by the method of (Chen and Shao, 1999) for these posterior samples.

5.5 Simulation Study

In this section, a simulation study is done to check the performance of the proposed estimators. We generate $T_1 \sim \mathcal{E}(\theta_1)$ and $T_2 \sim \mathcal{R}(\theta_2)$ for arbitrary values of $\theta_1 = 0.95$ and $\theta_2 = 0.87$. These T_1 and T_2 are considered as the time to events of component 1 and component 2 respectively. Under competing risk setup, the observed time to event becomes $T = min(T_1, T_2)$. The cause of the event of interest is observed associated with minimum time to event. Here, we consider two different patterns uniform and exponential for generation of monitoring time points. In first case, S is generated from uniform distribution with parameters min(T) and max(T). We set the non-recall probability parameters $(\lambda_1, \lambda_2) = \{(0.21, 0.16), (0.35, 0.27)\}$ in such a way that we can maintain a proportion of non-recall at around 16% and 25% while right-censored proportion at around 14% and 25% as per the sample of size n. In second case, we generate S using exponential distribution with rate parameter 0.3. We set the non-recall probability parameters $(\lambda_1, \lambda_2) = \{(0.10, 0.11), (0.12, 0.14)\}$ in such a way that we can maintain proportion of non-recall at 21% and 25% while right-censored observations at 15% and 16% receptively for the sample size n.

For this setup, the point estimates of the parameters of the model are obtained by using the E-M algorithm along with MSE and AB. For interval estimation, the variance-covariance matrix is calculated using the missing information principle. In Bayesian analysis, we set the values of hyper-parameters by the moment matching method. For these values of hyper-parameters, we obtain the Bayes estimates of $\theta_1, \theta_2, \lambda_1$, and λ_2 using the Gibbs Sampling method and give values of MSE and AB. We report ALs and CP for interval estimation from BCI and HPD based on posterior samples. The coverage probabilities represent the proportion that the interval contains the true value of interest. This process is simulated $N = 1000$ times to provide simulation tables for classical and Bayes estimates. The simulated results are reported in Tables 5.1 to 5.4.

Table 5.1: Mean square error and absolute bias for set of parameters $(\theta_1, \theta_2) = (0.95, 0.87)$ for varying sample size n for uniform monitoring pattern.

			ML			*Bayes*		
(λ_1, λ_2)		$n \to$	50	150	250	50	150	250
(0.21, 0.16)	$\hat{\theta}_1$	*MSE*	0.0174	0.0075	0.0068	0.0120	0.0036	0.0044
		AB	0.1022	0.0718	0.0665	0.1005	0.0688	0.0573
	$\hat{\theta}_2$	*MSE*	0.0221	0.0093	0.0085	0.0110	0.0043	0.0045
		AB	0.1133	0.0770	0.0689	0.1009	0.0651	0.0561
	$\hat{\lambda}_1$	*MSE*	0.0153	0.0022	0.0014	0.0111	0.0011	0.0010
		AB	0.0994	0.0371	0.0292	0.0720	0.0229	0.0117
	$\hat{\lambda}_2$	*MSE*	0.0304	0.0018	0.0010	0.0285	0.0011	0.0007
		AB	0.1092	0.0343	0.0255	0.0850	0.0251	0.0186
(0.35, 0.27)	$\hat{\theta}_1$	*MSE*	0.0188	0.0086	0.0079	0.0126	0.0039	0.0062
		AB	0.1034	0.0730	0.0689	0.1013	0.0695	0.0668
	$\hat{\theta}_2$	*MSE*	0.0228	0.0127	0.0115	0.0198	0.0116	0.0107
		AB	0.1139	0.0949	0.0827	0.1020	0.0878	0.0680
	$\hat{\lambda}_1$	*MSE*	0.0164	0.0042	0.0026	0.0139	0.0038	0.0018
		AB	0.0840	0.0510	0.0397	0.0760	0.0494	0.0265
	$\hat{\lambda}_2$	*MSE*	0.0379	0.0046	0.0022	0.0281	0.0037	0.0017
		AB	0.1174	0.0509	0.0376	0.1016	0.0483	0.0287

Table 5.2: Mean square error and absolute bias for set of parameters $(\theta_1, \theta_2) = (0.95, 0.87)$ for varying sample size n under exponential monitoring patterns.

			ML			*Bayes*		
(λ_1, λ_2)	$n \rightarrow$		50	150	250	50	150	250
$(0.10, 0.11)$	$\hat{\theta}_1$	*MSE*	0.0083	0.0042	0.0038	0.0058	0.0034	0.0032
		AB	0.0783	0.0506	0.0478	0.0685	0.0500	0.0404
	$\hat{\theta}_2$	*MSE*	0.1747	0.1428	0.1375	0.1574	0.1325	0.1322
		AB	0.0659	0.0589	0.0566	0.0642	0.0578	0.0555
	$\hat{\lambda}_1$	*MSE*	0.0008	0.0007	0.0006	0.0005	0.0003	0.0002
		AB	0.0236	0.0222	0.0214	0.0230	0.0220	0.0210
	$\hat{\lambda}_2$	*MSE*	0.0049	0.0048	0.0046	0.0042	0.0038	0.0030
		AB	0.0682	0.0680	0.0658	0.0679	0.0600	0.0579
$(0.12, 0.14)$	$\hat{\theta}_1$	*MSE*	0.0085	0.0045	0.0042	0.0068	0.0043	0.0042
		AB	0.0786	0.0546	0.0537	0.0689	0.0700	0.0434
	$\hat{\theta}_2$	*MSE*	0.1848	0.1494	0.1462	0.1625	0.1552	0.1422
		AB	0.3850	0.3798	0.3788	0.0649	0.0588	0.0566
	$\hat{\lambda}_1$	*MSE*	0.0022	0.0018	0.0014	0.0008	0.0006	0.0005
		AB	0.0390	0.0369	0.0332	0.0330	0.0226	0.0231
	$\hat{\lambda}_2$	*MSE*	0.0070	0.0067	0.0060	0.0046	0.0042	0.0050
		AB	0.0778	0.0758	0.0742	0.0688	0.0706	0.0588

Table 5.3: Average lengths and coverage probabilities for ML and Bayes intervals for the set of parameters $(\theta_1, \theta_2) = (0.95, 0.87)$ with varying sample size n under uniform monitoring pattern.

			ACI			*BCI*			*HPD*		
(λ_1, λ_2)		$n \rightarrow$	50	150	250	50	150	250	50	150	250
$(0.21, 0.16)$	$\hat{\theta}_1$	*AL*	0.7379	0.4057	0.3091	0.7157	0.4013	0.2921	0.7061	0.3927	0.2860
		CP	0.8947	0.8822	0.9180	0.9974	0.9955	0.9840	0.9974	0.9866	0.9820
	$\hat{\theta}_2$	*AL*	0.7902	0.3725	0.2826	0.7809	0.5889	0.3617	0.7656	0.5766	0.3430
		CP	0.8974	0.8800	0.9259	0.9711	0.9532	0.9600	0.9947	0.9777	0.9220
	$\hat{\lambda}_1$	*AL*	0.4562	0.2121	0.1617	0.4496	0.2004	0.1411	0.4205	0.1811	0.1243
		CP	0.9020	0.9144	0.9000	0.9447	0.9655	0.9600	1.0000	0.9755	0.9560
	$\hat{\lambda}_2$	*AL*	0.6629	0.2524	0.1871	0.6581	0.2037	0.1744	0.6311	0.1976	0.1681
		CP	0.8921	0.8889	0.8920	0.9895	1.0000	0.9980	0.9774	1.0000	0.9840
$(0.35, 0.27)$	$\hat{\theta}_1$	*AL*	0.7441	0.4188	0.3311	0.7364	0.4042	0.3050	0.7126	0.3945	0.2940
		CP	0.8968	0.9500	0.9405	0.9200	0.9406	0.9752	0.9558	0.9704	0.9620
	$\hat{\theta}_2$	*AL*	0.7984	0.4142	0.3314	0.7836	0.4095	0.3002	0.7565	0.3804	0.2945
		CP	0.9779	0.9558	0.9659	0.9779	0.9735	0.9535	0.9874	0.9912	0.9871
	$\hat{\lambda}_1$	*AL*	0.5451	0.2899	0.2265	0.5307	0.2752	0.2036	0.5291	0.2586	0.1971
		CP	0.8937	0.9082	0.9182	0.9674	0.9735	0.9647	0.9605	0.9823	0.9888
	$\hat{\lambda}_2$	*AL*	0.8447	0.3561	0.2619	0.8247	0.3444	0.2540	0.8054	0.3201	0.2453
		CP	0.9268	0.9035	0.9253	0.9968	1.0000	0.9794	0.9856	0.9741	0.9847

Table 5.4: Average lengths and coverage probabilities for ML and Bayes intervals for set of parameters $(\theta_1, \theta_2) = (0.95, 0.87)$ with varying sample size n under exponential monitoring patterns.

			ACI			*BCI*			*HPD*		
(λ_1, λ_2)			50	150	250	50	150	250	50	150	250
$(0.10, 0.11)$	$\hat{\theta}_1$	*AL*	0.3116	0.3102	0.2210	0.3001	0.2905	0.2011	0.2860	0.2720	0.1904
		CP	0.9067	0.9290	0.9200	0.9667	0.9590	0.9800	0.9577	0.9701	0.9700
	$\hat{\theta}_2$	*AL*	0.5622	0.5462	0.5069	0.5135	0.5025	0.4851	0.5021	0.4814	0.4625
		CP	0.9012	0.9120	0.9195	0.9512	0.9720	0.9895	0.9610	0.9750	0.9652
	$\hat{\lambda}_1$	*AL*	0.4847	0.4641	0.4589	0.4783	0.4443	0.4281	0.4685	0.4341	0.4079
		CP	0.8920	0.8862	0.9045	0.9620	0.9562	0.9745	0.9702	0.9442	0.9852
	$\hat{\lambda}_2$	*AL*	0.5788	0.5679	0.5582	0.5579	0.5378	0.5286	0.5388	0.5219	0.5080
		CP	0.9219	0.9387	0.9247	0.9819	0.9887	0.9747	0.9739	0.9907	0.9957
$(0.12, 0.14)$	$\hat{\theta}_1$	*AL*	0.3316	0.3305	0.2221	0.3230	0.3152	0.2021	0.3010	0.2950	0.1930
		CP	0.9267	0.9167	0.9467	0.9650	0.9526	0.9548	0.9654	0.9547	1.0000
	$\hat{\theta}_2$	*AL*	0.5822	0.5760	0.5100	0.5755	0.5432	0.5065	0.5321	0.5121	0.5013
		CP	0.8800	0.8900	0.9580	0.9654	0.9354	0.9756	0.9500	0.9546	0.9854
	$\hat{\lambda}_1$	*AL*	0.4855	0.4644	0.4599	0.4787	0.4454	0.4387	0.4698	0.4439	0.4158
		CP	0.8747	0.9155	0.9387	0.9933	0.9854	0.9933	0.9800	0.9900	1.0000
	$\hat{\lambda}_2$	*AL*	0.5890	0.5688	0.5670	0.5688	0.5480	0.5309	0.5581	0.5278	0.5098
		CP	0.9333	0.9200	0.9353	0.9755	0.9655	0.9456	0.9600	0.9500	0.9766

From the simulated tables, we observe that values of MSE and AB decrease as sample size increases which validates the consistency of estimators. ALs also decrease with an increase in sample sizes. As we increase the proportion of missing data in the sample, MSE, AB, and AL increases which is obvious as we have more missing data in the sample. The performance of estimators under Bayesian is better than that of classical estimators in terms of MSE and AB. Coverage probabilities are satisfactory for all parameters.

5.6 Concluding Remarks

In this chapter, we have considered the situation two different distributions for different competing causes. The study is done under recall-based competing setup. For point estimation, the E-M algorithm is used. For estimation of confidence intervals, observed Fisher information matrix is calculated using missing information principle. The study is extended to Bayesian paradigm with suitable choice of conjugate priors. The samples from full conditionals are drawn with help of Gibbs Sampling algorithm. An extensive simulation study is done under exponential and uniform monitoring patterns to check the performance of the proposed estimators. The results under all cases found to be satisfactory.

References

Abu-Shawiesh, M. O. A. Adjusted confidence interval for the population median of the exponential distribution. *Journal of Modern Applied Statistical Methods* **9(2)**: 13 (2010).

Aksglaede, L., Sørensen, K., Petersen, J. H., Skakkebæk, N. E. and Juul, A. Recent decline in age at breast development: The Copenhagen puberty study. *Pediatrics* **123(5)**: e932–e939 (2009).

Allison, P. D. Discrete-time methods for the analysis of event histories. *Sociological Methodology* **13**: 61–98 (1982).

Andersen, P. K., Borgan, O., Gill, R. D. and Keiding, N. *Statistical Models Based on Counting Processes*. Springer Science & Business Media (2012).

Ashour, S. and Nassar, M. Inference for Weibull distribution under adaptive type-I progressive hybrid censored competing risks data. *Communications in Statistics-Theory and Methods* **46(10)**: 4756–4773 (2017).

Baghestani, A. R. and Hosseini-Baharanchi, F. S. An improper form of Weibull distribution for competing risks analysis with Bayesian approach. *Journal of Applied Statistics* **46(13)**: 2409–2417 (2019).

Balakrishnan, N. Progressive censoring methodology: An appraisal. *Test* **16(2)**: 211 (2007).

Balakrishnan, N. and Aggrawalla, R. *Progressive Censoring: Theory, Methods and Applications*. Birkhauser, Boston (2000).

Balakrishnan, N., So, H. Y. and Ling, M. H. EM algorithm for one-shot device testing with

competing risks under Weibull distribution. *IEEE Transactions on Reliability* **65(2)**: 973–991 (2015).

Banerjee, A. and Kundu, D. Inference based on type-II hybrid censored data from a Weibull distribution. *IEEE Transactions on Reliability* **57(2)**: 369–378 (2008).

Banerjee, M. and Wellner, J. A. Confidence intervals for current status data. *Scandinavian Journal of Statistics* **32(3)**: 405–424 (2005).

Banik, P., Jaiswal, J., Koley, T. and Chatterjee, I. Explanatory analysis of tertiary audiological data. *Asia Pacific J. Res* **1**: 33–45 (2018).

Barlow, R. E. and Campo, R. Total time on test processes and applications to failure data analysis. In *Reliability and Fault Tree Analysis* (1975).

Bergsten-Brucefors, A. A note on the accuracy of recalled age at menarche. *Annals of Human Biology* **3(1)**: 71–73 (1976).

Bousquet, N., Bertholon, H. and Celeux, G. An alternative competing risk model to the Weibull distribution for modelling aging in lifetime data analysis. *Lifetime Data Analysis* **12(4)**: 481 (2006).

Chacko, M. and Mohan, R. Bayesian analysis of Weibull distribution based on progressive type-ii censored competing risks data with binomial removals. *Computational Statistics* **34(1)**: 233–252 (2019).

Chen, D.-G. D., Sun, J. and Peace, K. E. *Interval Censored Time-to-Event Data: Methods and Applications*. CRC Press (2012).

Chen, M.-H. and Shao, Q.-M. Monte Carlo estimation of Bayesian credible and HPD intervals. *Journal of Computational and Graphical Statistics* **8(1)**: 69–92 (1999).

Childs, A., Chandrasekar, B., Balakrishnan, N. and Kundu, D. Exact likelihood inference based on type-I and type-II hybrid censored samples from the exponential distribution. *Annals of the Institute of Statistical Mathematics* **55(2)**: 319–330 (2003).

Chumlea, W. C., Schubert, C. M., Roche, A. F., Kulin, H. E., Lee, P. A., Himes, J. H. and Sun, S. S. Age at menarche and racial comparisons in US girls. *Pediatrics* **111(1)**: 110–113 (2003).

Clements, M., Mitchell, E., Wright, S., Esmail, A., Jones, D. and Ford, R. Influences on breastfeeding in southeast england. *Acta paediatrica* **86(1)**: 51–56 (1997).

Cohen, A. C. Progressively censored samples in life testing. *Technometrics* **5(3)**: 327–339 (1963).

Cooney, M. A., Louis, G. M. B., Sundaram, R., McGuiness, B. M. and Lynch, C. D. Validity of self-reported time to pregnancy. *Epidemiology (Cambridge, Mass.)* **20(1)**: 56 (2009).

Crowder, M. J. *Classical Competing Risks*. CRC Press. (2001).

David, H. A. and Moeschberger, M. L. *The Theory of Competing Risks: HA David, ML Moeschberger*. C. Griffin. (1978).

De Gruttola, V. and Lagakos, S. W. Analysis of doubly-censored survival data, with application to AIDS. *Biometrics* pp. 1–11 (1989).

Dempster, A. P., Laird, N. M. and Rubin, D. B. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)* **39(1)**: 1–22 (1977).

Deshpande, J. V. and Purohit, S. G. *Lifetime Data: Statistical Models and Methods*, volume 16. World Scientific Publishing Company (2015).

Dewanji, A. and Kalbfleisch, J. Nonparametric methods for survival/sacrifice experiments. *Biometrics* pp. 325–341 (1986).

Dey, S. and Dey, T. Bayesian estimation and prediction intervals for a Rayleigh distribution under a conjugate prior. *Journal of Statistical Computation and Simulation* **82(11)**: 1651–1660 (2012).

Diamond, I. D., McDonald, J. W. and Shah, I. H. Proportional hazards models for current status data: Application to the study of differentials in age at weaning in Pakistan. *Demography* **23(4)**: 607–620 (1986).

Dietz, K., Gail, M., Krickeberg, K., Samet, J. and Tsiatis, A. Statistics for biology and health. *Survival Analysis, Edition Springer* **7** (2002).

Dinse, G. E. and Lagakos, S. Regression analysis of tumour prevalence data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **32(3)**: 236–248 (1983).

Epstein, B. Truncated life tests in the exponential case. *The Annals of Mathematical Statistics* pp. 555–564 (1954).

Fan, T.-H., Wang, Y.-F. and Ju, S.-K. A competing risks model with multiply censored reliability data under multivariate Weibull distributions. *IEEE Transactions on Reliability* **68(2)**: 462–475 (2019).

Finkelstein, D. M. A proportional hazards model for interval-censored failure time data. *Biometrics* pp. 845–854 (1986).

Flehinger, B. J., Reiser, B. and Yashchin, E. Parametric modeling for survival with competing risks and masked failure causes. *Lifetime Data Analysis* **8(2)**: 177–203 (2002a).

Flehinger, B. J., Reiser, B. and Yashchin, E. Parametric modeling for survival with competing risks and masked failure causes. *Lifetime Data Analysis* **8(2)**: 177–203 (2002b).

Friedman, L. and Gertsbakh, I. B. Maximum likelihood estimation in a mininum-type model with exponential and Weibull failure modes. *Journal of the American Statistical Association* **75(370)**: 460–465 (1980).

Gail, M., Krickeberg, K., Samet, J., Tsiatis, A. and Wong, W. *Statistics for Biology and Health.* Springer (2007).

Gehan, E. A. A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. *Biometrika* **52(1-2)**: 203–224 (1965).

Gelman, A. and Rubin, D. B. A single series from the Gibbs sampler provides a false sense of security. *Bayesian Statistics* **4**: 625–631 (1992).

Gillespie, B., d'Arcy, H., Schwartz, K., Bobo, J. K. and Foxman, B. Recall of age of weaning and other breastfeeding variables. *International Breastfeeding Journal* **1(1)**: 4 (2006).

Groeneboom, P. and Wellner, J. A. *Information Bounds and Nonparametric Maximum Likelihood Estimation*, volume 19. Springer Science & Business Media (1992).

Grummer-Strawn, L. M. Regression analysis of current-status data: An application to breast-feeding. *Journal of the American Statistical Association* **88(423)**: 758–765 (1993).

Gupta, A. Estimation of the mean and standard deviation of a normal population from a censored sample. *Biometrika* **39(3/4)**: 260–273 (1952).

Gupta, A., Ranjanb, R. and Upadhyaya, S. K. Classical and Bayes analysis of a competing risk model based on two Weibull distributions with increasing and decreasing hazards. *Reliability: Theory & Applications* **13(3 (50))** (2018).

Haghighi, F. and Castagliola, P. Control chart for monitoring the Weibull shape parameter under two competing risks. *Communications in Statistics-Simulation and Computation* **48(7)**: 2125–2137 (2019).

Hashemi, R. and Amiri, L. Analysis of progressive Type-ii censoring in the Weibull model for competing risks data with binomial removals. *Applied Mathematical Sciences* **5(22)**: 1073–1087 (2011).

Hastings, W. K. Monte Carlo sampling methods using Markov Chains and their applications. *Biometrika* **57(1)**: 97–109 (1970).

Hope, A. C. A simplified Monte Carlo significance test procedure. *Journal of the Royal Statistical Society: Series B (Methodological)* **30(3)**: 582–598 (1968).

Huang, J. and Wellner, J. A. Interval censored survival data: A review of recent progress. In *Proceedings of the First Seattle Symposium in Biostatistics*, pp. 123–169. Springer (1997).

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

Jewell, N. P. and van der Laan, M. Current status data: Review, recent developments and open problems. *Handbook of Statistics* **23**: 625–642 (2003).

Jiang, R. and Murthy, D. Parametric study of competing risk model involving two Weibull distributions. *International Journal of Reliability, Quality and Safety Engineering* **4(01)**: 17–34 (1997).

Keiding, N. Age-specific incidence and prevalence: A statistical perspective. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **154(3)**: 371–396 (1991).

Keiding, N., Begtrup, K., Scheike, T. H. and Hasibeder, G. Estimation from current-status data in continuous time. *Lifetime Data Analysis* **2(2)**: 119–129 (1996).

Klein, J. P. and Moeschberger, M. L. *Survival Analysis: Techniques for Censored and Truncated Data*. Springer Science & Business Media (2006).

Klein, J. P., Van Houwelingen, H. C., Ibrahim, J. G. and Scheike, T. H. *Handbook of Survival Analysis*. CRC Press (2016).

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

Koley, T. and Dewanji, A. Current status data with two competing risks and missing failure types: A non-parametric approach. `https://www.isical.ac.in/~asu/TR/TechRepASU201901.pdf` (2019a).

Koley, T. and Dewanji, A. Revisiting non-parametric maximum likelihood estimation of current status data with competing risks. *Sankhya B* **81(1)**: 39–59 (2019b).

Koley, T. and Dewanji, A. Current status data with two competing risks and missing failure types: Time dependent missing probabilities. `https://asu.isical.ac.in/sites/default/files/TechRepASU202004.pdf` (2020a).

Koley, T. and Dewanji, A. Use of additional information for current status data with two competing risks and missing failure types. `https://asu.isical.ac.in/sites/default/files/TechRepASU202003.pdf` (2020b).

Koley, T. and Dewanji, A. Current status data with two competing risks and missing failure types: A parametric approach. *Journal of Applied Statistics* pp. 1–15 (2021).

Koo, M. and Rohan, T. Accuracy of short-term recall of age at menarche. *Annals of human biology* **24(1)**: 61–64 (1997).

Krailo, M. and Pike, M. Estimation of the distribution of age at natural menopause from prevalence data. *American Journal of Epidemiology* **117(3)**: 356–361 (1983).

Kumar, A., Kumar, S., Kishun, J., Singh, U. and Pushpraj, P. Estimation of the distribution of duration of breastfeeding from cross-sectional data: Some methodological issue. *Journal of Biostatistics and Epidemiology* **6(4)**: 290–299 (2020).

Kundu, D. Bayesian inference and life testing plan for the Weibull distribution in presence of progressive censoring. *Technometrics* **50(2)**: 144–154 (2008).

Kundu, D. and Pradhan, B. Bayesian analysis of progressively censored competing risks data. *Sankhya B* **73(2)**: 276–296 (2011).

Kuo, L. and Yang, T. Y. Bayesian reliability modeling for masked system lifetime data. *Statistics & Probability Letters* **47(3)**: 229–241 (2000).

Lawless, J. F. *Statistical Models and Methods for Lifetime Data*. John Wiley and Sons Inc, New York, 2nd edition (2003).

Lee, E. T. and Wang, J. *Statistical Methods for Survival Data Analysis*, volume 476. John Wiley & Sons (2003).

Li, Z., Zhou, S., Choubey, S. and Sievenpiper, C. Failure event prediction using the cox proportional hazard model driven by frequent failure signatures. *IIE transactions* **39(3)**: 303–315 (2007).

Lin, D. K. and Guess, F. M. System life data analysis with dependent partial knowledge on the exact cause of system failure. *Microelectronics Reliability* **34(3)**: 535–544 (1994).

Liu, F. and Shi, Y. Inference for a simple step-stress model with progressively censored competing risks data from weibull distribution. *Communications in Statistics-Theory and Methods* **46(14)**: 7238–7255 (2017).

Louis, T. A. Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)* **44(2)**: 226–233 (1982).

Maathuis, M. H. *Nonparametric Estimation for Current Status Data with Competing Risks*. University of Washington Seattle (2006).

Maathuis, M. H. and Hudgens, M. G. Nonparametric inference for competing risks current status data with continuous, discrete or grouped observation times. *Biometrika* **98(2)**: 325–340 (2011).

MacMahon, B., Worcester, J. et al. Age at menopause, United States, 1960-1962 (1966).

Mann, N. R., Singpurwalla, N. D. and Schafer, R. E. Methods for statistical analysis of reliability and life data (1974).

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. and Teller, E. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics* **21(6)**: 1087–1092 (1953).

Metropolis, N. and Ulam, S. The Monte Carlo method. *Journal of the American Statistical Association* **44(247)**: 335–341 (1949).

Miller, B. C., Norton, M. C., Curtis, T., Hill, E. J., Schvaneveldt, P. and Young, M. H. The timing of sexual intercourse among adolescents: Family, peer, and other antecedents. *Youth & Society* **29(1)**: 54–83 (1997).

Mirzaei Salehabadi, S., Sengupta, D. and Ghosal, R. Estimating menarcheal age distribution from partially recalled data. *Biostatistics* **21(4)**: 876–894 (2020).

Miyakawa, M. Analysis of incomplete data in a competing risks model. *IEEE Transactions on Reliability* **33**: 293–296 (1984).

Mukhopadhyay, C. and Basu, A. Competing risks with k independent exponentials: A Bayesian analysis. Technical report, Tech. Report, Univ. of Missouri, Columbia (1993).

Nath, D. C., Land, K. C. and Singh, K. K. Birth spacing, breastfeeding, and early child mortality in a traditional Indian society: A hazards model analysis. *Social Biology* **41(3-4)**: 168–180 (1994).

Nelson, W. Graphical analysis of accelerated life test data with the inverse power law model. *IEEE Transactions on Reliability* **21(1)**: 2–11 (1972).

Pareek, B., Kundu, D. and Kumar, S. On progressively censored competing risks data for Weibull distributions. *Computational Statistics & Data Analysis* **53(12)**: 4083–4094 (2009).

Pascual, F. Accelerated life test planning with independent Weibull competing risks with known shape parameter. *IEEE Transactions on Reliability* **56(1)**: 85–93 (2007).

Pascual, F. Accelerated life test planning with independent Weibull competing risks. *IEEE Transactions on Reliability* **57(3)**: 435–444 (2008).

Prentice, R. L., Kalbfleisch, J. D., Peterson Jr, A. V., Flournoy, N., Farewell, V. T. and Breslow, N. E. The analysis of failure times in the presence of competing risks. *Biometrics* **34(4)**: 541–554 (1978).

Rai, H., Panwar, M. and Tomer, S. K. Analysis of masked data with Lindley failure model. *Communications in Statistics-Simulation and Computation* pp. 1–25 (2021).

Ranjan, R., Singh, S. and Upadhyay, S. K. A bayes analysis of a competing risk model based on gamma and exponential failures. *Reliability Engineering & System Safety* **144**: 35–44 (2015).

Ranjan, R. and Upadhyay, S. K. Classical and Bayesian estimation for the parameters of a competing risk model based on minimum of exponential and gamma failures. *IEEE Transactions on Reliability* **65(3)**: 1522–1535 (2016).

Reiser, B., Guttman, I., Lin, D. K., Guess, F. M. and Usher, J. S. Bayesian inference for masked system lifetime data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **44(1)**: 79–90 (1995).

Rossini, A. and Tsiatis, A. A semiparametric proportional odds regression model for the analysis of current status data. *Journal of the American Statistical Association* **91(434)**: 713–721 (1996).

Roy, S., Gijo, E. and Pradhan, B. Inference based on progressive type-I interval censored data from log-normal distribution. *Communications in Statistics-Simulation and Computation* **46(8)**: 6495–6512 (2017).

Salehabadi, S. M. and Sengupta, D. Regression under Cox's model for recall-based time-to-event data in observational studies. *Computational Statistics & Data Analysis* **92**: 134–147 (2015).

Salehabadi, S. M. and Sengupta, D. Nonparametric estimation of time-to-event distribution based on recall data in observational studies. *Lifetime Data Analysis* **22(4)**: 473–503 (2016).

Salehabadi, S. M. and Sengupta, D. Recent advances in the statistical analysis of retrospective time-to-event data. *Advances in Growth Curve and Structural Equation Modeling* pp. 137–150 (2018).

Salehabadi, S. M., Sengupta, D. and Das, R. Parametric estimation of menarcheal age distribution based on recall data. *Scandinavian Journal of Statistics* **42(1)**: 290–305 (2015).

Sarhan, A. M. Estimations of parameters in Pareto reliability model in the presence of masked data. *Reliability Engineering & System Safety* **82(1)**: 75–83 (2003).

Singh, A. K. and Tomer, S. K. Maximum likelihood estimation of component reliability using masked series system lifetime data. *Calcutta Statistical Association Bulletin* **63(1-4)**: 181–192 (2011).

Singh, K., Pandey, N. and Gautam, A. Effect of breastfeeding and maternal health care programme on infant mortality. *Demography India* **36(2)**: 253–266 (2007).

Sinha, S. K. *Reliability and Life Testing*. John Wiley & Sons, Inc. (1986).

Soland, R. M. Bayesian analysis of the Weibull process with unknown scale and shape parameters. *IEEE Transactions on Reliability* **18(4)**: 181–184 (1969).

Sukumaran, A. and Dewan, I. Modelling and analysis of recall-based competing risks data. *Journal of Applied Statistics* **46(9)**: 1621–1635 (2019).

Sun, J. A nonparametric test for current status data with unequal censoring. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **61(1)**: 243–250 (1999).

Sun, J. *The Statistical Analysis of Interval Censored Failure Time Data*. Springer (2007).

Sun, J. and Kalbfleisch, J. D. The analysis of current status data on point processes. *Journal of the American Statistical Association* **88(424)**: 1449–1454 (1993).

Tan, Z. Estimation of exponential component reliability from uncertain life data in series and parallel systems. *Reliability Engineering & System Safety* **92(2)**: 223–230 (2007).

Tan, Z. A new approach to MLE of Weibull distribution with interval data. *Reliability Engineering & System Safety* **94(2)**: 394–403 (2009).

Tomer, S. K., Singh, A. K. and Panwar, M. Bayesian analysis of masked series system lifetime data from a family of lifetime distributions. *International Journal of System Assurance Engineering and Management* **5(4)**: 495–502 (2014).

Trussell, J., Hankinson, R. and Tilton, J. *Demographic Applications of Event History Analysis*. Oxford University Press (1992).

Tsokos, C. P. A Bayesian approach to reliability- Theory and simulation(Bayesian analysis application to reliability and life parameter estimation for Weibull failure model, using Monte Carlo simulation). In *Annual Reliability and Maintainability Symposium, San Francisco, Calif*, pp. 78–87 (1972).

Turnbull, B. W. The empirical distribution function with arbitrarily grouped, censored and truncated data. *Journal of the Royal Statistical Society: Series B (Methodological)* **38(3)**: 290–295 (1976).

Turnbull, B. W. and Weiss, L. A likelihood ratio statistic for testing goodness of fit with randomly censored data. *Biometrics* pp. 367–375 (1978).

Usher, J. S. Weibull component reliability-prediction in the presence of masked data. *IEEE Transactions on Reliability* **45(2)**: 229–232 (1996).

Usher, J. S. and Hodgson, T. J. Maximum likelihood analysis of component reliability using masked system life-test data. *IEEE transactions on reliability* **37(5)**: 550–555 (1988).

Wang, L. Estimation for exponential distribution based on competing risk middle censored data. *Communications in Statistics-Theory and Methods* **45(8)**: 2378–2391 (2016).

Yamane, T. Statistics: An introductory analysis. Technical report (1967).

Zeglen, M., Marini, E., Cabras, S., Kryst, L., Das, R., Chakraborty, A. and Dasgupta, P. The relationship among the age at menarche, anthropometric characteristics, and socio-economic factors in Bengali girls from Kolkata, India. *American Journal of Human Biology* **32(4)**: 23380 (2020).

Zhang, Z., Sun, J. and Sun, L. Statistical analysis of current status data with informative observation times. *Statistics in Medicine* **24(9)**: 1399–1407 (2005).